

B. Bhushan
Editor

NANOSCIENCE AND TECHNOLOGY

Scanning Probe Microscopy in Nanoscience and Nanotechnology

 Springer

NANOSCIENCE AND TECHNOLOGY

NANO SCIENCE AND TECHNOLOGY

Series Editors:

P. Avouris B. Bhushan D. Bimberg K. von Klitzing H. Sakaki R. Wiesendanger

The series NanoScience and Technology is focused on the fascinating nano-world, mesoscopic physics, analysis with atomic resolution, nano and quantum-effect devices, nanomechanics and atomic-scale processes. All the basic aspects and technology-oriented developments in this emerging discipline are covered by comprehensive and timely books. The series constitutes a survey of the relevant special topics, which are presented by leading experts in the field. These books will appeal to researchers, engineers, and advanced students.

Please view available titles in *NanoScience and Technology* on series homepage
<http://www.springer.com/series/3705/>

Bharat Bhushan

Editor

Scanning Probe Microscopy in Nanoscience and Nanotechnology

With 462 Figures

 Springer

Professor Bharat Bhushan

Ohio State University

Nanoprobe Laboratory for Bio- & Nanotechnology and Biomimetics (NLB²)

201 W. 19th Avenue

Columbus, OH 43210-1142, USA

bhushan.2@osu.edu

Series Editors:

Professor Dr. Phaedon Avouris

IBM Research Division

Nanometer Scale Science & Technology

Thomas J. Watson Research Center

P.O. Box 218

Yorktown Heights, NY 10598, USA

Professor Dr. Bharat Bhushan

Ohio State University

Nanoprobe Laboratory

for Bio- & Nanotechnology

and Biomimetics (NLB²)

201 W. 19th Avenue

Columbus, OH 43210-1142, USA

Professor Dr. Dieter Bimberg

TU Berlin, Fakultät Mathematik/

Naturwissenschaften

Institut für Festkörperphysik

Hardenbergstr. 36

10623 Berlin, Germany

Professor Dr., Dres. h.c. Klaus von Klitzing

Max-Planck-Institut

für Festkörperforschung

Heisenbergstr. 1

70569 Stuttgart, Germany

Professor Hiroyuki Sakaki

University of Tokyo

Institute of Industrial Science

4-6-1 Komaba, Meguro-ku

Tokyo 153-8505, Japan

Professor Dr. Roland Wiesendanger

Institut für Angewandte Physik

Universität Hamburg

Jungiusstr. 11

20355 Hamburg, Germany

NanoScience and Technology ISSN 1434-4904

ISBN 978-3-642-03534-0 e-ISBN 978-3-642-03535-7

DOI 10.1007/978-3-642-03535-7

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009934500

© Springer-Verlag Berlin Heidelberg 2010

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: SPI Publisher Services

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Nature is the best example of a system functioning on the nanometer scale, where the materials involved, energy consumption, and data handling are optimized. Opening the doors to the nanoworld, the emergence of the scanning tunneling microscope in 1982 and the atomic force microscope in 1986 led to a shift of paradigm in the understanding and perception of matter at its most fundamental level. As a consequence, new revolutionary concepts stimulated a number of new technologies. The current volume *Scanning Probe Methods in Nanoscience and Nanotechnology* shows that these methods are still making a tremendous impact on many disciplines that range from fundamental physics and chemistry through information technology, spintronics, quantum computing, and molecular electronics, all the way to life sciences. Indeed, over 6,000 AFM-related papers were published in 2008 alone, bringing the total to more than 70,000 since its invention, according to the web of science, and the STM has inspired a total of 20,000 papers. There are also more than 500 patents related to the various forms of scanning probe microscopes. Commercialization of the technology started at the end of the 1980s, and approximately 12,000 commercial systems have been sold so far to customers in areas as diverse as fundamental research, the car industry, and even the fashion industry. There are also a significant number of home-built systems in operation. Some 60–80 companies are involved in manufacturing SPM and related instruments. Indeed, not even the sky seems to be the limit for AFM technology. The Rosetta mission to comet 67P launched by the European Space Agency in 2004 includes an AFM in its MIDAS (Micro-Imaging Dust Analysis System) instrument. The goal of this mission, which is expected to touch down on 67P in 2014, is to analyze particle size distributions in comet material. NASA's Phoenix mission to Mars in 2008 included an AFM for similar studies (collaboration between the Universities of Neuchâtel and Basel, as well as with Nanosurf GmbH).

What does the future hold? Nanotechnology is still dominated to a certain extent by the top down approach where miniaturization plays a crucial role. However, there is a worldwide effort to meet the bottom-up approach of

self-assembly and self-organization that has been so successfully implemented in the natural world. Researchers are trying to unravel nature's secrets on a nanometer scale to create a new generation of materials, devices, and systems that will spectacularly outperform those we have today in information technology, medicine and biology, environmental technologies, the energy industry, and beyond. As we understand better how nature is doing 'things' on a fundamental level, achievements such as clean chemistry or clean processing will emerge along with the ability to handle waste problems and not polluting the environment. New smart materials, hybrid or heterostructured, as well as carbon nanotubes, a variety of nanowires or graphene could be ingredients for novel energy-saving devices. In order to understand the whole functionality of a cell, Systems Biology Institutes have been established with the hope of artificially synthesizing a cell in a bottom-up approach. Nanomedicine, including non-invasive diagnostics, will be more and more on the agenda, fighting diseases on the molecular level, e.g. new kinds of drug delivery systems based on peptides or block co-polymer nanocontainers are being investigated as possible carriers to target carcinogenic cells. Biology is driven by chemistry; however, the scaffold, the gears, the nuts and bolts, e.g. in cell membranes, is nanomechanics, a template that nature has orchestrated during eons of evolution, and worthwhile trying to copy and implement in novel nanodevices.

However, to keep this worldwide effort alive, the interdisciplinary structure of nanoscience requires a new breed of scientists educated in all science disciplines with no language barriers, ready to make an impact on all the global challenges ahead where nanotechnology can be applied¹. Scanning Probe Microscopy and related methods will still play an important role in many of these investigations, helping as to capitalize on this fundamental knowledge, beneficial for future technologies and to mankind.

Basel,
August 2009

Christoph Gerber

¹ Nano-Curriculum at the University of Basel, www.nccr.nano.org/NCCR/study

Preface

The emergence and proliferation of proximal probes, in particular tip-based microscopies, has found applications in a large number of fields of scientific and industrial interest. These allow investigations down to the atomic scale. The recent focus on nanotechnology has made probe-based methods indispensable. The present editor has co-edited several volumes on applied scanning probe methods (SPM). The first volume came out in January 2004, the second to fourth volumes in early 2006, the fifth to seventh volumes in late 2006, the eighth to tenth volumes in late 2007, and eleventh to thirteenth volumes in 2009. These volumes have provided a timely comprehensive overview of SPM applications. These volumes also include fundamental developments in SPMs.

The success of *Applied Scanning Probe Methods I–XIII* (published in the Springer Series *Nanoscience and Technology*) and the rapidly expanding activities in scanning probe development and applications worldwide make it a natural step to collect further specific results in the fields of development of scanning probe microscopy techniques, characterization, and industrial applications, particularly in nanoscience, nanotechnology, and biomimetics. This volume provides insight into the recent work of leading specialists in their respective fields. The focus in this volume is on the fundamental developments in SPM techniques.

The volume introduces many technical concepts and improvements in existing scanning probe techniques and covers a broad and impressive spectrum of recent SPM development and application in many fields of technology, biology, and medicine. The chapters are broken down under three major headings: *Scanning Probe Microscopy Techniques*, *Characterization*, and *Industrial Applications*. After introducing new developments in scanning probe microscopy, including sensor technology and cantilever calibration techniques, characterization data in various applications of scientific and technological interest is presented. Next, chapters on various industrial applications are presented. Characterization data and industrial applications include studies of biological materials, nanostructures, and nanotubes.

The chapters are written by leading researchers and application scientists from all over the world and from various industries to provide a broad perspective. The field is progressing so fast that there is a need for a set of volumes every 12 to 18 months to capture the latest developments.

We gratefully acknowledge the support of all the authors, who are representing leading scientists in academia and industry, for their highly valuable contribution to this volume. We also cordially thank the series editor Claus Ascheron and his staff member Adelheid Duhm from Springer for their continued support during the publication process.

We sincerely hope that readers will find this volume to be scientifically stimulating and rewarding.

Columbus (Ohio),
August 2009

Bharat Bhushan

Contents

Part I Scanning Probe Microscopy Techniques

1 Dynamic Force Microscopy and Spectroscopy Using the Frequency-Modulation Technique in Air and Liquids

Hendrik Hölscher, Daniel Ebeling, Jan-Erik Schmutz, Marcus

<i>M. Schäfer, and Boris Anczykowski</i>	3
1.1 Introduction	3
1.2 Basic Principles of the FM Technique	4
1.2.1 The Equation of Motion	4
1.2.2 Oscillation Behavior of a Self-Driven Cantilever	6
1.2.3 Theory of FM Mode Including Tip-Sample Forces	7
1.2.4 Measuring the Tip-Sample Interaction Force	9
1.2.5 Experimental Comparison of the FM Mode with the Conventional Amplitude-Modulation-mode in Air	11
1.3 Mapping of the Tip-Sample Interactions on DPPC Monolayers in Ambient Conditions	12
1.4 Force Spectroscopy of Single Dextran Monomers in Liquid	15
1.5 Summary	18
References	19

2 Photonic Force Microscopy: From Femtonewton Force Sensing to Ultra-Sensitive Spectroscopy

O.M. Maragò, P.G. Gucciardi, and P.H. Jones

2.1 Introduction	24
2.2 Principles of Optical Trapping	24
2.2.1 Theoretical Background	24
2.3 Experimental Implementation	29
2.3.1 Optical Tweezers Set-up	29
2.3.2 Brownian Motion and Force Sensing	31
2.3.3 Optical Trapping of Linear Nanostructures	33

2.4	Photonic Force Microscopy	39
2.4.1	Bio-Nano-Imaging	39
2.4.2	Bio-Force Sensing at the Nanoscale	42
2.5	Raman Tweezers	45
2.5.1	The Raman Effect	45
2.5.2	Experimental Configuration	46
2.5.3	Applications	48
2.6	Conclusions	53
	References	53

3 Polarization-Sensitive Tip-Enhanced Raman Scattering

	<i>Pietro Giuseppe Gucciardi, Marc Lamy de La Chapelle, Jean-Christophe Valmalette, Gennaro Picardi, and Razvigor Ossikovski</i>	57
3.1	Introduction	57
3.2	Tip-Enhanced Raman Spectroscopy	58
3.2.1	Concept and Advantages	58
3.2.2	Experimental Implementations of TERS with Side Illumination Optics	60
3.2.3	Probes for Tip-Enhanced Raman Spectroscopy	61
3.3	Polarized Raman Scattering from Cubic Crystals	64
3.3.1	Model for Backscattering Raman Emission in c-Silicon . . .	64
3.3.2	Selection Rules	67
3.4	Tip-Enhanced Field Modeling	67
3.4.1	Phenomenological Model	67
3.4.2	Numerical Models and Results	70
3.5	Depolarization of Light Scattered by Metallic Tips	73
3.6	Polarized Tip-Enhanced Raman Spectroscopy of Silicon Crystals .	75
3.6.1	Background Suppression	75
3.6.2	Selective Enhancement of the Raman Modes Induced by Depolarization	80
3.6.3	Evaluation of the Field Enhancement Factor	84
3.7	Conclusions	85
	References	86

4 Electrostatic Force Microscopy and Kelvin Force Microscopy as a Probe of the Electrostatic and Electronic Properties of Carbon Nanotubes

	<i>Thierry Mélin, Mariusz Zdrojek, and David Brunel</i>	89
4.1	Introduction	89
4.2	Electrostatic Measurements at the Nanometer Scale	90
4.2.1	Electrostatic Force Microscopy	90
4.2.2	Kelvin Force Microscopy	93
4.2.3	Lateral Resolution in EFM and KFM	94

4.3 Electrostatic Imaging of Carbon Nanotubes 97

4.3.1 Capacitive Imaging of Carbon Nanotubes
in Insulating Layers 98

4.3.2 EFM Imaging of Carbon Nanotubes and DNA 100

4.3.3 Imaging of Native Charges in Carbon Nanotube Loops . . . 102

4.4 Charge Injection Experiments in Carbon Nanotubes 103

4.4.1 Charge Injection and Detection Techniques 103

4.4.2 Experimental Illustration of EFM Signals 104

4.4.3 Inner-Shell Charging of CNTs 112

4.4.4 Electrostatic Interactions in SWCNTs 115

4.5 Probing the Band Structure of Nanotubes on Insulators 116

4.5.1 Imaging the Semiconductor/Metal Character of
Carbon Nanotubes 116

4.5.2 Imaging the Density of States of Carbon Nanotubes 118

4.6 KFM Studies of Nanotube Devices 119

4.6.1 Charge Transfers at Nanotube–Metal Interfaces 119

4.6.2 Diffusive and Ballistic Transport in Carbon Nanotubes . . 121

4.6.3 Kelvin Force Microscopy of CNTFETs 121

4.7 Conclusion 125

References 126

**5 Carbon Nanotube Atomic Force Microscopy
with Applications to Biology and Electronics**

*Edward D. de Asis, Jr., You Li, Alex J. Austin, Joseph Leung,
and Cattien V. Nguyen* 129

5.1 Carbon Nanotube Introduction 129

5.2 Carbon Nanotube Synthesis 134

5.3 Fabrication of Carbon Nanotube Atomic Force Microscopy Probes 135

5.3.1 Fabrication of Carbon Nanotube Atomic Force
Microscopy Probes by Gluing 135

5.3.2 Mechanical Attachment in Scanning Electron Microscopy . 135

5.3.3 Fabrication of Carbon Nanotube Atomic Force
Microscopy Probes by In Situ Pick-Up 137

5.3.4 Miscellaneous Methods for Post-Growth Attachment
of Carbon Nanotube to Atomic Force Microscopy Tips . . . 137

5.3.5 Metal Catalyst-Assisted Direct-Growth of Carbon
Nanotube Atomic Force Microscopy Probes 137

5.3.6 Post-Growth Attachment is Currently the Most
Optimal Fabrication Process 139

5.4 Characteristics and Characterization of Carbon Nanotube
Atomic Force Microscopy Tips 141

5.5 Applications of Carbon Nanotube Scanning Probe Microscopy . . 148

5.5.1 Functionalization of Carbon Nanotube Tips
for Chemical Force Microscopy 148

5.5.2 Carbon Nanotube Friction Force Microscopy 152

5.5.3	Carbon Nanotube Electric Force Microscopy	152
5.5.4	Carbon Nanotube Scanning Tunneling Microscopy	154
5.5.5	Carbon Nanotube Magnetic Force Microscopy	155
5.5.6	Carbon Nanotube Scanning Near-Field Optical Microscopy	158
5.5.7	Biological Applications of Carbon Nanotube Atomic Force Microscopy	159
References	165

6 Novel Strategies to Probe the Fluid Properties and Revealing its Hidden Elasticity

<i>Laurence Noirez</i>	169
6.1	Introduction	170
6.2	Basic Theoretical Considerations: Conciliating Simple Liquid Approach to the Viscoelasticity Theory?	172
6.2.1	Simple Liquid Description	172
6.2.2	The Viscoelastic Approach	173
6.3	Conventional Procedure to Determine the Dynamic Properties of Fluids	174
6.3.1	Linear Rheology	174
6.3.2	Non-Linear Rheology	176
6.4	Unpredicted Phenomena and Unsolved Questions: Flow Instabilities, Non-Linearities, Shear Induced Transitions, Extra-Long Relaxation Times, Elasticity in the Liquid State	177
6.5	From Macro to Micro and Nanofluidics	181
6.6	Analysis of the Viscoelasticity Scanning Method	183
6.7	The Question of the Boundary Conditions: Surface Effects, Wetting, and Slippage	186
6.8	Novel Description of Conventional Fluids: from Viscous Liquids, Glass Formers to Entangled Polymers. Experiments in Narrow Gap Geometry: Extracting the Shear Elasticity in Viscous Fluids	187
6.9	Tribology Meets Rheology. Novel Methods for the Determination of Bulk Dynamic Properties of a Soft Solid or a Fluidic Material	188
6.10	Elasticity and Dimensionality in Fluids	192
6.11	General Summary and Perspectives	192
References	194

7 Combining Atomic Force Microscopy and Depth-Sensing Instruments for the Nanometer-Scale Mechanical Characterization of Soft Matter

<i>Davide Tranchida and Stefano Piccarolo</i>	199
7.1	Introduction	199
7.2	Determining Elastic Modulus of Compliant Materials from Nanoindentations	201

7.3	Determining Elastic Modulus of Compliant Materials from Nanoindentations	206
7.4	Modulus Estimate of a Challenging Set of Samples	211
	References	220

8 Static and Dynamic Structural Modeling Analysis of Atomic Force Microscope

	<i>Yin Zhang and Kevin D Murphy</i>	225
8.1	Introduction	226
8.2	Working Principle and Modes	227
8.3	Statics of Atomic Force Microscope Cantilever: Effective Stiffness Approach	230
8.4	Electrostatic, Surface and Residual Stress Influence on the Atomic Force Microscope Initial Deflection	234
8.5	Modeling Tip-Sample Contact	237
8.6	Non-Contact Atomic Force Microscope Dynamics: Damping and Influence of Tip-Surface Interaction	242
8.7	Dynamics of Intermittent Contact	248
8.8	Summary	252
	References	253

9 Experimental Methods for the Calibration of Lateral Forces in Atomic Force Microscopy

	<i>Martin Munz</i>	259
9.1	Introduction	260
9.2	Basic Definitions and Relationships	264
	9.2.1 The Calibration Constants Involved in a Lateral Force Measurement	264
	9.2.2 Basic Relationships Involving the Calibration Constants ..	266
	9.2.3 The Lateral and the Normal Spring Constant of a Rectangular CL	268
	9.2.4 The Case of In-Plane Deformations	270
9.3	Calibration of the Lateral Sensitivity of the PSD	271
	9.3.1 Available Methods	271
	9.3.2 Optical Crosstalk	275
9.4	Methods Relying on a Scanning Motion	277
	9.4.1 The Wedge Method	277
	9.4.2 Methods Involving the Normal Spring Constant	283
9.5	Methods Relying on a Force Balance upon Contact with a Rigid Structure	284
	9.5.1 Normal Loading upon Contact with a Sloped Substrate ..	284
	9.5.2 Normal Loading with the Contact Point off the CL Long Axis	286
	9.5.3 Lateral Loading of a Horizontal Surface	288
	9.5.4 Lateral Loading of a Vertical Surface	290
	9.5.5 Mechanical Crosstalk	290

9.6	Methods Relying on a Force Balance Upon Contact with a Compliant Structure	294
9.6.1	The Case of a Vertical Reference Beam	294
9.6.2	The Case of a Horizontal Reference Beam	298
9.6.3	The Case of a Mechanically Suspended Platform	299
9.6.4	The Case of a Magnetically Suspended Platform	302
9.7	Methods Relying on Torsional Resonances of the CL	304
9.8	Discussion	306
9.9	Concluding Remarks	318
	References	319

Part II Characterization

10 Simultaneous Topography and Recognition Imaging

	<i>A. Ebner, L.A. Chtcheglova, J. Preiner, J. Tang, L. Wildling, H.J. Gruber, and P. Hinterdorfer</i>	325
10.1	Introduction	326
10.2	AFM Tip Chemistry	328
10.3	Operating Principles of TREC	331
10.3.1	Half-Amplitude Versus Full-Amplitude Feedback	334
10.3.2	Adjusting the Amplitude	337
10.3.3	Adjusting the Driving Frequency	340
10.3.4	Proofing the Specificity of the Detected Interactions	342
10.4	Applications of TREC: Single Proteins, Membranes, and Cells	344
10.4.1	Antibiotin Antibodies Adsorbed to an Organic Semiconductor	344
10.4.2	Bacterial S-Layer Lattices	346
10.4.3	RBC Membranes	349
10.4.4	Cells	351
10.5	Conclusion	357
	References	357

**11 Structural and Mechanical Mechanisms of Ocular
Tissues Probed by AFM**

	<i>Noël M. Ziebarth, Felix Rico, and Vincent T. Moy</i>	363
11.1	Introduction	363
11.2	Atomic Force Microscopy	364
11.2.1	Principle of Operation	364
11.2.2	Instrumentation	367
11.2.3	Mechanical Measurements	368
11.3	Atomic Force Microscopy in Ophthalmology	370
11.3.1	Cornea	370
11.3.2	Contact Lenses	374

11.3.3 Lens	376
11.3.4 Retinal Tissue	381
11.4 Summary and Conclusions	383
References	383

12 Force-Extension and Force-Clamp AFM Spectroscopies in Investigating Mechanochemical Reactions and Mechanical Properties of Single Biomolecules

<i>Robert Szoszkiewicz</i>	395
12.1 Introduction	396
12.2 Experimental Techniques for Measuring Displacements and Forces at the Single Molecule Level	397
12.2.1 Centroid Tracking	397
12.2.2 Fluorescence Resonance Energy Transfer	398
12.2.3 Magnetic Tweezers	399
12.2.4 Optical Traps	399
12.2.5 Single Molecule AFM Force Spectroscopy	401
12.3 Displacement and Force as Control Parameters in Small Systems ..	402
12.3.1 Displacement Sensitivity and Resolution	402
12.3.2 Force Sensitivity and Resolution	404
12.4 AFM Force Spectroscopy with a Few Piconewton Sensitivity and at a Single Molecule Level	404
12.4.1 Fingerprinting the Biomolecules	405
12.4.2 Optimizing the AFM System	405
12.5 FX-AFM Probes Mechanical Stability of Proteins and Polysaccharides	406
12.5.1 Details of the FX Trace	406
12.5.2 What can be Inferred from the FX Trace?	407
12.5.3 Applications of FX Force Spectroscopy	408
12.6 FC-AFM Probes the Details of Protein (Un)folding and Force-Induced Disulfide Reductions in Proteins	409
12.6.1 Details of the FC Trace	409
12.6.2 What can be Inferred from the FC Trace?	410
12.6.3 Applications of the FC Spectroscopy	412
12.7 Some Shortcomings of the FX/FC-AFM Spectroscopies	413
References	414

13 Multilevel Experimental and Modelling Techniques for Bioartificial Scaffolds and Matrices

<i>F. Consolo, F. Mastrangelo, G. Ciardelli, F.M. Montevecchi, U. Morbiducci, M. Sassi, and C. Bignardi</i>	425
13.1 Scaffolds for Tissue-Engineering Applications	426
13.2 Multi-Scale Computer-Aided Approach in Designing and Modelling Scaffold for Tissue Regeneration	429
13.2.1 CATE: Computer-Aided Anatomical Tissue Representation, CT and MRI Techniques	429

13.2.2	CATE: From Computer-Aided Anatomic 3D Reconstruction to Scaffolds Modelling and Design	432
13.2.3	CATE: FEM and CFD-Based Scaffolds Modelling and Design Methods	448
13.3	Understanding the Cell and Tissue Mechanics: A Multi-Scale Approach	464
13.4	Experimental Techniques for Scaffolds Characterisations	468
	References	476

14 Quantized Mechanics of Nanotubes and Bundles

	<i>Nicola M. Pugno</i>	487
14.1	Introduction	487
14.2	Quantized Fracture Mechanics Approaches	488
14.3	Fracture Strength	492
14.4	Impact Strength	493
14.5	Hyper-Elasticity, Elastic-Plasticity, Fractal Cracks, and Finite Domains	494
14.6	Fatigue Life	494
14.7	Elasticity	495
14.8	Atomistic Simulations	496
14.9	Nanotensile Tests	499
14.10	Thermodynamic Limit	502
14.11	Hierarchical Simulations and Size Effects: from a Nanotube to a Megacable	503
14.12	Conclusions	505
	References	505

15 Spin and Charge Pairing Instabilities in Nanoclusters and Nanomaterials

	<i>Armen N. Kocharian, Gayanath W. Fernando, and Chi Yang</i>	507
15.1	From Atoms to Solids	507
15.1.1	Discreteness of Spectrum	509
15.1.2	Electron Spectroscopy	510
15.1.3	Electron Correlations in Clusters	511
15.2	Transition Metal Oxides	513
15.2.1	Spin-Charge Separation	513
15.2.2	BCS Versus High T_c Superconductivity	518
15.2.3	Localized Versus Itinerant Behavior	520
15.3	Scanning Tunneling Experiments	521
15.3.1	Pseudogap and Gap	521
15.3.2	Two Energy (Temperature) Scales	523
15.3.3	Coherent Versus Incoherent Condensation	524
15.3.4	Modulated Pairs in Cuprates	526
15.3.5	Inhomogeneities	526
15.4	Bethe-Ansatz and GSCF Theories	528

15.5	Hubbard Model	529
15.5.1	GSCF Decoupling Scheme	529
15.5.2	Canonical Transformation	531
15.5.3	Order Parameter $\Delta_q^{(+)}$	532
15.5.4	Quasi-Particle Spectrum	534
15.5.5	Chemical Potential	535
15.5.6	Ground State Phase Diagram	537
15.5.7	GSCF Phase Diagram at $T \neq 0$	539
15.6	Bottom up Approach	540
15.6.1	The Cluster Formalism	542
15.7	General Methodology	542
15.7.1	The Canonical Charge and Spin Gaps	543
15.7.2	Quantum Critical Points: Level Crossings	545
15.7.3	Symmetry Breaking	546
15.7.4	The Charge and Spin Instabilities	548
15.7.5	The Charge and Spin Susceptibility Peaks	550
15.7.6	Charge and Spin Inhomogeneities	551
15.7.7	The Coherent Charge and Spin Pairings	553
15.8	Ground State Properties	554
15.8.1	Bipartite Clusters	554
15.8.2	Tetrahedrons	556
15.8.3	Square Pyramids	559
15.9	Phase T- μ Diagram	560
15.9.1	Tetrahedrons at $t = 1$	560
15.10	Conclusion	563
	References	565

16 Mechanical Properties of One-Dimensional Nanostructures

	<i>Gheorghe Stan and Robert F. Cook</i>	571
16.1	Introduction	571
16.2	Mechanical Property Measurements of One-Dimensional Nanostructures	572
16.2.1	Electric Field-Induced Mechanical Resonance of One-Dimensional Nanostructures	573
16.2.2	Axial Tensile Loading of One-Dimensional Nanostructures	574
16.2.3	Three-Point Bending Test of Bridge-Suspended One-Dimensional Nanostructures	575
16.2.4	Beam-Bending of One-End-Clamped One-Dimensional Nanostructures	576
16.2.5	Instrumented Indentation of One-Dimensional Nanostructures	577
16.2.6	Contact Modulation AFM-Based Techniques	578

16.3	Contact-Resonance Atomic Force Microscopy	579
16.3.1	Cantilever Dynamics in CR-AFM	580
16.3.2	Contact Mechanics in CR-AFM	583
16.3.3	Precision and Accuracy in CR-AFM Measurements (Dual-Reference Calibration Method for CR-AFM)	585
16.4	Contact-Resonance Atomic Force Microscopy Applied to Elastic Modulus Measurements of 1D Nanostructures	587
16.4.1	Normal Contact Stiffness of the Tip–Nanowire Contact	588
16.4.2	Lateral Contact Stiffness of the Tip–Nanowire Contact	589
16.5	Elastic Moduli of ZnO and Te Nanowires Measured by CR-AFM	590
16.5.1	CR-AFM Measurements on ZnO Nanowires	590
16.5.2	CR-AFM Measurements on Te Nanowires	596
16.6	Surface Effects on the Mechanical Properties of 1D Nanostructures	601
16.7	How Important Are the Mechanical Properties of 1D Nanostructures in Applications?	604
	References	605

17 Colossal Permittivity in Advanced Functional Heterogeneous Materials: The Relevance of the Local Measurements at Submicron Scale

	<i>Patrick Fiorenza, Raffaella Lo Nigro, and Vito Raineri</i>	613
17.1	Introduction	614
17.2	Physical Properties of Heterogeneous Materials	616
17.2.1	Theory of the Dielectric Relaxation: Basic Principles	616
17.2.2	Separation of Charges: Maxwell/Wagner/Sillars Polarization	619
17.2.3	Ultimate Theories on the Dielectric Relaxation	621
17.3	Conventional Macroscopic Techniques	627
17.3.1	Basic Principles	627
17.3.2	Dielectric Spectroscopy	628
17.4	Scanning Probe Microscopy	629
17.4.1	Scanning Tunnelling Microscopy on Giant- κ Materials	629
17.4.2	Kelvin Probe Force Microscopy on Giant- κ Materials	633
17.4.3	SIM on Giant- κ Materials	636
17.5	Summary and Conclusions	643
	References	644

18 Controlling Wear on Nanoscale

	<i>Mario D’Acunto</i>	647
18.1	Introduction	648
18.2	Molecular and Supra-Molecular Features for Basic Wear Mechanism	651
18.2.1	Abrasive Wear Mechanisms for Viscoelastic Materials	666

18.3 Modelling Wear as an Activated Process	669
18.3.1 Self-assembled Monolayers as a Frame for Modelling Wear in Viscoelastic Materials	675
18.4 Conclusions	683
References	684
 19 Contact Potential Difference Techniques as Probing Tools in Tribology and Surface Mapping	
<i>Anatoly Zharin</i>	687
19.1 Introduction	687
19.2 Electron Work Function as a Parameter for Surfaces Characterization	688
19.3 Measurements of Contact Potential Difference	690
19.3.1 Kelvin-Zisman Probe	691
19.3.2 Nonvibrating Probe	692
19.3.3 Ionization Probe	693
19.3.4 Atomic Force Microscope in Kelvin Mode	694
19.4 Typical Electron Work Function Responses	696
19.4.1 Surface Deformation	696
19.4.2 Friction	698
19.4.3 Experimental Examples of Kelvin Technique Application .	701
19.5 Periodic Electron Work Function Changes During Friction	704
19.5.1 Phenomenology	704
19.6 Surface Mapping Examples	713
19.7 Closure	716
References	717

Part III Industrial Applications

**20 Modern Atomic Force Microscopy and Its Application
to the Study of Genome Architecture**

<i>Kunio Takeyasu, Hugo Maruyama, Yuki Suzuki, Kohji Hizume, and Shige H. Yoshimura</i>	723
20.1 Introduction: History of AFM Applications to Biological Macromolecules	724
20.1.1 Nanometer Scale Imaging of DNA–Protein Complexes ...	724
20.1.2 Visualization of Various Biological Macromolecules	725
20.1.3 Challenges Toward Technical Advancement	726
20.2 Trends in Biological AFM	727
20.2.1 Analyses of Biological Macromolecules in Motion	727
20.2.2 Measurement of Pico-Newton Mechanical Forces in Biological Systems	729

20.2.3	Cantilever Modification and Application to Force Measurements	730
20.2.4	Recognition Imaging: Integration of Force Measurements and Imaging	731
20.3	Eukaryotic Genome Architecture	731
20.3.1	Biophysical Properties of DNA and DNA-Binding Proteins	733
20.3.2	Fundamental Structures of Eukaryotic Genomes	735
20.3.3	Chromosome Structure in the Mitotic Phase	738
20.3.4	Chromatin Structure Inside Nuclei	738
20.4	Prokaryotic Genome Architecture	739
20.4.1	Bacterial DNA-Binding Proteins	739
20.4.2	Bacterial Genome Structure and Dynamics	741
20.4.3	Archaeal DNA-Binding Proteins, Genome Structure, and Dynamics	742
20.5	Conclusion/Perspectives	746
	References	746
21 Near-Field Optical Lithography		
	<i>Eugenio Cefali, Salvatore Patané, and Maria Allegrini</i>	757
21.1	Introduction	758
21.2	Lithography: Principles and Materials	758
21.2.1	Photolithography	760
21.2.2	Electron Beam Lithography	762
21.2.3	Ion Beam Lithography	763
21.2.4	Materials	764
21.3	Scanning Near-Field Optical Microscopy and Lithography	767
21.3.1	Aperture and Apertureless SNOM Lithography	771
21.3.2	Near-Field Optical Lithography Achievements on Azo – Polymers	781
21.4	Conclusions	788
	References	788
22 A New AFM-Based Lithography Method: Thermochemical Nanolithography		
	<i>Debin Wang, Robert Szoszkiewicz, Vamsi Kodali, Jennifer Curtis, Seth Marder, and Elisa Riedo</i>	795
22.1	Introduction	796
22.2	Thermochemical Nanolithography	797
22.3	Thermal Unmasking of Chemical Groups on a Polymer	799
22.3.1	Unmasking Carboxylic Acid Groups	799
22.3.2	Unmasking Amines Groups	802
22.4	Two-Step Wettability Modification	802
22.5	Covalent Functionalization and Molecular Recognition	805
	References	809

23 Scanning Probe Alloying Nanolithography

<i>Luohan Peng, Hyungoo Lee, and Hong Liang</i>	813
23.1 Brief Review of Nanolithography	813
23.1.1 Introduction	813
23.1.2 Probe-Based Lithography	815
23.1.3 Probe Materials and Properties	816
23.1.4 Probe Wear	817
23.2 Nanoalloying and Nanocrystallization	819
23.2.1 Background	819
23.2.2 Synthesis of Nanoalloys	819
23.3 Probe-Based Nanoalloying and Nanocrystalizations	820
23.3.1 Background	820
23.3.2 Scanning Probe-Based Alloying Nanolithography	821
References	827

24 Structuring the Surface of Crystallizable Polymers with an AFM Tip

<i>Cvetlin Vasilev, Günter Reiter, Khalil Jradi, Sophie Bistac, and Marjorie Schmitt</i>	833
24.1 Introduction	833
24.2 Experimental Part	836
24.2.1 Characteristics of the Polymers Used	836
24.2.2 Sample Preparation	837
24.2.3 The Employed AFM Working Mode	837
24.3 Melting of Confined, Nanometer-Sized Polymer Crystals	841
24.3.1 Self-Assembly and Non-Periodic Patterns	841
24.3.2 The AFM Set-Up Employed for Local Heating	843
24.3.3 Local Melting of Confined Polymer Crystals	843
24.4 Lowering the Crystal Nucleation Barrier by Deforming Polymer Chains	856
24.4.1 Stretched Chains Resulting from Friction Transfer	856
24.4.2 Stretched Chains Resulting from Rubbing with an AFM Tip	859
24.5 Conclusions: Controlling Polymer Properties at a Molecular Scale	862
References	863

25 Application of Contact Mode AFM to Manufacturing Processes

<i>Michael A. Giordano and Steven R. Schmid</i>	867
25.1 Introduction	867
25.2 Review of Atomic Force Microscope Capabilities Relevant to Manufacturing	869
25.2.1 Evaluation of Mechanical Properties	869
25.2.2 Friction/Lubricant Evaluation	882

25.3 Applications to Metal Forming	886
25.3.1 Evaluation of Lubricants	886
25.3.2 Bulk and Sheet Forming	887
25.3.3 Powder Processing	894
25.4 Abrasive Machining Processes	896
25.4.1 Grinding and Polishing	896
25.4.2 Chemical Mechanical Polishing	897
25.4.3 Miscellaneous Applications	901
25.5 Polymer Processing	902
25.6 Conclusions	904
References	905
26 Scanning Probe Microscopy as a Tool Applied to Agriculture	
<i>Fabio Lima Leite, Alexandra Manzoli, Paulo Sérgio de Paula Her- rmann Jr,</i>	
<i>Oswaldo Novais Oliveira Jr, and Luiz Henrique Capparelli Mattoso</i>	
26.1 Applications of Nanotechnology in Agriculture	915
26.2 Applications of AFM in Agriculture	916
26.2.1 Introduction	916
26.2.2 Some Examples and Results of Agricultural Research	916
26.3 Conclusions and Perspectives	940
References	940
Index	945

Contributors

Mario D'Acunto

NSR, NanoSystems Research
via Cecco di Pietro 5
56123 Pisa, Italy

and

Department of Materials Science
University of Pisa
Pisa, Italy
m.dacunto@ing.unipi.it

Maria Allegrini

Dipartimento di Fisica
"E. Fermi" Università di Pisa
and CNISM
Largo Pontecorvo 3
56127 Pisa, Italy
maria.allegrini@df.unipi.it

Boris Anczykowski

Nano Analytics GmbH
Heisenbergstr. 11
48149 Münster, Germany

Alex J. Austin

Stanford University
Stanford, CA 94305, USA
alexjaustin@gmail.com

C. Bignardi

Department of Mechanics
Politecnico di Torino

Corso Duca degli Abruzzi 24
10129 Torino, Italy
cristina.bignardi@polito.it

Sophie Bistac

Université de Haute Alsace – ISC
12 rue des Frères Lumière
68093 Mulhouse cedex, France
sophie.bistac-brogly@uha.fr

David Brunel

Institute of Electronics
Microelectronics and Nanotechnology
IEMN - CNRS-UMR8520
Avenue Poincaré, BP60069
59652 Villeneuve d'Ascq, France
david.brunel@gmail.com

Eugenio Cefali

Department of Physics
University of Messina
Salita Sperone 31
98166 Messina, Italy
eugcef@libero.it

L.A. Chtcheglova

Institute for Biophysics
University of Linz
Altenbergerstr. 69
4040 Linz, Austria
Lilia.Chtcheglova@jku.at

G. Ciardelli

Department of Mechanics
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino, Italy
gianluca.ciardelli@polito.it

F. Consolo

Department of Mechanics
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino, Italy
filippo.consolo@polito.it

Robert F. Cook

Ceramics Division
National Institute of Standards
and Technology
100 Bureau Drive, Stop 8520
Gaithersburg, MD 20899, USA
robert.cook@nist.gov

Jennifer Curtis

Georgia Institute of Technology
School of Physics
837 South Street
Atlanta, GA 30332-0430, USA
Jennifer.curtis@physics.gatech.edu

Edward de Asis Jr

Department of Electrical Engineering
and Bioengineering
Santa Clara University
500 El Camino Real
Santa Clara, CA 95053, USA
edeasis@scu.edu

Daniel Ebeling

Center for Nanotechnology
(CeN Tech), Heisenbergstr. 11
48149 Münster, Germany

Andreas Ebner

Institute for Biophysics
University of Linz
Altenbergerstr. 69

4040 Linz, Austria
Andreas.Ebner@jku.at

Patrick Fiorenza

CNR-IMM, Stradale Primosole 50
95121 Catania, Italy
patrick.fiorenza@imm.cnr.it

Michael Giordan

Department of Aerospace
and Mechanical Engineering
University of Notre Dame
Notre Dame, IN 46556-5637, USA
mgiordan@nd.edu

Christoph Gerber

National Competence Center
for Research in Nanoscale Science
(NCCR)
Institute for Physics
University of Basel
4056 Basel, Switzerland
Christoph.Gerber@unibas.ch

H.J. Gruber

Institute for Biophysics
University of Linz
Altenbergerstr. 69
4040 Linz, Austria
Hermann.Gruber@jku.at

P.G. Gucciardi

CNR-Istituto per i Processi
Chimico-Fisici sez. Messina
Salita Sperone
C.da Papardo,
98158 Messina, Italy
gucciardi@me.cnr.it

**Paulo Sérgio de Paula
Herrmann Jr**

Embrapa Instrumentação
Agropecuária

Laboratório Nacional de
Nanotecnologia para o Agronegócio
Rua 15 de novembro,
1452 São Carlos
13561-206, Sao Paulo, Brasil
herrmann@cnpdia.embrapa.br

Peter Hinterdorfer

Institute for Biophysics
University of Linz
Altenbergerstr. 69
4040 Linz, Austria
peter.hinterdorfer@jku.at

Kohji Hizume

Laboratory of Plasma Membrane
and Nuclear Signaling
Kyoto University
Graduate School of Biostudies
Sakyo-ku, Yoshida Konoe-cho
Kyoto 606-8501, Japan
hizume@lif.kyoto-u.ac.jp

Hendrik Hölscher

Institute for Microstructure
Technology (IMT)
Karlsruhe Institute for Technology
(KIT)
P.O. box 3640
76021 Karlsruhe, Germany
hendrik.hoelscher@kit.edu

P.H. Jones

University College London
Gower Street,
London WC1E 6BT, UK
philip.jones@ucl.ac.uk

Khalil Jradi

Centre Integre en Pates et Papiers
Universite du Quebec a Trois
Rivieres,
3351 boulevard de Forges, C.P. 500
Trois Rivieres, QC, G9A-5H7,
Canada

Armen N. Kocharian

Department of Physics
and Astronomy
California State University
Los Angeles
5151 State University Dr.
Los Angeles, CA 90032, USA
armen.kocharian@calstatela.edu

Vamsi Kodali

Georgia Institute of Technology
School of Physics
837 State Street
Atlanta, GA 30332-0430, USA
vamsi.kodali@physics.gatech.edu

Marc Lamy de La Chapelle

Laboratoire de Chimie, Structure
et Proprietes de Biomateriaux et
c'Agents Therapeutiques
Université Paris 13, SFR SMBH
74 rue Marcel Cachin
93017 Bobigny, France
marc.lamydelachapelle@univ-
paris13.fr

Hyungoo Lee

M.S. 3123 Texas A&M University
College Station
TX 77843-3123, USA
thanku7@gmail.com

Fabio de Lima Leite

Embrapa Agricultural
Instrumentation
C. P. 741, 13560-970,
Sao Carlos, Sao Paulo, Brazil
fabioleite@ufscar.br
and

Federal University of Sao Carlos
(UFSCar)
Rodovia Joao Leme dos Santos
km 110 (SP 264)
C. P. 3031, 18052-780
Sorocaba, Sao Paulo, Brazil
leung@merlin.arc.nasa.gov

Joseph Leung

NASA Ames Research Center
Moffett Field
CA 94035-1000, USA
leung@merlin.arc.nasa.gov

You Li

Stanford University
Stanford
CA 94305, USA
reasonli@gmail.com

Hong Liang

Mechanical Engineering
Texas A&M University
College Station
TX 77843-3123, USA
hliang@tamu.edu

Alexandra Manzoli

Alan G. MacDiarmid Institute
for Innovation and Business and
National Nanotechnology Laboratory
for Agribusiness (LNNA)
Embrapa Agricultural Instrumentation
C. P. 741, 13560-970
Sao Carlos, Sao Paulo, Brazil
Alexandra@cnpdia.embrapa.br

O. M. Maragò

CNR-Istituto per i Processi
Chimico-Fisici sez. Messina
Salita Sperone, C.da Papardo
98158 Messina, Italy
marago@me.cnr.it

Seth Marder

School of Chemistry
and Biochemistry
Georgia Institute of Technology
901 Atlantic Drive NW
30332-0400, Atlanta

GA 30332-0430, USA

seth.marder@chemistry.gatech.edu

Hugo Maruyama

Laboratory of Plasma Membrane
and Nuclear Signaling
Kyoto University
Graduate School of Biostudies
Yoshia-Konoe-cho, Sakyo-ku
Kyoto 606-8501, Japan

F. Mastrangelo

Department of Mechanics
Politecnico di Torino Corso Duca
degli Abruzzi 24
10129 Torino, Italy
francesco.mastrangelo@polito.it

Luiz H.C. Mattoso

Chefe de P&D, Embrapa
Instrumentação Agropecuária
Laboratório Nacional de
Nanotecnologia para o Agronegócio
Rua 15 de novembro
1452 São Carlos,
13561-206, Sao Paulo, Brasil
mattoso@cnpdia.embrapa.br

Thierry Mélin

Institute of Electronics
Microelectronics and Nanotechnology
IEMN-CNRS UMR8520
Avenue Poincaré BP 60069
59652 Villeneuve d'Ascq Cedex
France
thierry.melin@isen.iemn.
univ-lille1.fr

F.M. Montevocchi

Department of Mechanics
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino, Italy
franco.montevocchi@polito.it

U. Morbiducci

Department of Mechanics
 Politecnico di Torino Corso Duca
 degli Abruzzi 24
 10129 Torino, Italy
 umberto.morbiducci@polito.it

Vincent T. Moy

University of Miami
 Department of Physiology
 and Biophysics
 1600 NW 12th Ave.
 RMSB 5077, Miami
 FL 33136, USA
 vmoy@miami.edu

Martin Munz

Analytical Science Team
 Quality of Life Division
 National Physical Laboratory
 Teddington
 Middlesex TW11 0LW, UK
 martin.munz@npl.co.uk

Kevin Murphy

Department of Mechanical
 Engineering
 The University of Connecticut Storrs
 CT 06269, USA
 kdm@engr.uconn.edu

Cattien V. Nguyen

ELORET Corporation
 NASA Ames Research Center
 M/S 229-1, Moffett Field
 CA 94035-1000, USA
 Cattien.V.Nguyen@nasa.gov

Raffaella Lo Nigro

CNR-IMM
 Stradale Primosole 50
 95121 Catania, Italy
 raffaella.lonigro@imm.cnr.it

Laurence Noirez

Laboratoire Léon Brillouin
 (CEA-CNRS)
 CE-Saclay, 91191 Gif-sur-Yvette
 Cédex, France
 laurence.noirez@cea.fr

Oswaldo Novais de Oliveira Jr

Institute of Physics of Sao Carlos
 (IFSC)
 University of Sao Paulo
 PO Box 369, CEP: 13560-970
 Sao Carlos, Sao Paulo, Brazil
 chu@ifsc.usp.br

Razvigor Ossikovski

LPICM, Ecole Polytechnique
 CNRS, 91128 Palaiseau, France
 ossikovs@poly.polytechnique.fr

Salvatore Patanè

Dipartimento di Fisica della Materia
 e Ingegneria Elettronica
 Università di Messina
 Salita Sperone 31
 98100 Messina, Italy
 salvatore.patane@unime.it

Luohan Peng

M.S. 3123, Texas A&M University
 College Station
 TX 77843-3123, USA
 pengluohan@gmail.com

G. Picardi

LPICM, Ecole Polytechnique
 CNRS, 91128 Palaiseau, France
 gennaro.picardi@polytechnique.edu

Stefano Piccarolo

Dipartimento di Ingegneria Chimica
 dei Processi e dei Materiali
 Università di Palermo
 Viale delle Scienze
 90128 Palermo, Italy
 piccarolo@unipa.it

J. Preiner

Institute for Biophysics
University of Linz
Altenbergerstr. 69
4040 Linz, Austria
Johannes.Preiner@jku.at

Nicola Pugno

Department of Structural
Engineering
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino, Italy
nicola.pugno@polito.it

Vito Raineri

CNR-IMM, Stradale Primosele 50
95121 Catania, Italy
vito.raineri@imm.cnr.it

Günter Reiter

Physikalisches Institut
Albert-Ludwigs-Universität
Hermann-Herder-Str. 3,
79104 Freiburg, Germany
guenter.reiter@physik.uni-
freiburg.de

Félix Rico

Department of Physiology
and Biophysics
University of Miami Miller School
of Medicine
1638 NW 10th Avenue, Room 117A
Miami, FL 33136, USA
frico@med.miami.edu

Elisa Riedo

School of Physics
Georgia Institute of Technology
837 State Street
Atlanta, GA 30332-0430, USA
elisa.riedo@physics.gatech.edu

Mariapaola Sassi

Thermodynamic Division
Istituto di Ricerca Metrologica
(INRIM)
Strada delle Cacce 73
10135 Torino, Italy
m.sassi@inrim.it

Marcus M. Schaefer

nanoAnalytics GmbH
Heisenbergstr. 11
48149 Münster, Germany

Steven Schmid

Department of Aerospace
and Mechanical Engineering
University of Notre Dame
Notre Dame
IN 46556-5637, USA
schmid.2@nd.edu

Marjorie Schmitt

Université de Haute Alsace – ISC
12 rue des Frères Lumière
68093 Mulhouse cedex, France
marjorie.schmitt@uha.fr

Jan-Erik Schmutz

Center for Nanotechnology
(CeN Tech)
Heisenbergstr. 11
48149 Münster, Germany

Gheorghe Stan

Ceramics Division
National Institute of Standards
and Technology
100 Bureau Drive
Stop 8526, Gaithersburg
MD 20899, USA
gheorghe.stan@nist.gov

Yu-ki Suzuki

Laboratory of Plasma Membrane
and Nuclear Signaling
Kyoto University Graduate
School of Biostudies, Sakyo-ku
Yoshida Konoe-cho
Kyoto 606-8501, Japan
ysuzuki79.m07@lif.kyoto-u.ac.jp

Robert Szoszkiewicz

Department of Physics
Kansas State University
Manhattan
KS 66506-2601, USA
rs@phys.ksu.edu

Kunio Takeyasu

Laboratory of Plasma Membrane
and Nuclear Signaling
Kyoto University Graduate
School of Biostudies, Sakyo-ku
Yoshida Konoe-cho
Kyoto 606-8501, Japan
takeyasu@lif.kyoto-u.ac.jp

J. Tang

Institute for Biophysics University
of Linz
Altenbergerstr. 69
4040 Linz, Austria
jilintang@yahoo.com

Davide Tranchida

Department of Physical Chemistry
University of Siegen
Adolf-Reichwein-Str. 2
57076 Siegen, Germany
tranchida@chemie.uni-siegen.de

J.C. Valmalette

IM2NP CNRS-UMR 6242
Université du Sud Toulon Var
BP 20132 83957 La Garde
France
valmalette@univ-tln.fr

Cvetelin Vasilev

Department of Physics
and Astronomy
The University of Sheffield
Hounsfield Road
Sheffield S3 7RH, UK
c.vasilev@sheffield.ac.uk

Debin Wang

School of Physics
Georgia Institute of Technology
837 State Street
Atlanta, GA 30332-0430, USA
db_wang@gatech.edu

L. Wildling

Institute for Biophysics
University of Linz
Altenbergerstr. 69
4040 Linz, Austria
Linda.Wildling@jku.at

Shige H. Yoshimura

Laboratory of Plasma Membrane
and Nuclear Signaling
Kyoto University Graduate
School of Biostudies
Sakyo-ku Yoshida Konoe-cho
Kyoto 606-8501, Japan
yoshimura@lif.kyoto-u.ac.jp

Mariusz Zdrojek

Faculty of Physics
Warsaw University of Technology
Koszykowa 75
00-662 Warsaw, Poland
and
CIN2-ICN Barcelona, Spain
zdrojek@if.pw.edu.pl

Yin Zhang

State Key Laboratory
of Nonlinear Mechanics

Institute of Mechanics
Chinese Academy of Sciences
Beijing 1000190, China
zhangyin@lnm.imech.ac.cn

Anatoly Zharin
Belarusian National Technical
University
65 F. Skoriny Ave.
bld. 17.220013, Minsk, Belarus
anatoly.zharin@gmail.com

Noel Ziebarth
Department of Biomedical
Engineering
University of Miami
College of Engineering
1251 Memorial Drive
McArthur Annex Room 209
Coral Gables
FL 33146, USA
nziebarth@miami.edu

Part I Scanning Probe Microscopy Techniques

Dynamic Force Microscopy and Spectroscopy Using the Frequency-Modulation Technique in Air and Liquids

Hendrik Hölscher, Daniel Ebeling, Jan-Erik Schmutz, Marcus M. Schäfer, and Boris Anczykowski

Summary. The frequency-modulation (FM) mode was introduced in 1991 to increase the sensitivity of dynamic force microscopy in vacuum. However, it is also possible to use this technique in air and liquids which has several advantages compared with the conventional amplitude-modulation (AM) (“tapping”) mode. In this chapter, we review the fundamentals of the FM mode and analyze its basic theoretical background. Finally, we present experimental results obtained in air and liquids and compare them with the conventional AM technique.

Key words: dynamic force spectroscopy, frequency-modulation technique, chain-like molecules.

1.1 Introduction

Atomic force microscopy (AFM) [1] has been established as a standard tool in nano- and biotechnology for the imaging of surfaces. Beside this, it is possible to measure the tip-sample forces with high accuracy enabling the measurement of intermolecular [2] and intramolecular forces [3]. These results are typically obtained in the so-called contact mode of the AFM wherein the quasi-static bending of the cantilever is detected.

Another option to image surfaces is the use of the dynamic mode where the cantilever oscillates near the sample surface. Typically, the cantilever is vibrated at a fixed frequency and the oscillation characteristics of the cantilever – like amplitude and phase – change in dependence of the actual cantilever-sample distance. This mode is sometimes also called “tapping”-mode and is the standard technique in ambient conditions [4, 5] and liquids [6–8].

Since working under vacuum conditions allows the preparation of very clean surfaces, the resolution can be greatly improved in this environment. However, in vacuum, the cantilever is not damped by a surrounding medium and it exhibits very high Q -factors which results in very low bandwidth. Since

the so-called frequency-modulation (FM) mode [9] is well-suited for these conditions, it is commonly applied in this environment. This technique abandons the external excitation of the tapping-mode and uses the specific behavior of a self-driven oscillator instead [10, 11]. True atomic resolution became possible with this approach for the first time [12, 13].

Later on, this technique was extended to dynamic force spectroscopy (DFS) enabling the precise detection of tip-sample interaction forces in vacuum [14–16]. By using stable microscopes at low temperatures, it is even possible to measure the interaction forces at specific lattice sites [17] or to scan the three dimensional force field between the probing tip and the sample surface with atomic resolution [18, 19]. Compared to conventional force-vs-distance curves, dynamic force spectroscopy has the advantage that the so-called jump-to-contact [20] of the tip towards the sample surface can be prevented [21]. Later on, it has been shown that dynamic force spectroscopy works also under ambient conditions [22, 23] and in liquids [24] enabling the precise measurement of tip-sample forces.

In this chapter, we review the basic principles of the FM mode and focus on its application in ambient conditions and liquids. Experimental applications are presented for the imaging of DPPC monolayers and the stretching of single dextran monomers.

1.2 Basic Principles of the FM Technique

1.2.1 The Equation of Motion

A sketch of the experimental setup of a dynamic force microscope utilizing the FM technique is shown in Fig. 1.1. The deflection of the cantilever is typically measured with the laser beam deflection method as indicated [25, 26], but other displacement sensors such as interferometric sensors [27–29] can be applied as well.

A dynamic force microscope driven in the FM-mode has a feedback circuit consisting of an amplifier and a phase shifter (solid lines in Fig. 1.1). The properly amplified and phase-shifted displacement sensor signal is then used to excite the dither piezo vibrating the cantilever [9]. Two different driving techniques have been established for use with the FM-mode: The original constant-amplitude (CA) driving scheme, where the oscillation amplitude of the cantilever is held constant by the automatic gain control (AGC) [9] and the constant-excitation (CE) driving scheme [30, 31], where the excitation amplitude of the cantilever driving is kept constant. Both the CE driving scheme [22, 32–35] as well as the CA driving scheme [23, 36–38] are frequently used in air and liquids. However, since the amplitude can be used as a feedback signal for scanning in the CE driving scheme, its implementation is easily possible for an existing DFM build for AM-mode applications in air and liquid (cf. Fig. 1.1). Therefore, we focus on this driving scheme in the following.

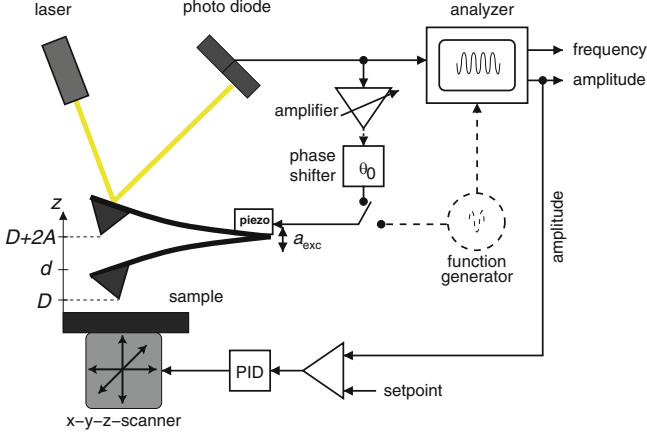


Fig. 1.1. Schematic drawing of the experimental setup of a dynamic force microscope where the cantilever is driven with the frequency-modulation (FM) technique. The FM-mode exhibits a feedback loop consisting of a phase (or time) shifter and an amplifier. For comparison we have plotted also the set-up for the conventional amplitude-modulation (AM) mode where the cantilever is externally driven with a frequency generator producing a fixed frequency (*dashed lines*). In both cases we assume that the laser beam deflection method is used to measure the oscillation of the tip oscillating between the nearest tip-sample position D and the upper turning point $D + 2A$. The equilibrium position of the tip is denoted as d

For comparison, we have also plotted the experimental set-up of the conventional amplitude-modulation (AM) mode [7, 39] – sometimes also named as “tapping”-mode [5] – by dashed lines in Fig. 1.1. In this historically older mode, the cantilever is driven by a sinusoidal signal with a fixed frequency and a constant amplitude originating from a function generator. The resulting oscillation amplitude and the phase shift are detected by a lock-in amplifier. The signal of the function generator is not only used for the cantilever driving; its signal serves also as an input for the lock-in amplifier. The focus of this chapter, however, is the FM-mode with constant-excitation and we discuss its specific features in the following.

Based on the forementioned description of the experimental setup, we can formulate the equation of motion describing the cantilever dynamics in the FM-mode [40]:

$$m\ddot{z}(t) + \frac{2\pi f_0 m}{Q_0} \dot{z}(t) + c_z(z(t) - d) + F_{ts}[z(t), \dot{z}(t)] + \frac{a_{exc}}{A} c_z(z(t - t_0) - d) = 0 \quad (1.1)$$

Here, $z(t)$ is the position of the tip at the time t ; c_z , m , and $f_0 = \sqrt{(c_z/m)/(2\pi)}$ are the spring constant, the effective mass, and the eigenfrequency of the cantilever, respectively. Somewhat simplifying, it is assumed that the quality

factor Q_0 unites the intrinsic damping of the cantilever and all influences from surrounding media such as air or liquid, if present, in a single overall value. The equilibrium position of the tip is denoted as d . The (non-linear) tip-sample interaction force F_{ts} is introduced by the fourth term. The self-excitation of the cantilever used in the FM mode is described by the retarded amplification of the displacement signal, i.e., the tip position z is measured at the retarded time $t - t_0$. Nonetheless, a consideration of the time shift by a phase difference θ_0 is also possible, giving equivalent results. Therefore, we use “time shift” and “phase shift” as synonyms throughout this review and notice that both parameters are scaled by $\theta_0 = 2\pi f_d t_0$.

1.2.2 Oscillation Behavior of a Self-Driven Cantilever

To analyze the specific features of the FM mode, it is instructive to first examine the specific features of the cantilever driving by self-excitation. For simplicity, we assume that the cantilever vibrates far away from the sample surface. Consequently, we can neglect tip-sample forces ($F_{ts} \equiv 0$), resulting in a greatly simplified equation of motion (1.1)

$$m\ddot{z}(t) + \frac{2\pi f_0 m}{Q_0} \dot{z}(t) + c_z(z(t) - d) + \frac{a_{\text{exc}}}{A} c_z(z(t - t_0) - d) = 0. \quad (1.2)$$

As the cantilever is not excited with a specific externally set frequency, the cantilever itself serves as the frequency determining element. Therefore, we make the ansatz [10, 40]

$$z(t \gg 0) = d + A \cos(2\pi f t), \quad (1.3)$$

and introduce it into (1.2). As a result, we obtain a set of two coupled trigonometric equations:

$$\frac{a_{\text{exc}}}{A} \cos(2\pi f t_0) = \frac{f^2 - f_0^2}{f_0^2}, \quad (1.4a)$$

$$\frac{a_{\text{exc}}}{A} \sin(2\pi f t_0) = \frac{1}{Q_0} \frac{f}{f_0}. \quad (1.4b)$$

The two equations can be decoupled with the assumption that the time shift t_0 is set to a value corresponding to $t_0 = 1/(4f_0)$ ($= 90^\circ$), which simultaneously corresponds to the by far most common choice for t_0 . For this value, the solution of (1.4) is given by

$$f = f_0, \quad (1.5a)$$

$$A_0 = a_{\text{exc}} Q_0. \quad (1.5b)$$

This simple calculation demonstrates the very specific behavior of a self-driven oscillator if the phase shift is set to 90° . In this case, the cantilever oscillates

exactly with its eigenfrequency f_0 . Due to this specific feature revealed by (1.5a), we define that the cantilever is in *resonance* if this condition is fulfilled. The linear relationship between the oscillation amplitude A_0 of the freely oscillating cantilever and excitation amplitude is described by (1.5b).

This simple analysis points out the difference to the well-known case of a driven and damped harmonic oscillator. As known from introductory physics courses, such an oscillator can be driven at any frequency which is not necessarily the actual resonance frequency. In the FM mode, however, the cantilever oscillates with its resonance frequency as long as the phase shift is set to 90° .

1.2.3 Theory of FM Mode Including Tip–Sample Forces

The mathematical form of realistic tip–sample forces is highly nonlinear for almost all the cases of tip–sample systems. This fact complicates the analytical solution of the equation of motion (1.1) even if we assume a simplified force model. However, for the analysis of dynamic force microscopy experiments, we need to focus on steady-state solutions of the equation of motion with sinusoidal cantilever oscillation. Therefore, it is advantageous to expand the tip–sample force into a Fourier series

$$\begin{aligned}
 F_{ts}[z(t), \dot{z}(t)] &\approx f \int_0^{1/f} F_{ts}[z(t), \dot{z}(t)] dt \\
 &+ 2f \int_0^{1/f} F_{ts}[z(t), \dot{z}(t)] \cos(2\pi ft) dt \times \cos(2\pi ft) \\
 &+ 2f \int_0^{1/f} F_{ts}[z(t), \dot{z}(t)] \sin(2\pi ft) dt \times \sin(2\pi ft) + \dots,
 \end{aligned}
 \tag{1.6}$$

where $z(t)$ is given by (1.3). In the following, we assume that the tip–sample force is so small and the Q -factor is so high that, as a consequence, higher harmonics can be neglected. It has been shown by Cleveland et al. [41] and Rodriguez and García [42] that this condition is well fulfilled in many practical cases.

The first term in the Fourier series reflects the averaged tip–sample force over one full oscillation cycle, which shifts the equilibrium point of the oscillation by a small offset Δd from d to d_0 . Actual values for Δd , however, are typically small. For amplitudes commonly used in dynamic force microscopy in air and liquids (some nm to some tens of nm), the averaged tip–sample force is in the range of some pN. The resulting offset Δd is less than 1 pm for typical sets of parameters [43]. Since this is well beyond the resolution limit of a typical dynamic force microscopy experiment in air and liquids, we neglect this effect in the following and assume $d \approx d_0$ and $D = d - A$.

Now we focus on the solution of (1.1) for the FM-mode. Assuming a sinusoidal cantilever oscillation and using again the ansatz (1.3) leads to a set of two coupled trigonometric equations

$$\frac{a_{\text{exc}}}{A} \cos(2\pi f t_0) = \frac{f^2 - f_0^2}{f_0^2} - I_+(d, A), \quad (1.7a)$$

$$\frac{a_{\text{exc}}}{A} \sin(2\pi f t_0) = \frac{1}{Q} \frac{f}{f_0} - I_-(d, A), \quad (1.7b)$$

where we defined the following two integrals:

$$\begin{aligned} I_+(d, A) &= \frac{2f_d}{c_z A} \int_0^{1/f} F_{\text{ts}}[z(t), \dot{z}(t)] \cos(2\pi f t) dt \\ &= \frac{1}{\pi c_z A^2} \int_{d-A}^{d+A} (F_{\downarrow} + F_{\uparrow}) \frac{z-d}{\sqrt{A^2 - (z-d)^2}} dz, \end{aligned} \quad (1.8a)$$

$$\begin{aligned} I_-(d, A) &= \frac{2f}{c_z A} \int_0^{1/f} F_{\text{ts}}[z(t), \dot{z}(t)] \sin(2\pi f t) dt \\ &= \frac{1}{\pi c_z A^2} \int_{d-A}^{d+A} (F_{\downarrow} - F_{\uparrow}) dz \\ &= -\frac{1}{\pi c_z A^2} \Delta E(d, A). \end{aligned} \quad (1.8b)$$

Both integrals are functions of the actual oscillation amplitude A and cantilever-sample distance d . Furthermore, they depend on the sum and the difference of the tip-sample forces during approach (F_{\downarrow}) and retraction (F_{\uparrow}) as manifested by the labels “+” and “-” for easy distinction. The integral I_+ is a weighted average of these tip-sample forces ($F_{\text{ts}} = (F_{\downarrow} + F_{\uparrow})/2$). On the other hand, the integral I_- is directly connected to ΔE , which reflects the energy dissipated during an individual oscillation cycle. Consequently, this integral vanishes for purely conservative tip-sample forces, where F_{\downarrow} and F_{\uparrow} are identical. A more detailed discussion of these integrals can be found in [44] and [45].

Both equations can be simplified for the conditions typically found in DFM experiments where the FM-mode is applied. First, we assume that the frequency shift $\Delta f := f - f_0$ caused by the tip-sample interaction and the damping is small compared to the resonance frequency of the free cantilever ($\Rightarrow f/f_0 \approx 1$ and $f^2 - f_0^2 \approx -2\Delta f f_0$). Second, we consider that the phase shift is typically set to 90° in the FM mode. In this case, the terms on the left side are given by $\cos(2\pi f t_0) = 0$ and $\sin(2\pi f t_0) = 1$. Due to these simplifications, the frequency shift and the driving amplitude can be calculated from

$$\Delta f(d, A) = -\frac{f_0}{2} I_+ = -\frac{f_0}{\pi c_z A^2} \int_{d-A}^{d+A} F_{\text{ts}} \frac{z-d}{\sqrt{A^2 - (z-d)^2}} dz, \quad (1.9a)$$

$$a_{\text{exc}} = \frac{A}{Q} + \frac{\Delta E(d, A)}{\pi c_z A}. \quad (1.9b)$$

Both equations are valid for every type of interaction as long as the resulting cantilever oscillation is nearly sinusoidal. Equation (1.9a) coincides with

the well-known result for the FM-mode with constant-oscillation amplitude (see, e.g., [16, 21, 46]), but it is coupled with (1.9b) through the oscillation amplitude.

1.2.4 Measuring the Tip-Sample Interaction Force

In the previous subsection, we have calculated the influence of the tip-sample interaction on the cantilever oscillation with the assumption that the tip-sample force is known. In practice, however, it is just the other way around: the tip-sample interaction is unknown. Therefore, the ability to measure the continuous tip-sample interaction force as a function of the tip-sample distance is of great interest for the microscopist.

We start our analysis by applying the transformation $D = d - A$ to the integral I_+ (1.8a), where D corresponds to the nearest tip-sample distance as defined in Fig. 1.1. Next, we note that due to the cantilever oscillation, the current method intrinsically recovers the values of the force that the tip experiences at its lower turning point, where F_{\downarrow} necessarily equals F_{\uparrow} . We thus define $F_{\text{ts}} = (F_{\downarrow} + F_{\uparrow})/2$, and (1.8a) subsequently reads as

$$I_+ = \frac{2}{\pi c_z A^2} \int_D^{D+2A} F_{\text{ts}} \frac{z - D - A}{\sqrt{A^2 - (z - D - A)^2}} dz. \quad (1.10)$$

In the next step, we assume that the used amplitude is considerably larger than the interaction range of the tip-sample force. Consequently, tip-sample forces in the integration range between $D + A$ and $D + 2A$ are practically insignificant. For this so-called ‘‘large amplitude approximation’’ [16, 21], the last term can be expanded at $z \rightarrow D$ to $(z - D - A)/\sqrt{A^2 - (z - D - A)^2} \approx -\sqrt{A/2(z - D)}$, resulting in

$$I_+ \approx -\frac{\sqrt{2}}{\pi c_z A^{3/2}} \int_D^{D+2A} \frac{F_{\text{ts}}}{\sqrt{z - D}} dz. \quad (1.11)$$

Introducing now this equation into (1.9a), we obtain the following integral equation:

$$\Delta f(D, A) \approx \frac{f_0}{\sqrt{2}\pi c_z A^{3/2}} \int_D^{D+2A} \frac{F_{\text{ts}}(z)}{\sqrt{z - D}} dz. \quad (1.12)$$

The inversion now leads to the following formula for the tip-sample interaction potential [16, 47]

$$V_{\text{ts}}(D) = \sqrt{2} \int_D^{D+2A} \frac{c_z A (z')^{3/2}}{f_0} \frac{\Delta f(z')}{\sqrt{z' - D}} dz'. \quad (1.13)$$

Consequently, the tip-sample force is given by

$$F_{\text{ts}}(D) = -\sqrt{2} \frac{\partial}{\partial D} \int_D^{D+2A} \frac{\Delta\gamma(z')}{\sqrt{z'-D}} dz', \quad (1.14)$$

where we defined the so-called *normalized frequency shift* [21]

$$\gamma(D) := \frac{c_z A^{3/2}}{f_0} \Delta f(D), \quad (1.15)$$

which is independent of the oscillation amplitude, but a function of the nearest tip-sample distance D [47]. To recover the tip-sample interaction force from a spectroscopy experiment, we have to measure the frequency shift and the driving amplitude as a function of the tip-sample distance before we calculate the normalized frequency shift. After that, we introduce this data into (1.14).

For completeness we would like to mention that the “large amplitude approximation” used for the deviation of (1.11) is not a prerequisite for the reconstruction of the tip-sample force. Inversion methods suitable for small amplitudes have been presented in [48–51]. The method of Sader and Jarvis [51] for example can also be applied to our case and adds two additional terms to our previous formula

$$F_{\text{ts}}^{\text{SJ}}(D) = -\sqrt{2} \frac{\partial}{\partial D} \int_D^\infty \frac{c_z A^{3/2}}{f_0} \frac{\Delta f(z)}{\sqrt{z-D}} + \sqrt{2} c_z \frac{\Delta f}{f_0} \left((z-D) + \frac{A^{1/2}}{4} \sqrt{\frac{z-D}{\pi}} \right) dz. \quad (1.16)$$

The calculation of the energy dissipation ΔE is straightforward using (1.9b)

$$\Delta E(D, A) \cong \pi c_z \left(A(D) a_{\text{exc}} - \frac{A(D)^2}{Q} \right). \quad (1.17)$$

It might be interesting to note that this equation follows also from the conservation of energy [41, 52].

Since the quantitative value of excitation amplitude a_{exc} is typically unknown in an experiment, it is convenient to determine a_{exc} from the freely oscillating cantilever with the help of (1.5b). As a result we get

$$\Delta E(D) = \frac{\pi c_z A(D)}{Q} (A_0 - A(D)), \quad (1.18)$$

where A_0 is the oscillation amplitude of the freely oscillating cantilever (i.e., with negligible tip-sample force).

1.2.5 Experimental Comparison of the FM Mode with the Conventional Amplitude-Modulation-mode in Air

The set-up presented in the previous subsections can be easily realized by combining a commercial dynamic force microscope (MultiMode AFM with NanoScope IIIa Controller, Veeco Instruments Inc.) with an additional electronics dedicated for the constant-excitation mode (QFM-Module, nanoAnalytics GmbH). In this way, it is also possible to switch between AM- and FM-mode using the same cantilever and sample. The spring constants c_z of the cantilever can be determined via the resonant frequency f_0 of the freely oscillating cantilever [53] while their quality factors Q can be obtained from resonance curves [54].

To illustrate the main differences between the “conventional” AM-mode and the presently much less used FM-mode in air, we present two spectroscopy experiments in Fig. 1.2, where the oscillating cantilever was approached to and retracted from a mica surface in both modes. The corresponding spectroscopy curves are presented in Fig. 1.2a, b. The measured quantities in the AM-mode are amplitude and phase, whereas amplitude and frequency shift are recorded in FM-mode.

As already discussed by others (see, e.g., [43,55,56] and references therein), the amplitude and phase shift curves recorded in AM-mode sometimes show a significant hysteresis during approach and retraction. At specific positions (marked by arrows in Fig. 1.2a), the oscillation becomes unstable and the cantilever jumps into another stable oscillation state. However, such a hysteresis is not present in the spectroscopy curves measured in the FM-mode due to the specific self-oscillation technique [11]. As shown in Fig. 1.2b, the particular amplitude and frequency shift curves are identical within noise limit and piezo hysteresis for approach and retraction. The amplitude is constant until the tip senses the interaction with the sample surface and decreases continuously during further approach. The frequency shift curves show a decrease and increase of the resonant frequency with a distinct minimum.

As already pointed out the continuous approach and retraction curves obtained in the FM mode allow the reconstruction of the tip-sample interaction by an inversion algorithm. An application of this procedure to the spectroscopy data is plotted in Fig. 1.2c and reveals the tip-sample force as well as the energy dissipation per oscillation cycle.

The tip-sample the force shows a distinct minimum of -7 nN. This is the minimal force needed to retract the tip from the sample surface. Therefore, we denote it as *adhesion force* in the following. During further approach the tip-sample force increases with a slope of 8.5 N m^{-1} as shown by a linear fit (solid line). This linear increase in the tip-sample force is caused by the contact of tip and sample. Justified by the almost linear increase, we use the term *contact stiffness* for the slope obtained by this linear fitting procedure. Interestingly, the energy dissipation curve begins to rise with the onset of the attractive tip-sample force and shows also a linear increase with a slope of

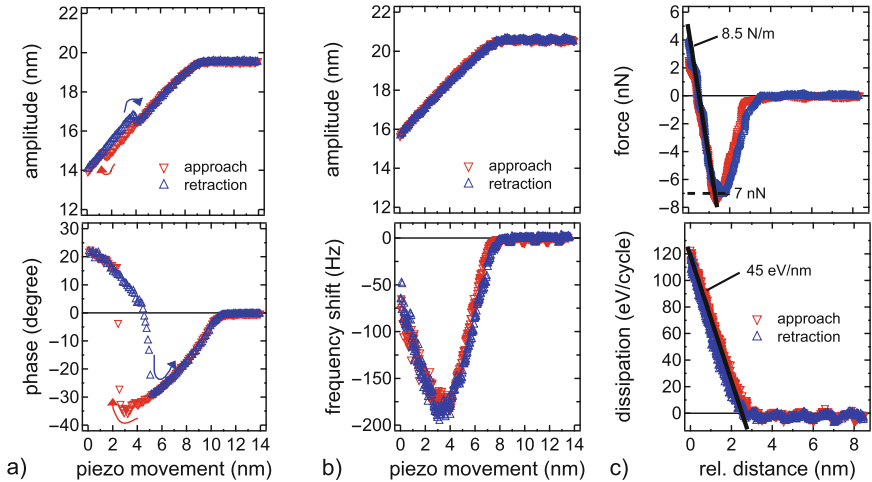


Fig. 1.2. Examples of “spectroscopy measurements” obtained on mica in ambient conditions. **(a)** Amplitude and phase vs. distance curves in the tapping mode. The instabilities during approach and retraction cause a hysteresis. **(b)** Such a behavior is not observed in the constant-excitation mode where the approach and retraction curves of the amplitude and frequency shift are identical within the noise limit. The parameters of the cantilever were $f_0 = 167\,224$ Hz, $c_z = 37.5$ N m $^{-1}$, and $Q = 465$. **(c)** Using the algorithm described in the text, we reconstruct the tip-sample potential and force from the data sets shown in **(b)**. The interaction force decreases until it reaches a minimum of -7 nN, and increases again with a slope of 8.5 N m $^{-1}$. The origin of all x -axes has been arbitrarily set to the left of the graphs

45 eV nm $^{-1}$. The physical origin of the energy dissipation, however, is still under discussion (see, e.g., [57–59]).

1.3 Mapping of the Tip–Sample Interactions on DPPC Monolayers in Ambient Conditions

In the previous section, we showed that the FM technique is well suited to measure the tip–sample force at an arbitrary position. However, it is also possible to record the tip–sample interaction in a more defined way as contour maps perpendicular to the sample surface. To examine the possible resolution of this approach under ambient conditions, we recorded sets of spectroscopy curves along predefined scan lines on DPPC (L- α -dipalmitoyl-phosphatidylcholine, Fluka), which frequently serves as a model for membranes [60]. The tip–sample interaction was subsequently calculated from the measured amplitude and frequency shift vs. distance curves with respect to the actual scan position. The obtained curves were then plotted in a color-coded contour map showing the potential of the tip–sample interaction.

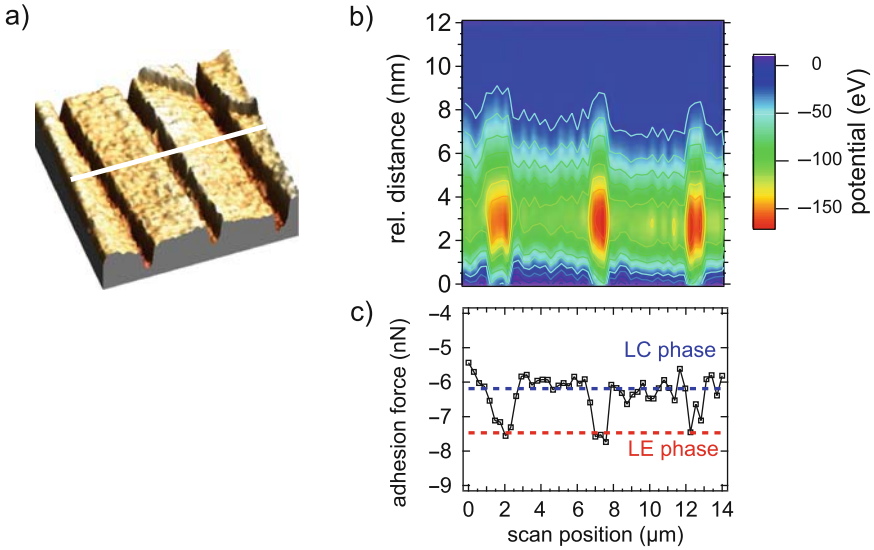


Fig. 1.3. (a) Surface plot (scan size: $14 \times 14 \mu\text{m}^2$) of the topography of the DPPC film prepared by the Langmuir–Blodgett technique. The monolayer shows alternating stripes and channels which consists of DPPC adsorbed in the LC- and LE-phase, respectively. The white line marks the position where we recorded the frequency shift and amplitude vs. distance curves for the construction of the contour map of the tip–sample interaction potential shown in (b). The graph in (c) displays the corresponding adhesion force obtained from the data shown in (b). The parameters of the cantilever were $f_0 = 170\,460 \text{ Hz}$, $c_z = 39.6 \text{ N m}^{-1}$, and $Q = 492$

Monolayers of DPPC were prepared with the Langmuir–Blodgett technique. As shown in the topography image in Fig. 1.3, the monolayers have a lateral structure of alternating stripes and channels. This specific pattern is obtained by rapidly withdrawing the mica substrate at a low monolayer surface pressure and constant temperature as described by Gleiche et al. [61]. The stripes consist of DPPC in a liquid condensed phase (LC-phase), whereas the channels between the stripes are filled with DPPC in the liquid expanded phase (LE-phase) [62]. The lateral periodicity of stripes and channels depends on the parameters used during the preparation of the sample.

We imaged the sample using the oscillation amplitude as a feedback signal in the FM-mode using constant-excitation before we recorded 50 spectroscopy curves along a predefined direction marked in Fig. 1.3a. All data sets were then transformed into tip–sample potential curves using the mathematical method described in Sect. 1.2.3. Finally, we computed the corresponding contour map as shown in Fig. 1.3. The complete procedure was done by a computer script using IGOR Pro software (Wavemetrics Inc.).

The resulting color-coded image reveals the different tip–sample interaction on the stripes (LC-phase) and in the channels (LE-phase). The potential

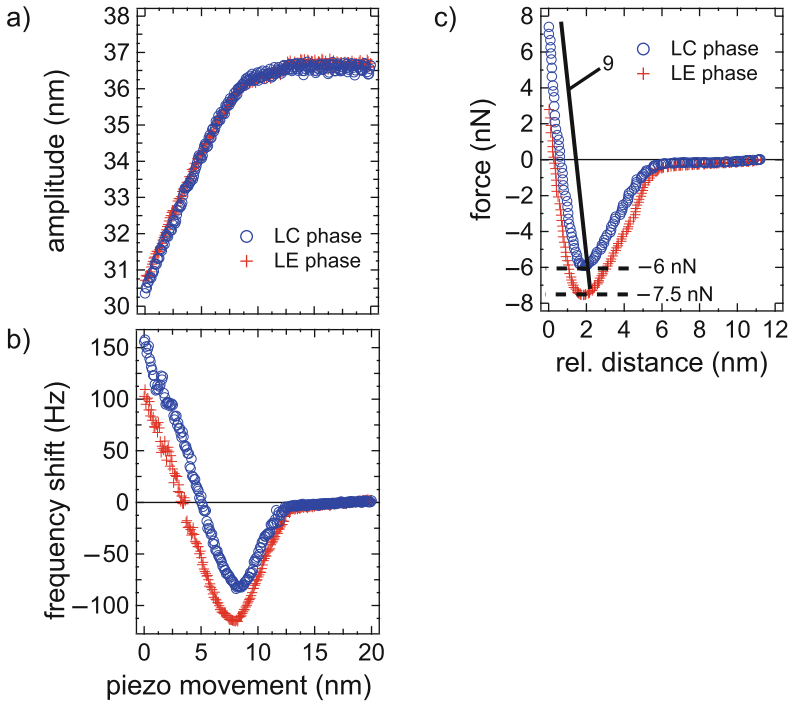


Fig. 1.4. (a) Spectroscopy curves obtained on the stripes (*circles*) and the channels (*crosses*). The tip-sample interaction is calculated from this amplitude and frequency shift vs. distance curves. (b) Using the numerical procedure described in the text we calculated the corresponding tip-sample interaction force and potential. A significant difference between the curves is observed for distances between 2 and 6 nm

is significantly larger above the stripes (≈ -100 eV) compared to the channels (≈ -150 eV), as it can be seen by the color coding in Fig. 1.3.

The local stiffness and adhesion force can be determined from the force curves plotted in Fig. 1.4. Here, we show amplitude and frequency shift vs. distance curves and the resulting tip-sample force measured on the LC- and LE-phase. The force curves reveal an adhesion force of -6 nN and -7.5 nN on the stripes and channels, respectively. However, the local stiffness is about 9 N m^{-1} for both positions, and we could not determine a significant difference in the stiffness between the two phases of the DPPC monolayers. This outcome can be attributed to the fact that the repulsive interaction forces for a thin film depends strongly on its thickness [63]. Therefore, the contact stiffness measured on the two phases of the DPPC monolayer is dominated by the mica substrate.

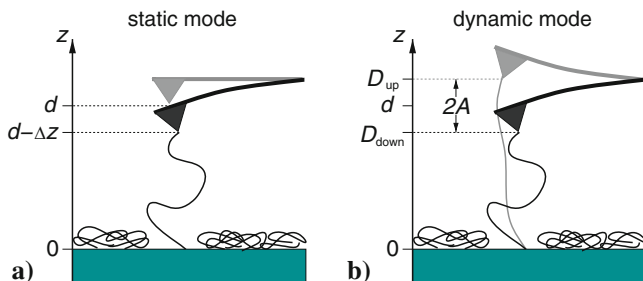


Fig. 1.5. Comparison between static and dynamic mode measurement of forces acting on a chain-like molecule. (a) Typical experimental set-up in the static mode where a chain-like molecule attached to a tip is stretched. Measured quantities are the bending Δz of the cantilever (i.e., the force) vs. the extension of the molecule. (b) In dynamic force spectroscopy the cantilever oscillates with an amplitude A around its equilibrium position d while it is retracted from the surface covered with chain-like molecules. If a molecule binds to the tip the resonance frequency f of the cantilever changes in dependence of the cantilever position. As explained in the text this signal is used to determine a force vs. extension curve

1.4 Force Spectroscopy of Single Dextran Monomers in Liquid

As already mentioned in the introduction, atomic force microscopy is frequently used for the precise measurement of the forces acting on chain-like molecules during their stretching (see, e.g., [3, 64–66] and references therein). In this section, we show how the forces acting on a chain-like molecule in liquid can be also measured in the constant-excitation mode of the FM technique. The difference between the conventional static mode and our dynamic approach is depicted in Fig. 1.5. In the static mode (Fig. 1.5a), the cantilever is approached and retracted from the surface covered with chain-like molecules (dextran monomers in our case). If now, one of these molecules has bound to the tip it is stretched during the retraction of the cantilever from the surface. A measurement of the cantilever bending Δz during the retraction is used to measure the force vs. extension curve of the molecule. In the dynamic mode, however, the cantilever oscillates between the positions D_{down} and D_{up} while it is approached and retracted from the surface (Fig. 1.5b). Since the chain-like molecule exerts an attractive force to the cantilever, its fundamental features like the oscillation amplitude A and resonance frequency f will be modified. In the following, we use this effect to record the force vs. extension curve together with the energy dissipation.

The dynamic force spectroscopy experiments were measured with the same set-up already described in the previous sections. For the sample preparation, the polysaccharide dextran (Dextran 500, Sigma Aldrich) was solved in pure water ($\approx 0.2 \text{ g ml}^{-1}$). Afterwards, a small amount ($100 \mu\text{l}$) of this solution was dropped onto a cleaned glass substrate. After drying the substrate was

rinsed several times with water to remove the loose molecule chains. The measurements were performed in a droplet (40–50 μl) of buffer solution (PBS Dulbecco, Biochrom AG) in an open liquid cell without an o-ring. All data sets shown here were obtained with a rectangular silicon cantilever (Mikro-Masch NSC36). Its spring constant in liquid was determined using the method given by Sader et al. [67]. The used cantilever had a spring constant of $c_z = 24.6 \text{ N m}^{-1}$, an eigenfrequency of $f_0 = 28.3 \text{ kHz}$, and a quality factor of $Q = 64$.

Figure 1.6 shows a typical dataset recorded with this experimental set-up. The oscillation amplitude (Fig. 1.6a) and the shift of the resonance frequency (Fig. 1.6b) are displayed as a function of the cantilever-sample distance d . Both the amplitude A and the frequency shift $\Delta f = f - f_0$ remain almost constant during the approach. Only when the tip comes very close to the surface the repulsive tip-sample interaction leads to a strong decrease (amplitude) and increase (frequency shift) of these two signals. During the retraction, however, a dextran strand adhered to the tip, resulting into a very characteristic pattern in the amplitude as well as in the frequency shift data. A minimum, a maximum, and a sudden jump appear as prominent features in the two curves. We observed that this pattern is a clear sign for a dextran molecule stretched during the retraction.

In the previous sections, we demonstrated how the tip-sample force can be reconstructed from the frequency shift and amplitude data via an integral equation. This approach, however, was based on the assumption that the tip-sample force decreases to zero if the tip is far away from the sample surface. In our case, however, it is just the other way around. The attractive force caused by the dextran molecule is practically zero at the surface and increases during the retraction from the sample surface. Consequently, we have to switch the integration limits of the integral equation used to calculate the extension force in dynamic force spectroscopy. Using the approach of Sader and Jarvis [51] we obtain

$$F(D_{\text{up}}) = -\frac{\partial}{\partial D_{\text{up}}} 2c_z \int_0^{D_{\text{up}}} \frac{\Delta f(z)}{f_0} \times \left[(D_{\text{up}} - z) + \sqrt{\frac{A(z)}{16\pi}} \sqrt{D_{\text{up}} - z} + \frac{A(z)^{3/2}}{\sqrt{2(D_{\text{up}} - z)}} \right] dz, \quad (1.19)$$

where the upper turning point D_{up} is easily calculated from the transformation $D_{\text{up}} = d + A$ (see Fig. 1.5b). The energy dissipated during one oscillation cycle, can be directly calculated using (1.18).

Due to the jump in the frequency shift and amplitude curve, we can apply these formulas only for the data on the left side of the jumps. (After the jump the molecule is ruptured from the tip and the tip-sample interaction is practically zero.) The solid lines on the left side of the arrows in Fig. 1.6a, b

mark this subset. Introducing now this data into (1.19), we obtain the force vs. distance curve displayed in Fig. 1.6c.

As it is well known from previous studies of dextran [3, 64, 68], the force curve reveals a kink in the range of about 700 pN. It originates from a conformational transition within each dextran monomer. Due to the external applied force each sugar ring of the dextran monomer flips into a new conformation [64] resulting in an additional elongation (about 10%) of the monomer. This is a thermodynamic process which can be described with a “single-click” model [68]. The solid line in Fig. 1.6c shows a fit of this model to the experimental data. The agreement is remarkable, because the only fit parameter is the number of monomers in the dextran molecule. All other parameters were fixed and chosen as given in [64] and [68].

This agreement was also observed for other stretching events. Figure 1.7 summarizes the measured force curves of three dextran strands of different length. Again, we could fit the “single-click” model to this data. Since this model is not time or velocity dependent and based on the assumption that the number of folded and unfolded molecules are in a thermodynamic equilibrium, this outcome proves that we indeed measure the equilibrium force of the dextran molecule even with our dynamic approach.

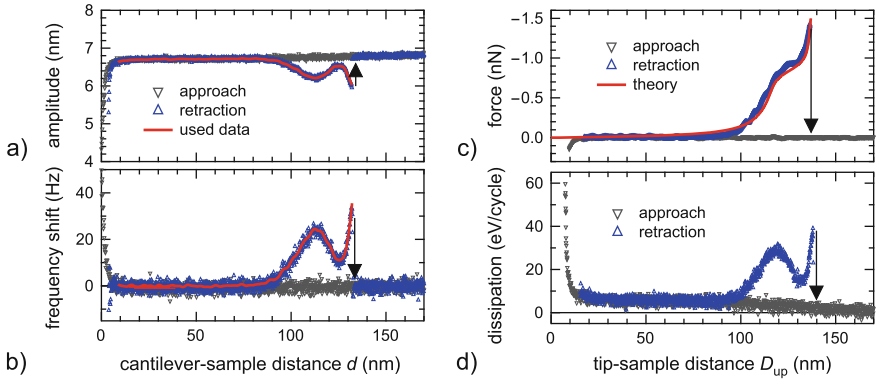


Fig. 1.6. An application of the introduced dynamic force spectroscopy technique. (a) Amplitude (A) and (b) frequency shift (Δf) curves measured during approach to and retraction from the surface covered with dextran molecules. During the retraction one dextran molecule bound to the tip as revealed by the change of the frequency and amplitude signal. At a position of about 135 nm (see *arrows*) the maximum binding force was exceeded and the cantilever oscillated freely again. Only the data before this jump is used for the subsequent analysis. (c) Using (1.19) the force acting on the dextran molecule (*symbols*) is reconstructed as a function of the actual tip position D_{up} . The experimental result is well described by a “single-click” model using only the number of molecules ($N = 266$) as fitting parameter (*solid line*). (d) The energy dissipated per oscillation cycle can be calculated from (1.18) for approach and retraction. The zero of the x -axes is arbitrarily set to the left side of the graphs

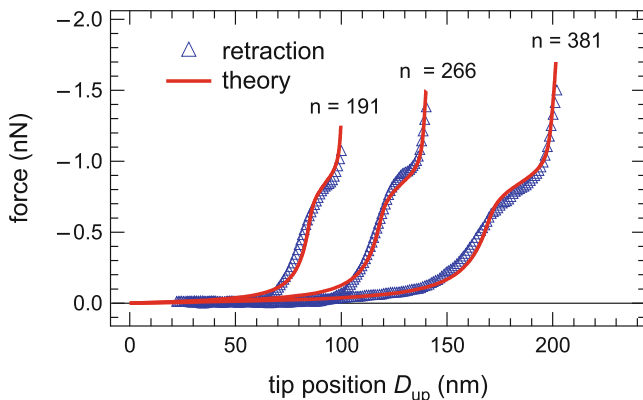


Fig. 1.7. (Color online) Force vs. extension curves of three different dextran strands measured with the proposed dynamic approach. The *symbols* represent experimental data which is well reproduced by the theoretical force curves where only the number of dextran monomers were used as a fit parameter ($n = 191$, 266 , and 381 , respectively)

As mentioned previously, it is also possible to calculate the dissipated energy per oscillation cycle via (1.18) in dynamic force spectroscopy. The resulting dissipation curve is shown in Fig. 1.6d as a function of the tip-sample distance and shows a maximum at a position corresponding to the kink in the force curve. We always observed this coincidence in the conservative and dissipative interaction curves. Since the kink appears at the position where most of the dextran monomers unfold, it is very likely that a large amount of energy dissipation is caused by this unfolding processes. However, the dissipation increases also shortly before the rupture of the dextran strand from the tip. At this position, the majority of the dextran monomers is already unfolded. Therefore, we conclude that other effects like hydrodynamic damping and/or the rupture of the molecule from the tip will contribute to the energy dissipation.

1.5 Summary

In summary, we presented an overview of the FM technique focusing on its application in air and liquids. After reviewing the basic theoretical foundations of this mode, we showed how the FM mode can be used to measure conservative as well as dissipative tip-sample interactions. The feasibility of this approach was demonstrated by its application in air (mica and DPPC monolayers) and liquids (stretching of dextran monomers).

Acknowledgements

The authors would like to thank X. Chen and L. Chi for the preparation of the DPPC sample. Furthermore, we acknowledge support from and many useful discussions with André Schirmeisen, Harald Fuchs (University of Münster), and Filipp Oesterhelt (University of Düsseldorf). This work was supported by the BMBF *Junior Researchers Competition Nanotechnology* (Grant No. 03N8704).

References

1. G. Binnig, C.F. Quate, C. Gerber, Phys. Rev. Lett. **56**, 930 (1986)
2. E.L. Florin, V.T. Moy, H.E. Gaub, Science **264**, 415 (1994)
3. M. Rief, F. Oesterhelt, B. Heymann, H.E. Gaub, Science **275**, 1295 (1997)
4. Y. Martin, C.C. Williams, H.K. Wickramasinghe, J. Appl. Phys. **61**, 4723 (1987)
5. Q.D. Zhong, D. Inness, K. Kjoller, V.B. Elings, Surf. Sci. Lett. **290**, L688 (1993)
6. R. Giles, J.P. Cleveland, S. Manne, P.K. Hansma, B. Drake, P. Maivald, C. Boles, J. Gurley, V.B. Elings, Appl. Phys. Lett. **63**, 617 (1993)
7. C.A.J. Putman, K.O. Vanderwerf, B.G. Degrooth, N.F. Vanhulst, J. Greve, Appl. Phys. Lett. **64**, 2454 (1994)
8. P.K. Hansma, J.P. Cleveland, M. Radmacher, D.A. Walters, P.E. Hillner, M. Bezanilla, M. Fritz, D. Vie, H.G. Hansma, C.B. Prater, J. Massie, L. Fukunaga, L. Gurley, V.B. Elings, Appl. Phys. Lett. **64**, 1738 (1994)
9. T.R. Albrecht, P. Grütter, D. Horne, D. Rugar, J. Appl. Phys. **69**, 668 (1991)
10. H. Hölscher, B. Gotsmann, W. Allers, U.D. Schwarz, H. Fuchs, R. Wiesendanger, Phys. Rev. B **64**, 075402 (2001)
11. H. Hölscher, B. Gotsmann, W. Allers, U.D. Schwarz, H. Fuchs, R. Wiesendanger, Phys. Rev. Lett. **88**, 019601 (2002)
12. F.J. Giessibl, Science **267**, 68 (1995)
13. Y. Sugawara, M. Otha, H. Ueyama, S. Morita, Science **270**, 1646 (1995)
14. B. Gotsmann, B. Ancykowski, C. Seidel, H. Fuchs, Appl. Surf. Sci. **140**, 314 (1999)
15. H. Hölscher, W. Allers, U.D. Schwarz, A. Schwarz, R. Wiesendanger, Phys. Rev. Lett. **83**, 4780 (1999)
16. U. Dürig, Appl. Phys. Lett. **75**, 433 (1999)
17. M.A. Lantz, H. Hug, R. Hoffmann, P.J.A. van Schendel, P. Kappenberger, S. Martin, A. Baratoff, H.J. Güntherodt, Science **291**, 2580 (2001)
18. H. Hölscher, Surf. Sci. **515**, 517 (2002)
19. A. Schirmeisen, D. Weiner, H. Fuchs, Phys. Rev. Lett. **97**, 136101 (2006)
20. N.A. Burnham, R.J. Colton, J. Vac. Sci. Technol. A **7**, 2906 (1989)
21. F.J. Giessibl, Phys. Rev. B **1997**, 16010 (1997)
22. H. Hölscher, B. Ancykowski, Surf. Sci. **579**, 21 (2005)
23. A.A. Farrell, T. Fukuma, T. Uchihashi, E.R. Kay, G. Bottari, D.A. Leigh, H. Yamada, S.P. Jarvis, Phys. Rev. B **72**, 125430 (2005)
24. T. Uchihashi, M.J. Higgins, S. Yasuda, S.P. Jarvis, S. Akita, Y. Nakayama, J.E. Sader, Appl. Phys. Lett. **85**, 3575 (2004)
25. G. Meyer, N.M. Amer, Appl. Phys. Lett. **53**, 1045 (1988)

26. S. Alexander, L. Hellemans, O. Marti, J. Schneir, V. Elings, P.K. Hansma, J. Appl. Phys. **65**, 164 (1988)
27. D. Rugar, H.J. Mamin, P. Guethner, Appl. Phys. Lett. **55**(25), 2588 (1989)
28. C. Schönenberger, S.F. Alvarado, Rev. Sci. Instrum. **60**, 3131 (1989)
29. A. Moser, H.J. Hug, T. Jung, U.D. Schwarz, H.J. Güntherodt, Meas. Sci. Technol. **4**, 769 (1993)
30. S. Kitamura, M. Iwatsuki, Jpn. J. Appl. Phys. **34**, L145 (1995)
31. H. Ueyama, Y. Sugawara, S. Morita, Appl. Phys. A **66**, S295 (1998)
32. K. Kobayashi, H. Yamada, K. Matsushige, Appl. Surf. Sci. **188**, 430 (2002)
33. T. Okajima, H. Sekiguchi, H. Arakawa, A. Ikai, Appl. Surf. Sci. **210**, 68 (2003)
34. D. Ebeling, H. Hölscher, B. Anczykowski, Appl. Phys. Lett. **89**, 203511 (2006)
35. J.E. Schmutz, H. Hölscher, D. Ebeling, M.M. Schäfer, B. Anczykowski, Ultramicroscopy **107**, 875 (2007)
36. T. Fukuma, T. Ichii, K. Kobayashi, H. Yamada, K. Matsushige, Appl. Phys. Lett. **86**, 034103 (2005)
37. T. Fukuma, K. Kobayashi, K. Matsushige, H. Yamada, Appl. Phys. Lett. **86**, 193108 (2005)
38. B.W. Hoogenboom, H.J. Hug, Y. Pellmont, S. Martin, P.L.T.M. Frederix, D. Fotiadis, A. Engel, Appl. Phys. Lett. **88**, 193109 (2006)
39. Y. Martin, H.K. Wickramasinghe, Appl. Phys. Lett. **50**, 1455 (1987)
40. H. Hölscher, B. Gotsmann, A. Schirmeisen, Phys. Rev. B **68**, 153401 (2003)
41. J.P. Cleveland, B. Anczykowski, A.E. Schmid, V.B. Elings, Appl. Phys. Lett. **72**, 2613 (1998)
42. T.R. Rodríguez, R. García, Appl. Phys. Lett **80**, 1646 (2002)
43. H. Hölscher, U.D. Schwarz, Int. J. Nonlinear Mech. **42**, 608 (2007)
44. U. Dürig, N. J. Phys. **2**, 5.1 (2000)
45. J.E. Sader, T. Uchihashi, M.J. Higgins, A. Farrell, Y. Nakayama, S. Jarvis, Nanotechnology **16**, S94 (2005)
46. A.I. Livshitz, A.L. Shluger, A. Rohl, Appl. Surf. Sci **140**, 327 (1999)
47. H. Hölscher, A. Schwarz, W. Allers, U.D. Schwarz, R. Wiesendanger, Phys. Rev. B **61**, 12678 (2000)
48. U. Dürig, Appl. Phys. Lett. **76**, 1203 (2000)
49. F.J. Giessibl, Appl. Phys. Lett. **78**, 123 (2001)
50. O. Pfeiffer, Ph.D. thesis, University of Basel, Switzerland, 2004
51. J.E. Sader, S.P. Jarvis, Appl. Phys. Lett. **84**, 1801 (2004)
52. B. Gotsmann, B. Anczykowski, C. Seidel, H. Fuchs, Phys. Rev. B **60**, 11051 (1999)
53. J.E. Sader, I. Larson, P. Mulvaney, L.R. White, Rev. Sci. Instrum. **66**, 3789 (1995)
54. D. Ebeling, H. Hölscher, H. Fuchs, B. Anczykowski, U.D. Schwarz, Nanotechnology **17**, S221 (2006)
55. B. Anczykowski, D. Krüger, H. Fuchs, Phys. Rev. B **53**, 15485 (1996)
56. A. San Paulo, R. García, Phys. Rev. B **66**, 041406 (2002)
57. L.N. Kantorovich, T. Trevelyan, Phys. Rev. Lett. **93**, 236102 (2004)
58. A. Schirmeisen, H. Hölscher, Phys. Rev. B **72**, 045431 (2005)
59. R. Hoffmann, A. Baratoff, H.J. Hug, H.R. Hidber, H.v. Löhneysen, H.J. Güntherodt, Nanotechnology **18**, 395503 (2007)
60. E. Sackmann, Science **271**, 43 (1996)
61. M. Gleiche, L.F. Chi, H. Fuchs, Nature **403**, 173 (2000)

62. X. Chen, S. Lenhert, M. Hirtz, N. Lu, H. Fuchs, L. Chi, *Acc. Chem. Res.* **40**, 393 (2007)
63. J. Domke, M. Radmacher, *Langmuir* **14**, 3320 (1998)
64. M. Rief, J.M. Fernandez, H.E. Gaub, *Phys. Rev. Lett.* **81**, 4764 (1998)
65. A. Janshoff, M. Neitzert, Y. Oberdörfer, H. Fuchs, *Angew. Chem. Int. Ed.* **39**, 3212 (2002)
66. P. Hinterdorfer, Y.F. Dufrene, *Nature Methods* **3**, 347 (2006)
67. J.E. Sader, J.W.M. Chon, P. Mulvaney, *Rev. Sci. Instrum.* **70**(10), 3967 (1999)
68. R.G. Haverkamp, A.T. Marshall, M.A.K. Williams, *Phys. Rev. E* **75**, 021907 (2007)

Photonic Force Microscopy: From Femtonewton Force Sensing to Ultra-Sensitive Spectroscopy

O.M. Maragò, P.G. Gucciardi, and P.H. Jones

Summary. Photonic force microscopy is a scanning probe technique based on the use of optical tweezers. An optically trapped probe particle is scanned over a surface or in a three-dimensional structured environment to probe force interactions. This technique is capable of force sensing at the sub-piconewton level with spatial resolution in the nanometre range depending on the dimensions of the trapped particle used as probe. Here we review some recent advances in the technique focussing on applications based on the use of linear nanostructures, bio-force sensing and imaging, and integration with ultra-sensitive spectroscopy such as Raman tweezers.

Key words: Optical Tweezers, Photonic Force Microscopy, Force Sensing, Linear Nanostructures, Bionanotechnology, Raman Tweezers.

Abbreviations

AFM	Atomic force microscopy
BFP	Back focal plane
CCD	Charge-coupled device
NA	Numerical aperture
NIR	Near infrared
OT	Optical tweezers
PFM	Photonic force microscopy
QPD	Quadrant photodiode
RBM	Radial breathing mode
RT	Raman tweezers
SERS	Surface-enhanced Raman scattering
SLM	Spatial light modulator
SWNTs	Single wall carbon nanotubes

2.1 Introduction

Light moves matter. Although this simple concept has been known for centuries (Kepler's explanation of comets tail is based on it), the advent of the laser age has led to tremendous experimental advances and understanding of this phenomenon. Optical Tweezers (OT) [1, 2], instruments based on a tightly focussed laser beam, have been used to trap, manipulate, control and assemble dielectric particles, single atoms, cells, metal and semiconducting nanostructures, leading to a real *optical revolution* in Physics, Biology and Nanotechnology [3, 4]. When used as a force transducer, OT are able to measure forces in the piconewton range [5, 6]. In this context, the concept of Photonic Force Microscopy (PFM) has been developed by scanning a dielectric sphere over surfaces in a liquid environment and sensing the force interaction between trapped sphere and surface [7]. Since the lateral resolution in PFM applications is related to particle size, it is desirable to extend conventional optical trapping to nanoparticles. This is a difficult task since the radiation force scales as the volume of the trapped particle. Recently, optical trapping of semiconductor nanowires [8–10] and carbon nanotubes bundles [11–16] has been achieved in aqueous environments. The almost linear geometry of these nanostructures with a sub-wavelength cross-section and very high aspect ratio is of crucial importance for bridging the gap between the micro- and nano-world. Their small (sub-optical wavelength) transverse size is the key to achieving nanometric resolution for PFM applications [10], while an axial dimension in the micron range ensures stable trapping. Together these features allow to tailor force sensing with 1D nanostructures in the femtonewton regime [16].

In this chapter, we will first introduce the concepts underpinning optical trapping of matter and the principal characteristics of an OT (Sect. 2.2). We will then go on to discuss how OT are implemented experimentally, and describe the particle tracking methods by which they can be usefully calibrated (Sect. 2.3). Here, we also discuss developments in optical trapping of linear nanostructures with a view to using them as a PFM probe. In Sect. 2.4, we will introduce the idea of calibrated OT used as a scanning probe, paying particular attention to examples of imaging and force spectroscopy in life science applications. Section 2.5 introduces Raman spectroscopy to OT, and looks ahead to possibilities arising from integrating these techniques.

2.2 Principles of Optical Trapping

2.2.1 Theoretical Background

Light forces are generated by the scattering of electromagnetic fields incident on a particle, hence the natural starting point for a quantitative understanding of optical trapping is scattering theory of electromagnetic radiation [17, 18].

The difficulties arising from the use of the full scattering theory are generally overcome by solving the problem in different regimes depending on the size of the scatterer. While technical improvements in OT have become more sophisticated throughout the years, only recently have careful optical trapping calculations on spherical particles found a quantitative correspondence with experiments [19–21]. Moreover, the models traditionally used for calculating optical forces are based on approximations which often limit the discussion only to spherical particles. Here, we give a brief review of these approaches and of recent advances in force calculations through electromagnetic theory in the transition matrix (T-matrix) approach [17, 21]. This latter approach is able to give a size independent description of optical trapping and an extension to nonspherical particles.

For dielectric particles with size $d \gg \lambda$ (ray optics regime), Ashkin used *ray optics* to attempt a quantitative description of light forces [22]. He considered ‘rays’ of light being refracted at the interface between dielectric media. For simplicity, consider the case of a spherical particle of refractive index n_p , suspended in a medium (generally water) of refractive index n_m and illuminated by a Gaussian laser beam (Fig. 2.1a). Refraction of the ray as it crosses the sphere implies a transfer of momentum from the sphere to the ray, and hence an equal and opposite transfer of momentum from the ray to the sphere.

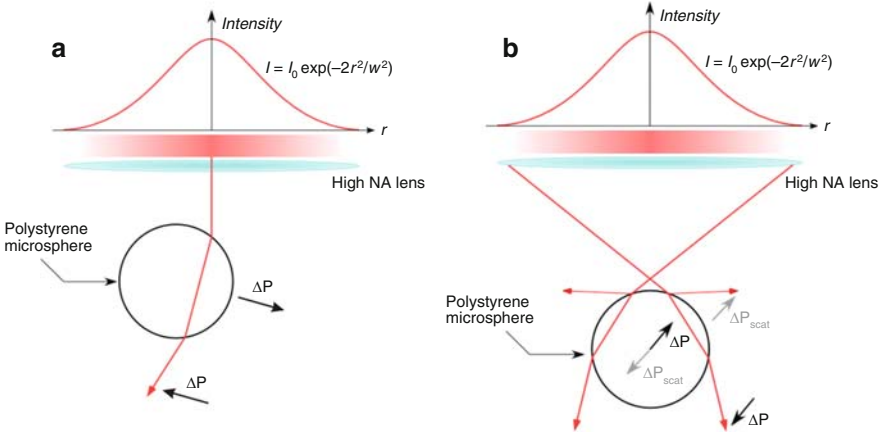


Fig. 2.1. Principles of optical trapping in ray optics approximation for dielectric particles with size $d \gg \lambda$. (a) A spherical particle of refractive index n_p is suspended in a medium (generally water) of refractive index n_m and illuminated by a Gaussian laser beam. Refraction of the ‘ray’ as it crosses the sphere implies a transfer of momentum from the sphere to the ‘ray’, and hence an equal and opposite transfer of momentum from the ‘ray’ to the sphere. The gradient in light intensity (number of ‘rays’) across the sphere produces a net transverse force towards the beam axis, that is, an *optical gradient* force. (b) To achieve trapping in the axial (z -) direction requires focussing of the beam where a similar argument for refraction providing an optical gradient force towards the focus can be made

The gradient in light intensity (number of rays) across the sphere produces a net transverse (radial) force towards the beam axis, that is, an *optical gradient* force. To achieve trapping in the axial (beam propagation) direction requires focussing of the beam as shown in Fig. 2.1b. In this case, a similar argument for refraction providing an optical gradient force directed towards the focus can be made. Axial trapping must also overcome the ‘pushing’ effect of the small reflection of light due to the mismatch in refractive indices and absorption, that is, the *optical scattering* force. Stable 3D trapping requires that the gradient force exceeds the scattering force, which is achieved in single beam OT with strong (high numerical aperture) focussing.

In the opposite regime, when $d \ll \lambda$ (dipole regime), the particle is polarised by the electric field of the laser and the *dipole approximation* can be used to derive the resulting force directly from the dipole interaction potential. The dependence of optical forces on the refractive index of the trapped particle can be evaluated by treating the particle as a point dipole in an inhomogeneous electromagnetic field and separating the scattering force (due to absorption and re-radiation) [23, 24]:

$$\mathbf{F}_{\text{scatt}}(\mathbf{r}) = \frac{n_m}{c} \sigma_{\text{scatt}} \langle \mathbf{S} \rangle = \frac{n_m}{c} \sigma_{\text{scatt}} I_0 \mathbf{g} \quad (2.1)$$

$$\sigma_{\text{scatt}} = \frac{128 \cdot \pi^5 \cdot d^6}{3 \cdot \lambda^4} \left(\frac{m^2 - 1}{m^2 + 2} \right)^2 \quad (2.2)$$

where I_0 is the field intensity, n_m is the refractive index, σ_{scatt} is the cross-section of the radiation pressure onto the particle, c is the light speed in vacuum, while $\langle \mathbf{S} \rangle$ represents the time averaged Poynting vector of the light scattered from the bead, and \mathbf{g} is the vector transfer that indicates the momentum transfer direction. The cross-section depends on the ratio $m = n_p/n_m$ between the refractive index of the bead and that of the medium, and the radiation wavelength λ . On the other hand, the electric field intensity gradient $\nabla \langle E^2 \rangle$ generates the gradient force:

$$\mathbf{F}_{\text{grad}}(\mathbf{r}) = \frac{n_m}{2} \alpha \nabla \langle E^2 \rangle \quad (2.3)$$

$$\alpha = d^3 \cdot \left(\frac{m^2 - 1}{m^2 + 2} \right) \quad (2.4)$$

where α is the polarisability of the particle. This separation into two components is a direct consequence of the dipole approximation and can also be derived for an atom in a quantum mechanics framework leading to laser cooling and atom optical trapping guiding [25]. The intensity distribution is typically Gaussian and for small displacements from the focal point we have a restoring force proportional to the gradient in intensity. Strong confinement (typically within 100 nm of positional fluctuations) is therefore achieved by strong focussing (see Fig. 2.2 for calculated intensity plots of a tightly focussed laser beam).

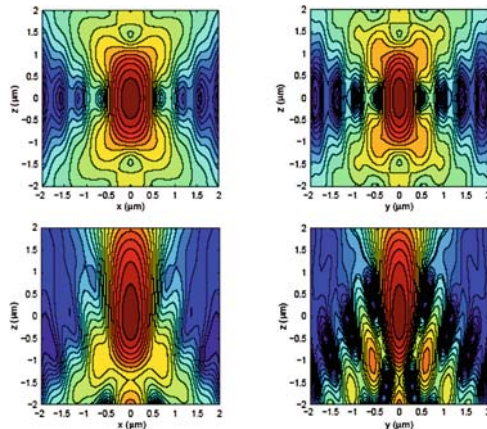


Fig. 2.2. Contour plots (log scale) of the intensity $|E(r)|^2$ of the unaberrated Gaussian (*top panels*) and aberrated Gaussian field (*lower panels*) in the xz (*left panels*) and in the yz plane (*right panels*). The origin of the z axis (intensity maximum) in the lower panels is shifted by $\Delta F = -4.0 \mu\text{m}$ due to aberration (reprinted from [21])

The *intermediate regime*, when $d \approx \lambda$ is more difficult to treat with a simple approximation, but is actually the usual regime of operation for OT experiments using infra-red laser light. In this case, a complete wave-optical modelling of the particle light interaction is necessary for calculating the optical trapping forces. The situation becomes even more complicated for nonspherical or optically anisotropic particles. Such particles are of particular importance because most biological structures of interest are nonspherical and often anisotropic. To compute the interaction of a particle with a focussed laser beam in the intermediate regime, many approaches have been developed. The most accurate procedures start by calculating the configuration of the field in the focal region in the absence of any particle, as done by Richards and Wolf [24,26]. An example of this procedure is given in Fig. 2.2 which shows intensity plots of a laser beam focussed by an oil immersion objective with numerical aperture (NA) of 1.3 are shown together with the effect of aberration when passing through a coverslip [21]. Once the field is known, the radiation force exerted on any particle can be calculated by resorting to momentum conservation for the combined system of field and particles. Although the resulting expression of the radiation force requires consideration of Maxwell stress tensor, several procedures have been reported to avoid using it when dealing with optical trapping [19,20]. Recently, a more rigorous approach has been developed based on the Transition Matrix formalism [17,27], resulting in a calculation of optical trapping (as well as optical torque) that can be applied to particles of any size and symmetry [21,28]. The radiation force exerted on particles is given by the integral [29]:

$$\vec{F}_{\text{rad}} = r'^2 \int_{\Omega'} \hat{r}' \cdot \langle \vec{\mathbf{T}}_M \rangle d\Omega', \quad (2.5)$$

where the integration is over the full solid angle, r' is the radius of a large sphere centred at $\mathbf{R}_{O'}$ surrounding the particle centre, and $\langle \vec{\mathbf{T}}_M \rangle$ is the time averaged Maxwell stress tensor. When the incident field is a polarised plane wave, the components of the radiation force along the direction of the unit vector \hat{v}_ξ are given by [21]

$$\vec{F}_{\text{Rad}\xi} = -\frac{r'^2}{16\pi} \text{Re} \int_{\Omega'} (\hat{r}' \cdot \hat{v}_\xi) [n^2 (|\vec{E}'_S|^2 + 2\vec{E}'_I^* \cdot \vec{E}'_S) + (|\vec{B}'_S|^2 + 2\vec{B}'_I^* \cdot \vec{B}'_S)] d\Omega' \quad (2.6)$$

where \vec{E}_I and \vec{B}_I are the incident fields, while \vec{E}_S and \vec{B}_S are the fields scattered by the particle. Expanding the incident field in a series of vector spherical harmonics with (known) amplitudes W_{ilm}^p , the scattered field can be expanded on the same basis with amplitudes $A_{l'm'}^{p'}$. The relation between the two amplitudes is given by $A_{l'm'}^{p'} = \sum_{plm} S_{l'm'lm}^{p'p} W_{ilm}^p$, where $S_{l'm'lm}^{p'p}$ is the transition (T-)matrix of the particle [27]. In this framework, any nonspherical particle can be modelled as an aggregate of spheres with size well below the radiation wavelength. The elements of the T-matrix are calculated in a given frame of reference through the inversion of the matrix of the linear system obtained by imposing to the field's boundary conditions across each spherical surface [30,31]. A comprehensive treatment of all these topics, related to the calculation of the transition matrix, can be found in [17]. Here we stress that these elements turn out to be independent both on the direction of propagation and on the polarisation of the incident field. Thus, they do not change when the incident field is a superposition of plane waves with different direction of propagation, that is, the description of a focussed laser beam in the angular spectrum representation [24].

This establishes a complete formalism based on the transition matrix approach that can be thoroughly applied to OT experiments. An example is shown in Fig.2.3, the trapping efficiency (i.e. the fraction of optical momentum transferred to the object) for a spherical latex bead trapped in water with a size in the intermediate regime ($d = 830\text{nm}$, $\lambda_0 = 1,064\text{nm}$ in vacuo) is plotted in both unaberrated (left panel) and aberrated (right panel) case. Using this approach, excellent agreement has been found between theory and experiments by comparing trapping efficiencies and polarisation effects [21]. Furthermore calculations have been extended to linear nanostructures [32] (such as polymer nanofibres and semiconductor nanowires) where a sub-wavelength cross-section and a high aspect ratio are of fundamental importance for bridging the gap between the microworld and the nanoworld [33]. It has been shown that for size much smaller than the trapping wavelength, the trapped linear nanostructure is stably oriented with the

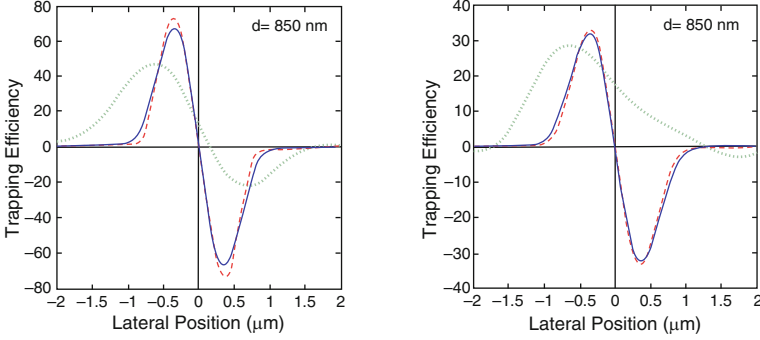


Fig. 2.3. Optical trapping efficiency $Q_x(x, 0, z_0)$ (blue solid line), $Q_y(0, y, z_0)$ (red dashed line) and $Q_z(0, 0, z)$ (green dotted line) for single spheres with diameter $d = 850$ nm and $\lambda_0 = 1,064$ nm as a function of the position of their centres. Optical trapping in both un aberrated (left panels) and aberrated beam (right panels) is considered. Note that the origin of the z axis in the right panels is shifted by $\Delta F = -4.0 \mu\text{m}$ because of aberration (reprinted from [21])

polarisation axis. Otherwise radiation torque changes the orientation along the optical axis (propagation direction) [32].

2.3 Experimental Implementation

2.3.1 Optical Tweezers Set-up

Experimentally, optical trapping is accomplished by tightly focussing a laser beam through a high NA lens. To this end, water or oil immersion microscope objectives are generally used. In Fig. 2.4, a typical set-up is shown that is built around a commercial inverted microscope. The inverted configuration gives more stability because it helps to counteract gravity with radiation pressure for weak traps. The actual optical trap in this system is realised using a microscope objective with $\text{NA} = 1.3$ and an 830 nm, 150 mW laser diode as a radiation source. Beam steering is often attained through a pair of orthogonally mounted, computer controlled galvomirrors [34] or acousto-optic deflectors [35], so that dynamic multiple traps and time-averaged traps can be created via time-sharing of the trapping beam. A telescope is used to enlarge the beam so to overfill the microscope objective back aperture leading to a diffraction limited focal spot. Samples are placed in a small chamber attached on a piezo-stage that enables positioning of the sample with nanometric resolution. A charged-coupled device (CCD) camera is used to image the trapped particles. Size calibration of the image can be obtained by imaging a calibrated slide or by optical trapping size-standard latex beads.

Finally, force sensing is achieved through back focal plane interferometry (BFP) [36, 37]. The BFP of the condenser is imaged onto a four quadrant

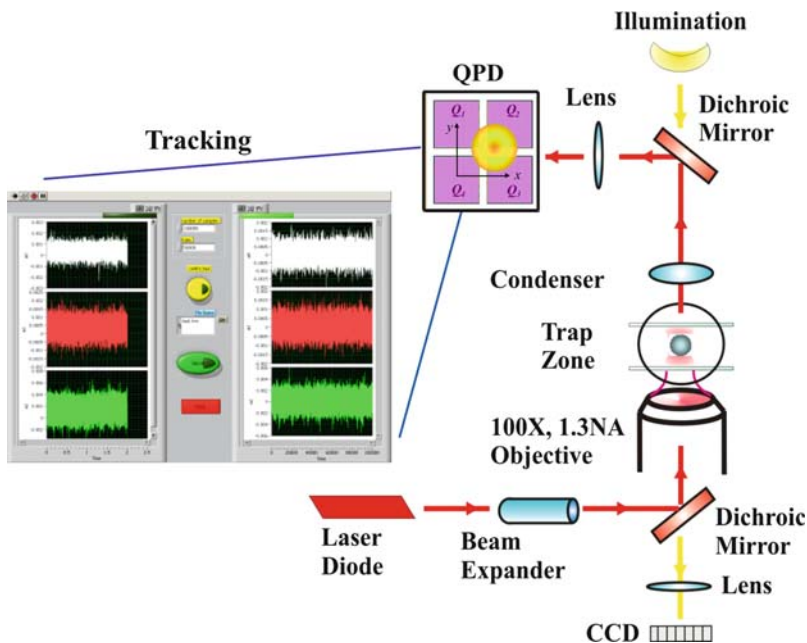


Fig. 2.4. Optical tweezers set-up. A laser beam (typically in the NIR) is expanded to overfill the back aperture of a high numerical aperture lens. The light is focussed in a chamber, with the sample solution. A CCD is used to image the trapped sample. A quadrant photodiode (QPD) detects the interference from the back focal plane of the condenser that collects the light from the trapping region. Particle tracking signals are obtained by a linear combination of the QPD output voltages

photodiode (QPD). This is oriented with the polarisation axis of the laser beam in order to have sensitivity over polarisation effects on the optical trap [21]. The laser light transmitted through the microscope carries an interference pattern between unscattered and forward scattered light from the trapped particle. The spatial distribution of the interference pattern depends on the position of the particle relative to the waist of the focussed laser beam. The outputs Q_i from all quadrants are processed through an analogue circuit board that generates electrical signals proportional to the trapped particle displacement in the three spatial directions (particle tracking). In the system shown in Fig. 2.4, for example, the signal $S_x = (Q_1 + Q_3) - (Q_2 + Q_4)$ is proportional to the displacement in the x -direction, while $S_y = (Q_1 + Q_2) - (Q_3 + Q_4)$ is proportional to the one in the y -direction and the sum $S_z = Q_1 + Q_2 + Q_3 + Q_4$ is proportional to the displacement in the z -direction. The analysis of these signals gives a wealth of information on the calibration of the trapping potential, on the measure of spring constants, on micro-rheology and more generally on Brownian motion spectroscopy [38].

More recent versions of OT set-ups include the use of spatial light modulators (SLM) in the beam path that enables the creation of more complex optical landscapes. The SLM permits arbitrary phase patterning over the trapping laser wavefront that is dynamically reconfigurable, and addressable in real time [39–41], replicating the effect of optical elements such as diffraction gratings, lenses or spiral phase plates. These allow the generation of multiple trapping sites by splitting a single beam, movement of the trap location in three dimensions or trapping in an optical vortex [42, 43] which transfers orbital angular momentum to the trapped particle, resulting in controlled rotations.

2.3.2 Brownian Motion and Force Sensing

A crucial element in force sensing with OT is Brownian motion [44, 45], that is, the random motion of particles in a fluid due to thermal fluctuations. For an optically trapped particle, position fluctuations due to Brownian motion are tracked by the QPD, appearing as randomly fluctuating signal voltages, similar to those shown in the traces of Figs. 2.4 and 2.5a. The statistical analysis of these fluctuations enables force measurements [46–50]. As will be discussed in Sect. 2.4, the Brownian dynamics can be utilised in PFM as a means of 3D scanning of the probe nanoparticle, and therefore the extent of positional fluctuations does not necessarily limit the imaging resolution.

The starting point for OT calibration based on Brownian motion analysis is the Langevin equation [51] in a confining harmonic potential $V(x) = \sum \frac{1}{2} k_i x_i^2$. For a trapped spherical particle of size d , this can be written as

$$\partial_t x_i(t) = -\omega_i x_i(t) + \xi_i(t), \quad i = x, y, z \quad (2.7)$$

Where the relaxation frequency $\omega_i = k_i/\gamma$ are related to the force constants and hydrodynamic (viscous) damping $\gamma = 3\pi\eta d$, being η the dynamic viscosity of the surrounding liquid. The terms $\xi_i(t)$ describe random uncorrelated fluctuations with zero mean $\langle \xi_i(t) \rangle = 0$ and

$$\langle \xi_i(t) \xi_j(t + \tau) \rangle = 2 \frac{k_B T}{\gamma} \cdot \delta(\tau) \delta_{ij} = 2D \cdot \delta(\tau) \delta_{ij} \quad (2.8)$$

Where the angled brackets indicate a time-averaged quantity and D is the diffusion constant (Einstein relation). Note that the delta-like correlation function is typical of an ideal white noise power spectrum $|\tilde{\xi}_i(\omega)|^2 = 2D$.

In order to get the force constants from experimental particle tracking signals, two major procedures are used that rely on a frequency domain (*power spectrum density*) [47, 48] and time domain (*autocorrelation function*) [49, 50] analysis of the thermal fluctuations in the trap. In the first procedure, we consider the Fourier transform of (2.7):

$$-i\omega \tilde{x}_i(\omega) = -\omega_i \tilde{x}_i(\omega) + \tilde{\xi}_i(\omega) \quad (2.9)$$

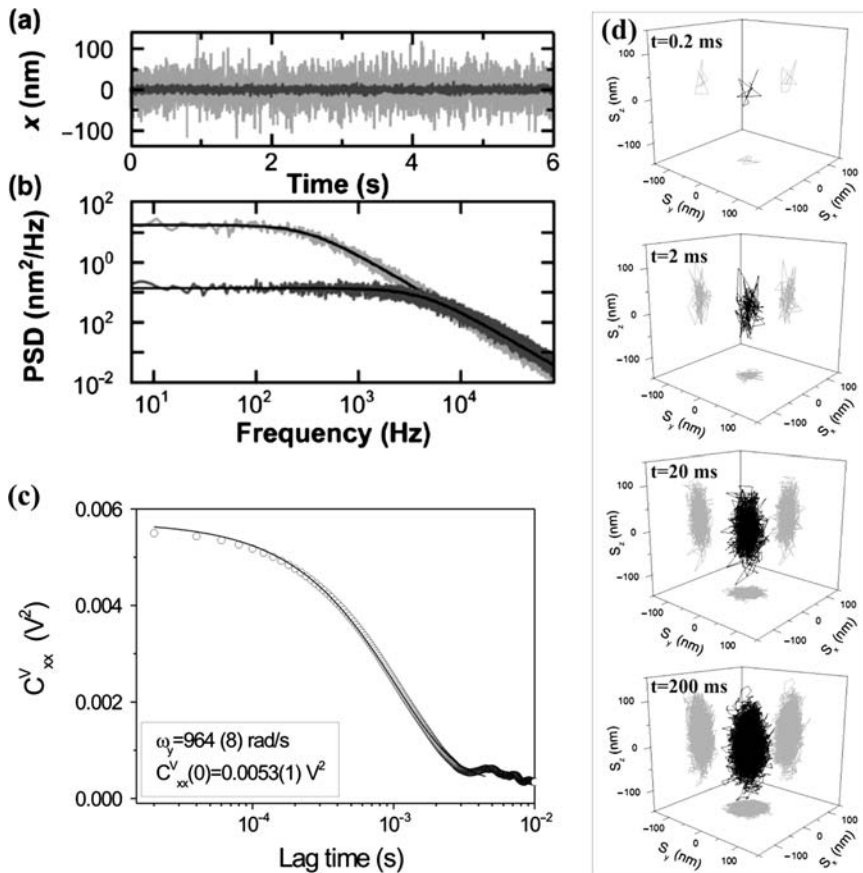


Fig. 2.5. Calibration of optical tweezers. (a) Position record of a 100 nm gold (*dark gray*) and a 110 nm polystyrene (*light gray*) bead. (b) Averaged power spectra fit for the same gold (*dark gray*) and polystyrene bead (*light gray*). Lorentzian fits (*solid curve*) yield relaxation frequencies and eventually force sensing calibration ((a) and (b) reprinted from [48]). (c) Autocorrelation analysis for a 2 μm latex bead and (d) reconstruction of Brownian motion at different times in the calibrated optical trap

and solve for:

$$\tilde{x}_i(\omega) = \frac{\tilde{\xi}_i(\omega)}{-i\omega + \omega_i} \quad (2.10)$$

Thus [using the Fourier transform of (2.8)] the corresponding power spectral density has a Lorentzian shape:

$$S(\omega) = \frac{2D}{\omega^2 + \omega_i^2} \quad (2.11)$$

with a half-width ω_i and a zero frequency point value $2D/\omega_i^2$ related to the force constants and thermal diffusion, respectively.

In time domain, it is instead useful to consider the autocorrelation functions of the position fluctuations:

$$C_{ii}(\tau) = \langle x_i(t)x_i(t + \tau) \rangle \quad (2.12)$$

That from (2.7) and (2.8) obey first-order uncoupled differential equations with the lag time τ :

$$\partial_\tau C_{ii}(\tau) = -\omega_i C_{ii}(\tau) \Rightarrow C_{ii}(\tau) = C_{ii}(0) \cdot e^{-\omega_i \tau} \quad (2.13)$$

Equation (2.13) can be easily integrated, giving exponential decays with relaxation frequencies ω_i and zero point value $C_{ii}(0) = D/\omega_i$. Examples of these procedures are shown in Fig. 2.5. In (a), the tracking signals for a gold nanoparticle (*dark gray trace*) and a polystyrene bead (*light gray*) are compared, while their power spectra are shown in (b). The Lorentzian curve fitting of these spectra yields the force constants and a full calibration of the optical trap. In Fig. 2.5c, we show a typical autocorrelation function of a tracking signal for a 2 μm latex bead. In this case (i.e. in the time domain), the force constants and the calibration of the OT can be obtained by fitting the autocorrelation function with an exponential decay. After calibration is achieved, a reconstruction of the Brownian motion in the trap is also obtained (Fig. 2.5d). In both procedures, obtaining the two fitting parameters enables measurements of displacements with nanometric precision, and force sensing with sub-piconewton precision.

2.3.3 Optical Trapping of Linear Nanostructures

Optical trapping has been recently extended to linear nanostructures. Semiconductor nanowires [8–10] and single wall carbon nanotubes (SWNTs) bundles [11–16] are ideally suited as probes in next generation PFM because of their intrinsic nanometric transverse size and micrometric axial length that ensure very stable trapping.

In Fig. 2.6, some examples of optical manipulation and trapping of linear nanostructures are shown. GaN, SnO₂, ZnO or Si nanowires with different transverse size (20–200 nm range) and length (1–100 μm) were suspended in water and the suspension was transferred by means of a pipette into a chamber. Nanowires were stably trapped [9] using near-infrared (NIR) light at 1,064 nm. Junctions and assemblies were then built using OT (Fig. 2.6a, b). Assembly of complex nanowire structures requires not only manipulation of individual wires, but also the controlled connection of one wire to another. It was observed that it is possible to locally fuse two wires by means of a focussed NIR beam. On laser irradiation of the crossing point of two wires using higher power (~ 1 W) than for trapping, the two wires were fused together.

Potassium niobate (KNbO₃) nanowires were also optically trapped [10]. These wires exhibit efficient second harmonic generation, and act as frequency converters, allowing the local synthesis of a wide range of colours via sum and

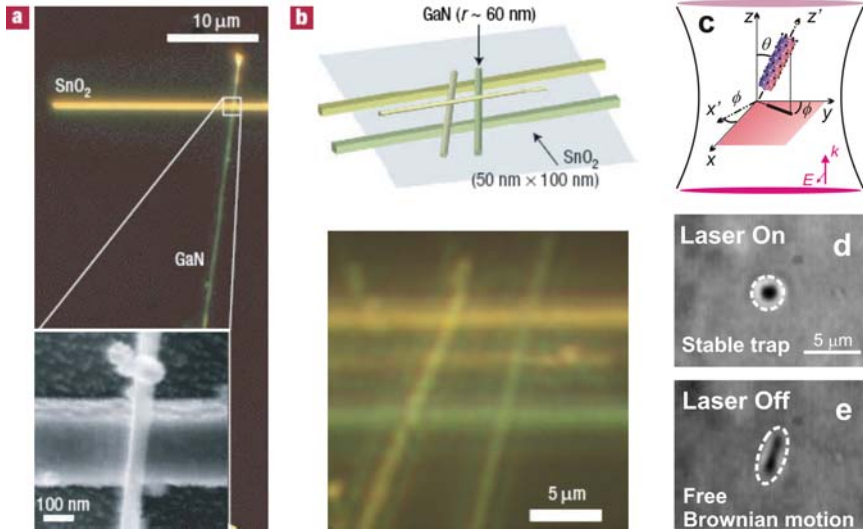


Fig. 2.6. Optical trapping and manipulation of linear nanostructures. (a) and (b) nanowires junctions and assemblies built using optical tweezers (reprinted from [9]). (c) Geometry with Euler angles relevant in the optical trapping of SWNT bundles. (d) A laser trapped SWNT bundle oriented by radiation torque along the optical axis. (e) The same bundle un-trapped (laser is off) and randomly oriented by Brownian motion (c), (d) and (e) (redrawn from [16])

difference frequency generation. Thus, an optically trapped KNbO_3 can be used as a tunable nanometric light source to implement a novel form of sub-wavelength microscopy, in which an OT is used to trap and scan a nanowire over a sample, suggesting a wide range of potential applications.

Figure 2.7 shows a proof-of-principle demonstration of this novel *nanowire scanning microscopy*. A test pattern was fabricated via electron beam lithography consisting of a series of 200-nm-wide, 50-nm-thick lines of gold on transparent glass coverslips with decreasing pitch between each line. An image of the gold pattern (Fig. 2.7b) was obtained by atomic force microscopy (AFM) and shows the progressively finer separation between each line, decreasing from 1,000 to 200 nm (Fig. 2.7c). The OT was then used to manipulate a single KNbO_3 nanowire over the gold pattern while a piezoelectric stage with nanometre positioning resolution was used to scan the pattern with respect to the wire. Each time the trapped wire passed over a gold feature, there was a reduction of transmitted second harmonic emission. Scanning the nanowire tip directly across the surface resulted in a transmitted intensity map with local spatial resolution of the order of the wire's diameter (Fig. 2.7d). Measurements of the pitch between lines within the gold pattern were made for both AFM and nanowire optical transmission line scans and agreed to within 10%. As expected, the minimum resolvable feature-size (pitch) decreases as the cross-section of the nanowire-probe decreases.

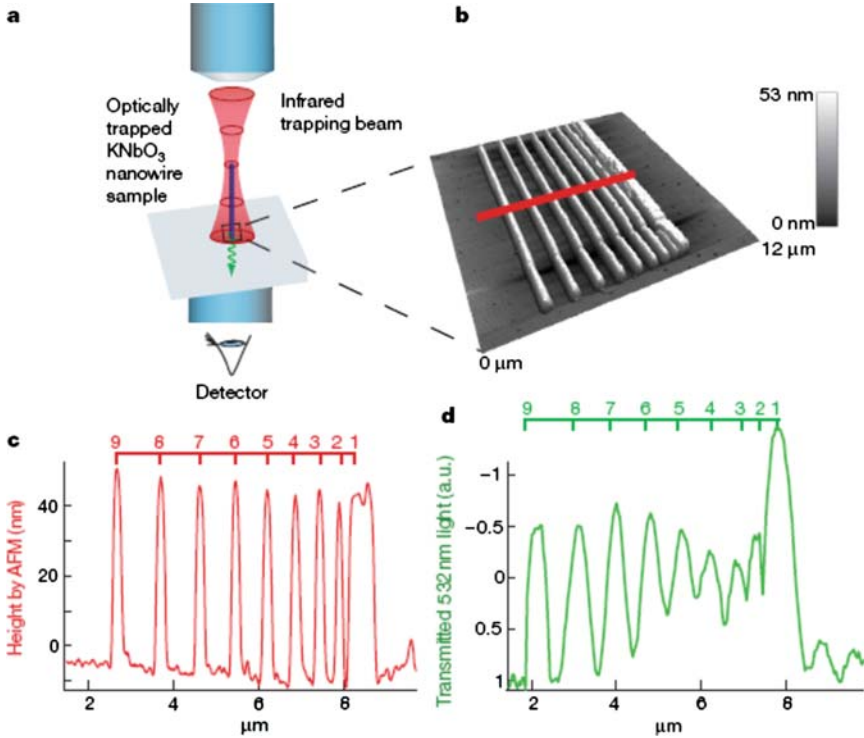


Fig. 2.7. Transmission line scan of metallic surface pattern with optically trapped KNbO₃ nanowire. (a) Schematic of inverted optical scanning configuration. (b) AFM topographic image of thermally evaporated pattern of gold stripes on a glass coverslip. (c) AFM line scan from region indicated in (b). (d) Optical transmission profile captured by scanning a single KNbO₃ nanowire over the metallic surface structure. The nanowire used to create the transmission line scan was measured by AFM to have these dimensions: width = 122 nm, length = 1.4 μm and height = 53 nm. The peak index (there are nine lines in total being measured) is shown at the top of (c) and (d) (reprinted from [10])

Carbon nanotube bundles were also trapped using NIR light. Figure 2.6d, e shows an SWNT bundle when optically trapped (Fig. 2.6d) and aligned with the laser propagation direction (the imaging axis) and when untrapped (Fig. 2.6e) freely floating and randomly oriented. In these experiments, carbon nanotubes dispersions were prepared using purified HiPCO single wall nanotubes (SWNTs) in water with sodium dodecyl benzene sulphonate. Solutions were ultrasonicated and filtered [14]. Few tens of micro-litres were then placed in a small chamber attached to a piezo-stage.

When dealing with quantitative measurements and calibration, the first issue to deal with is hydrodynamics of the trapped particle. For linear nanostructures (rigid rod-like structures), the viscous drag is described by an anisotropic hydrodynamic mobility tensor [52], the components of which

depend on the length-to-diameter ratio $p = L/d$ as [53]

$$\Gamma_{\perp} = \frac{\ln p + \delta_{\perp}}{4\pi\eta L}, \quad \Gamma_{\parallel} = \frac{\ln p + \delta_{\parallel}}{2\pi\eta L}, \quad \Gamma_{\Theta} = \frac{3(\ln p + \delta_{\theta})}{\pi\eta L^3}, \quad (2.14)$$

where Γ_{\perp} and Γ_{\parallel} are the translational mobilities, transverse and parallel to the main axis, Γ_{Θ} is the rotational mobility about midpoint, η is the water dynamical viscosity, and δ_i are end corrections (calculated in [53] as polynomial of $(\ln 2p)^{-1}$).

In a recent work [16], the distribution of both centre-of-mass and angular fluctuations from 3D tracking of optically trapped nanotubes was extracted. For the first time, the optical force and torque constants from auto and cross-correlation of the tracking signals were measured. The latter allowed the isolation of the angular Brownian motion, and hence a full calibration of the OT for using nanotubes (as well as any other linear nanostructure) as force sensing probes.

For an optically trapped bundle, positional and angular displacements from equilibrium are detected using BFP interferometry as for latex beads. While for spherical objects, the detector signals are combined so that they are proportional to the centre-of-mass displacements, for nonspherical particles they also contain angular information. The projections onto the laboratory axes (see Fig. 2.6c) are determined through the Euler angles (ϕ, θ) and the corresponding rotation matrix. Since the trapped bundle is aligned with z , fluctuations occur in the small polar angle limit, $\theta \ll 1$, and the QPD signals are [16]

$$S_x \sim \beta_x(X + L\Theta_x), \quad S_y \sim \beta_y(Y + L\Theta_y), \quad S_z \sim \beta_z Z, \quad (2.15)$$

where β_i values are the detector calibration factors, X , Y , and Z are the centre-of-mass coordinates, L is the bundle length, and $\Theta_x = \theta \sin \varphi$ and $\Theta_y = \theta \cos \varphi$ are the projections on the x and y axis, respectively (shown in Fig. 2.6c).

The centre-of-mass X_i and angular Θ_j coordinates are treated as stochastic variables. Thus, the Brownian dynamics of the trapped SWNTs can be described by a set of uncoupled Langevin equations:

$$\partial_t X_i(t) = -\omega_i X_i(t) + \xi_i(t), \quad i = x, y, z \quad (2.16)$$

$$\partial_t \Theta_j(t) = -\Omega_j \Theta_j(t) + \xi_j(t), \quad j = x, y, \quad (2.17)$$

where $\omega_i = \Gamma_i k_i$ and $\Omega_j = \Gamma_{\Theta} k_j$ are relaxation frequencies, related to the force and torque constants and mobility tensor components, while $\xi_i(t)$ are random noise sources. From (2.15)–(2.17), the autocorrelations of the transverse QPD signals $C_{ii}(\tau) = \langle S_i(t) S_i(t + \tau) \rangle$ now contain combined information on centre-of-mass and angular fluctuations, and decay with lag time τ as a double exponential with positional and angular relaxation frequencies ω_i, Ω_i ($i = x, y$) (see Fig. 2.8a, d). Furthermore, as the angular variables are geometrically correlated, the cross-correlations $C_{xy}(\tau) = \langle S_x(t) S_y(t + \tau) \rangle = C_{xy}(-\tau)$ of the

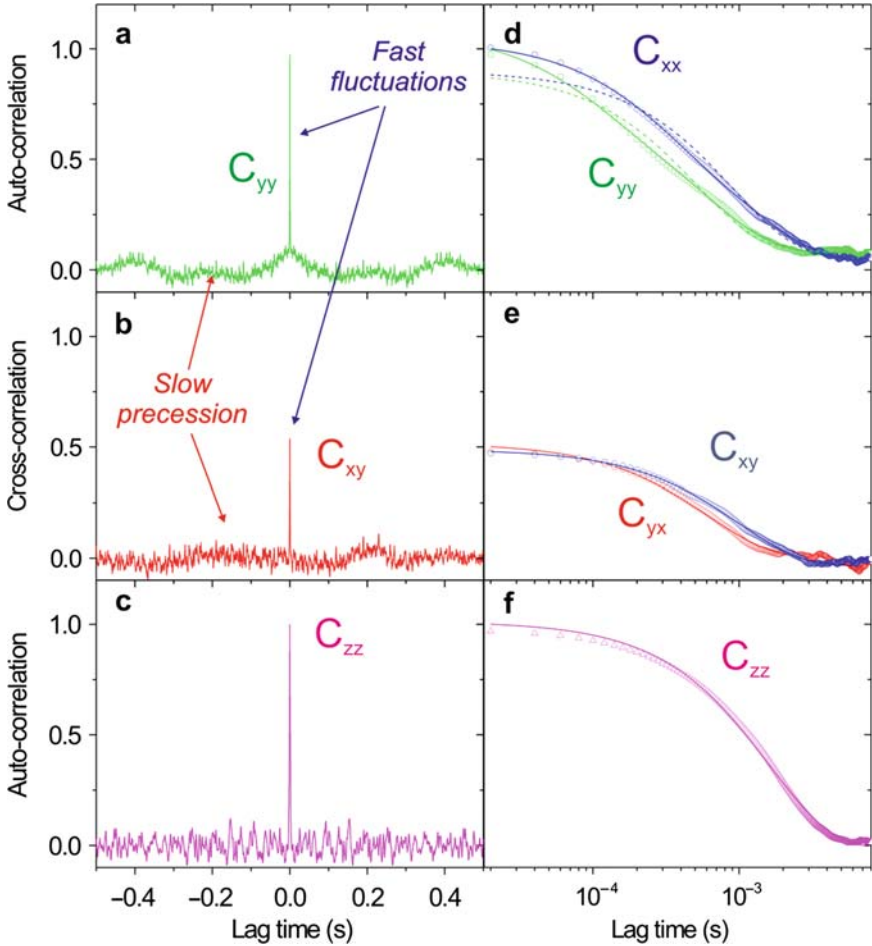


Fig. 2.8. Correlation analysis. (a) Transverse signal autocorrelation and (b) signal cross-correlation on a linear scale for lag time. The separation of time scales for fast centre-of-mass and polar angle fluctuations, and slow azimuthal precession is evident. (c) Longitudinal signal autocorrelation on linear scale. Only a fast centre-of-mass fluctuations signal is seen. (d) Signal autocorrelation and (e), signal cross-correlation, plotted on a logarithmic scale. Solid lines show a double-exponential fit for the autocorrelation of the transverse signals and a single exponential for the cross-correlation with a decay rate consistent with that found in the autocorrelation. (f) Autocorrelation of longitudinal signal on logarithmic time scale, showing a single exponential fit, with much longer decay time than for transverse signals (reprinted from [16])

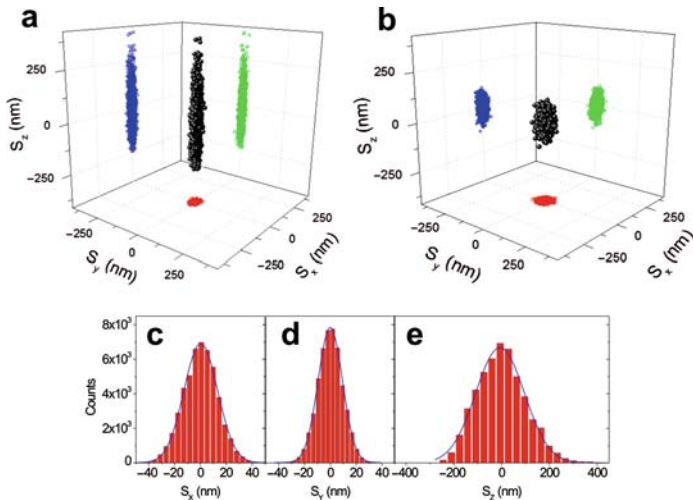


Fig. 2.9. Brownian motion in the optical trap. (a) Trapped SWNT bundle and (b) $2\ \mu\text{m}$ latex bead. In both cases, the tracking (*black dots*) is obtained by extracting 2,000 points from the QPD signals. For the trapped nanotubes, an increase in mobility compared to the spherical particle along the z axis is evident. (c,d) Histograms of the transverse signals. In the transverse direction, both translational and angular fluctuations are convoluted into a signal that includes both centre-of-mass displacement and the projection of the bundle onto the measurement axis. The root mean squares of the transverse displacements, extracted from a Gaussian fit, are 13 nm and 9 nm in the x and y direction, respectively. (e) Histogram of the longitudinal signal. This is only due to fluctuations of the centre of mass and has a measured root mean square of 100 nm (reprinted from [16])

transverse signals decay as single exponentials with relaxation rates different for positive and negative lag times, corresponding to Ω_x and Ω_y (see Fig. 2.8b, e). This also allowed measurement of an optical torque of $\sim 1\text{pN}\cdot\mu\text{m}$ on the trapped SWNT bundle.

Figure 2.9 shows the Brownian motion of a trapped nanotubes bundle with a transverse size of 10 nm and a length of $3\ \mu\text{m}$, in (a), to be compared with the one of a trapped $2\ \mu\text{m}$ latex bead, in (b), as reconstructed through the tracking of the fluctuating signals. Histograms of the 3D fluctuations (Fig. 2.9 c,d and e) ensure spatial resolution within 10 nm, while the force calibration makes the trapped bundle a very sensitive probe of forces in the axial direction with resolution <10 femtonewton.

These combined force and torque measurements enable the calibration of the OT when using linear nanostructures as force sensing probes in the femtonewton regime.

2.4 Photonic Force Microscopy

A scheme for scanning optical force microscopy was first described in 1993 by Ghislaine and Webb [54] based on observing the motion of an optically trapped probe particle held in a single beam optical gradient force trap (OT). By using the techniques of particle position sensing described in Sect. 2.3 with a trapped prolate glass shard used as the probe particle, substrate features of width 20 nm were resolved, with a lower limit imposed by the noise floor of an estimated 15 nm. By analogy with the AFM, this technique soon came to be known as Photonic Force Microscopy [7].

In AFM, the small forces between the nanometric tip and a sample surface are measured from the deflection of a mechanical cantilever with a calibrated spring constant [55]. In PFM, the tip is replaced by a trapped particle and the mechanical spring is replaced with an optical trap that has a spring constant about three orders of magnitude smaller. Consequently, this much softer spring permits the measurement of forces of a few piconewtons, with a resolution of tens of femtonewtons, in liquid environment.

Besides its spectacular force sensitivity, PFM shows several advantages with respect to other scanning probe microscopy techniques, among which its easy operation in liquid and the capacity of analysis ‘inside the structure’. The small load exerted by the optically trapped particle in PFM enables the imaging of soft materials, such as cells membranes, without deformation. Moreover, the lack of physical contact to the probe, which is held in a laser beam, presents advantages for intra-cellular imaging by trapping and manipulating particles located directly in the cell cytoplasm.

2.4.1 Bio-Nano-Imaging

Imaging of soft biological structures at the sub-diffraction scale is an attractive goal for PFM methodology. Three-dimensional particle tracking, as employed in PFM, offers a distinct advantage over conventional video microscopy not only in volumetric imaging versus the two-dimensional projection of a camera, but also in the high temporal resolution – the tracking signal may be sampled at rates of several tens of kilohertz, compared to 25 Hz video frame rate.

The potential of optically trapped particles for high-resolution imaging was demonstrated by Florin et al. in 1996 [56]. Fluorescent beads were held in an OT created by a strongly focussed Nd:YAG laser beam which also served to excite fluorescence emission via two-photon absorption. Displacement of the bead from the equilibrium position was observed by measuring the changes in fluorescence intensity when the bead was pressed against either a hard (glass) or soft (neuronal cell) surface. In this way, images of sub-optical wavelength structures can be obtained by using the trapped bead as a scanning probe in a manner analogous to the constant height mode of AFM. As with AFM, other modes of imaging can be realised by monitoring the fluctuating

particle position and closing a feedback loop to the position of the optical tweezer focus.

The spatial resolution of this imaging technique is determined by the size of the probe particle and the extent of its thermally driven fluctuations from the equilibrium position. The extent of the fluctuations may be of the order of several hundred nanometres, depending on the details of the trapping potential, but this need not provide the ultimate limitation on PFM resolution. Three-dimensional particle tracking methods [37], like those used in the calibration of the optical trap, can be used to reconstruct the position of the bead with a spatial resolution of better than 1 nm and a temporal resolution of a few microseconds [57]. This opens the possibility of using the thermal fluctuations to drive the probing particle to explore small volumes ($\sim 1 \mu\text{m}^3$) and track the extent and spatial distribution of the fluctuations to reveal information about the interaction of the probe with the local environment [58].

The so-called thermal noise imaging [59] can be separated from other optical tweezers-based techniques for measuring forces that rely on an initial calibration followed by the application and measurement of force along one direction. Tracking the PFM probe particle reveals the complete 3D interaction potential that arises from both the optical trap and the system being probed. The volume explored by the particle can be represented by a surface of constant frequency (frequency isosurface) derived from a histogram of the 3D tracking signal. In general, when the particle is interacting with the trapping potential only, this volume is ellipsoidal in shape (Fig. 2.10). The major axis of this ellipsoid lies along the direction of propagation of the trapping laser

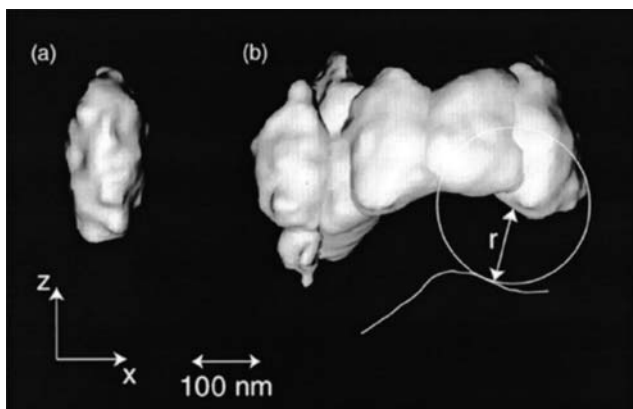


Fig. 2.10. Sensitivity of an optically trapped, thermally driven probe particle to restrictions in the local environment. On the left is a three-dimensional representation of the positional fluctuations of an optically trapped bead in solution. On the right is the same in a network of agar filaments. The filaments restrict the volume available to the probe particle, resulting in a fluctuation distribution that reflects the filament structure (reprinted from [59])

beam as this is the direction in which the curvature of the trapping potential is lowest. The presence of other objects within the probe volume, or a change in the interaction of the probe with its environment for example through a change in local viscosity or adhesion interaction, changes the form of the volume that is accessible to the particle. An instructive example is the photonic force imaging of a network of agar filaments using the technique described above with a latex sphere of 216 nm diameter [59]. As shown in Fig. 2.10, the frequency isosurface of the sphere fluctuations is radically altered from the ellipsoidal shape in solution by the presence of the agar filaments. It should be noted that the surface is the result of a convolution of the bead fluctuations and the filament shape, and therefore the extent of the bead must be accounted for in reconstructing the probed object.

An application of this method is the observation of diffusion of plasma membrane components. In [37], the motion of the membrane protein Thy1.1 in a neurite of a hippocampal neuron was observed via tracking of an attached latex sphere. The volume in which the bound sphere moves is shown in Fig. 2.11, clearly demonstrating the advantage of PFM as a 3D imaging technique.

The extent of the Brownian position fluctuations is also dependent on the viscosity of the medium surrounding the particle, and so the photonic force microscope is also a sensitive probe of local viscosity over the volume explored by the trapped particle during its thermal motion. For typical parameters, this volume can be as small as $1 \mu\text{m}^3$. This method was first applied to the

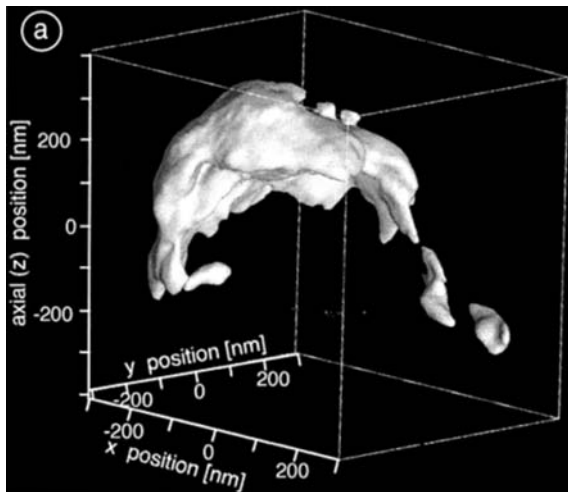


Fig. 2.11. Particle tracking of a probe particle tethered to a Thy1.1 molecule diffusing on a neurite membrane. From the shape of the volume in which the bead moves the outline structure of the neurite is clearly visible. From the motion, the diffusion constant of the membrane protein and hence the presence of membrane structures may be deduced (reprinted from [37])

change in diffusion coefficient due to the presence of a fixed boundary for a spherical particle in a fluid of finite extent [60], showing good agreement with the analytical Faxen correction to the Stokes diffusion coefficient for an infinite fluid [52]. Similarly, local diffusion and viscosity measurements on single membrane proteins can reveal the nature of their local environment. By tethering a membrane protein to a microsphere in a PFM OT [61], Pralle et al. were able to probe the dynamics and size of lipid raft structures within the membrane. The results were consistent with lipid rafts structures of size 26 ± 13 nm that were stable over a time scale of minutes.

A theoretical description of the operation of the photonic force microscope is a challenging task, not least in the calculation of the interaction potential between the micro- or nanoscale probe particle and the strongly focussed laser beam of the OT. However, significant recent progress has been made [19, 32] in the calculation of trapping potentials of both spherical and nonspherical objects and reliable models for particle tracking by BFP interferometry are now available [36, 37].

2.4.2 Bio-Force Sensing at the Nanoscale

As described above, a well-calibrated OT equipped with nanometre resolution particle position tracking is capable of measuring forces with magnitudes of a few piconewtons a resolution in sub-piconewton regime. This range of forces is relevant to a number of soft matter, colloidal and, in particular, biological systems ranging from the single molecule to the cellular level (for reviews, see e.g. [62, 63]). Significant examples of this method of bio-force sensing include measurements of the force generated by molecular motors and intermolecular binding and adhesion forces.

In a typical OT force measurement experiment, the displacement of a trapped particle in an OT under an externally applied force is measured (by the BFP interferometry technique), and if the trap spring constant is also known then the magnitude of the force can be determined (provided maximum displacements remain within the linear part of the force–distance curve). In bonding and adhesion force measurements usually the trapped particle will be a latex sphere, functionalised to target the system under investigation. Force is then applied by increasing the separation between the optical trap position and the target system, which displaces the latex sphere ‘handle’ from equilibrium. Applied force is increased until catastrophic rupture of the bond being tested occurs, and the sphere returns rapidly to the trap centre. Particle tracking reveals the movement of the latex sphere, from which the binding forces can be extracted as the optical trapping force on the sphere at its maximum displacement is assumed to be the force required to break the bond. A further advantage of this technique is that the loading rate, that is, the rate at which force is applied to the bond, can be varied by changing the speed at which the trap is withdrawn, potentially allowing the dynamic response of a bond to applied force to be determined [64].

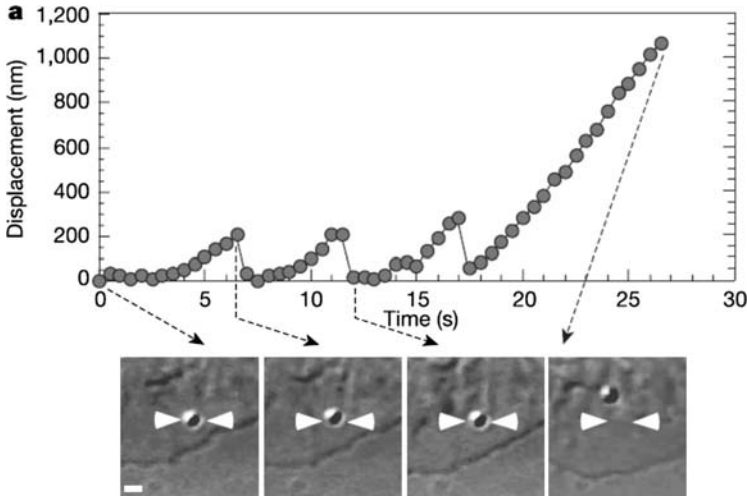


Fig. 2.12. The optically trapped bead is tethered to cellular integrins and trafficked rearwards, pulling it away from the trap centre. At a displacement from equilibrium corresponding to a load of 2 pN the talin mediated slip bond is broken and the bead returns quickly to the trap centre (reprinted from [65])

An example of this measurement technique is shown in Fig. 2.12 for the talin mediated slip bond between the cytoskeleton of migrating mouse fibroblasts and extracellular matrix (fibronectin) [65]. In this experiment, 0.64- μm -diameter beads were coated with fibronectin FN1117–10 trimer at low density and adhered to cellular integrins. The cytoskeleton–integrin–fibronectin linkage was stressed as the bead was trafficked rearwards by cytoskeletal flow away from the static laser trap at a rate of 60 nm s^{-1} , and breaking events occurred predominantly at an applied force of 2 pN corresponding to the force required to rupture the cytoskeleton–integrin linkage. This is supported by a probabilistic model [66] which finds that optical forces break only a single receptor cytoskeleton bond.

A second fruitful application for OT as a force measuring tool has been in the study of motor proteins such as kinesin [67]. The two-headed kinesin molecule is driven by ATP hydrolysis to step along microtubules. The length of the step and the force generated by the kinesin were revealed by adapting an OT to become a ‘molecular force clamp’. In this technique, a 0.5 μm silica bead was attached to a kinesin molecule and held in an optical trap. A feedback loop was implemented that maintained a fixed displacement of the bead from trap centre as the kinesin stepped along an immobilised microtubule, and particle tracking of the rapped silica bead revealed discrete displacement events with a magnitude of 8 nm, the length of the kinesin motor’s step. Using the force clamp, it was possible to alter the load against which the motor was forced to operate until the stepping velocity was reduced to zero – the ‘stalling force’.

This was found to be a function of the concentration of ATP, with a stalling force of up to 7 pN at the highest concentrations.

In contrast to the force clamp method, thermal noise probing keeps the applied optical forces to a minimum, thus allowing the 3D Brownian motion to dominate [68]. As discussed above, in a particle tracking experiment, the displacement frequency isosurface of a spherical particle in an OT is ellipsoidal, representing the 3D potential energy well of the trap, $V(x) = \sum \frac{1}{2} k_i x_i^2$. For a particle with restricted mobility, the volume explored by the particle changes, and from these changes a picture of the new energy landscape can be built up. In [68], this technique was applied to the kinesin motor protein by analysing the dynamics of a silica bead tethered to a microtubule via kinesin. As shown in Fig. 2.13 the 3D energy surface is radically different for a tethered bead,



Fig. 2.13. 3D thermal force probing of the kinesin motor protein. On the left is a representation of the volume explored by the particle in the optical tweezers. On the right are two examples of those of a particle attached to a kinesin molecule on a microtubule. From the bead fluctuation distribution, the kinesin length, orientation and rigidity can be deduced (reprinted from [68])

resembling a spherical cap. From this, it was deduced that the kinesin molecule behaves as a relatively rigid rod that pivots about an anchor point on the microtubule. Further experiments were able to elucidate the molecules stiffness to stretching and compression and locate the most flexible part of the molecule to the region close to the motor domain.

2.5 Raman Tweezers

In today's world of increasingly complex and refined analytical techniques, spectroscopy has maintained its place at the forefront. Raman spectroscopy, in particular, has proven especially useful in providing detailed analysis of a staggering variety of samples. Nowadays, it is routinely used for chemical/physical measurements in materials science, geology, microelectronic, biology [69]. Raman spectroscopy is able to detect and analyse extremely small molecular objects with high resolution. When combined with optical confocal microscopy, Raman spectroscopy allows for analysis and chemical imaging [70] with spatial resolution limited only by diffraction (~ 250 nm in the visible range) [71]. A derivative of Raman spectroscopy, termed Raman Tweezers (RT), has allowed for an even greater degree of analytical capability. RT use OT to suspend and manipulate micro/nanoparticles without direct contact, so that the molecular Raman spectra may be recorded while in their most natural state. As such, the spectra collected are more reflective of the true nature of the molecule under study and therefore of more significance. RT find natural application for biological analysis of a wide range of micron size particles, including living cells, viruses, bacteria, organelles, etc. [72–77]. Applications of RT in carbon nanotechnology include the possibility to analyse single nanotubes bundles dispersed in aqueous solutions and study the basic physical properties of individual, chirality-sorted nanotubes. A more intriguing application of the RT is given by the possibility to perform sub-diffraction near-field Raman imaging with optically trapped microparticles or to manipulate metallic nanoparticles to locally create 'hot' dimers for local surface-enhanced Raman spectroscopy, a methodology that may pave the way for controlled single molecule Raman analysis in situ.

2.5.1 The Raman Effect

When light is scattered from a molecule, most photons are elastically diffrused, that is, they have the same energy of the incident ones. However, a tiny fraction of the light (~ 1 in 10^7 photons) is scattered at energies slightly different from, and usually lower than, the incident light energy. The process leading to this inelastic scattering is referred to as the linear Raman effect. If the incident light beam has frequency ω_L , and the scattered radiation has frequency ω_R , the difference $\Delta\omega = \omega_L - \omega_R$ is called Raman shift. Raman spectroscopy relies upon the detection of inelastically scattered light from the

sample as a result of illumination of the sample by a laser beam. The energy shift provides information on the vibrational and rotational energies of molecular bonds and on chemical species in which those bonds are present. Raman spectra can be considered, therefore, as the chemical fingerprints of molecules and of their interaction with the environment. In the context of biological studies, this enables chemical identification of specific components in cells, tissues and organisms. In carbon-based nanotechnology, Raman spectroscopy is routinely applied to retrieve information on the diameter, electronic type (metallic semiconductor), number of walls of carbon nanotubes [79] or to determine the number of graphite layers in graphene samples [80].

2.5.2 Experimental Configuration

As shown in Fig. 2.14, there are four possible RT configurations. In the simplest one (A), a single laser beam allows for both trapping and Raman excitation. The back-scattered Raman light is collected by the same objective. This setup has been commonly employed among the RT community [81, 82]. Using two separate beams for trapping and Raman excitation [83, 84] with a single objective (B) has the advantage that the wavelength and the power of the trapping beam and the Raman beam can be adjusted independently to optimise the functionality. This technique also allows one to manipulate independently the trapped object and record Raman spectra from different regions. In (C), two counter-propagating laser beams are used for trapping and Raman excitation, each of the beams being focussed on the particle by a separate objective. The forward-scattered Raman emission is collected here. (D) Two counter-propagating optical beams have the advantage of creating

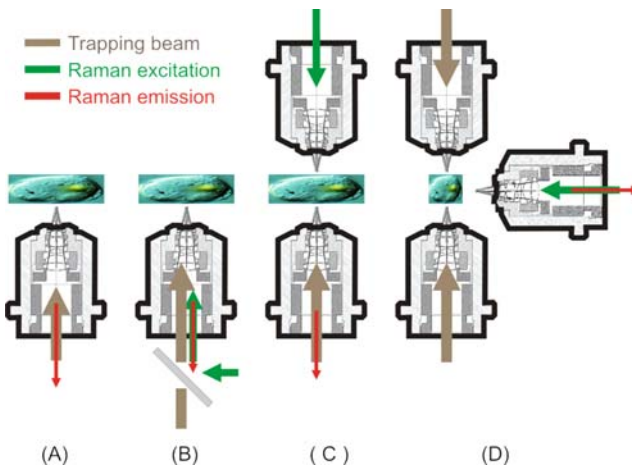


Fig. 2.14. The main four Raman tweezers configurations in an inverted microscope. *Brown lines* indicated the trapping beam, *green lines* the Raman excitation laser, *red lines* the Raman emission

a stable optical trap also using objectives with long working distance and relatively small numerical apertures, or even the divergent fields from two opposing optical fibres. This in particular permits to hold and manipulate large objects, like cells. Raman scattering is excited by a third beam focussed on the trapped particle through an individual objective, which also collects Raman signal [85].

Figure 2.15 shows a more detailed schematic of an RT apparatus employing the same beam to trap the particle and excite the sample for Raman analysis. Laser beams at different wavelengths can be used, depending on the specific applications. Visible lasers (454, 488, 514.5, 633 nm) are well suited for materials science applications, while NIR lasers (785, 830 nm) are better for biological samples because natural fluorescence is not excited. The laser beam (*brown lines*), after passing an interference filter for plasma lines removal, is expanded to a diameter of 6–8 mm in order to fill the entrance pupil of the objective. A notch filter, tuned at the laser emission wavelength, is used to first reflect back the laser beam towards the high NA microscope objective and, then, to cut out the elastic scattering component of the backscattered radiation which is collected by the same objective. The sample is illuminated

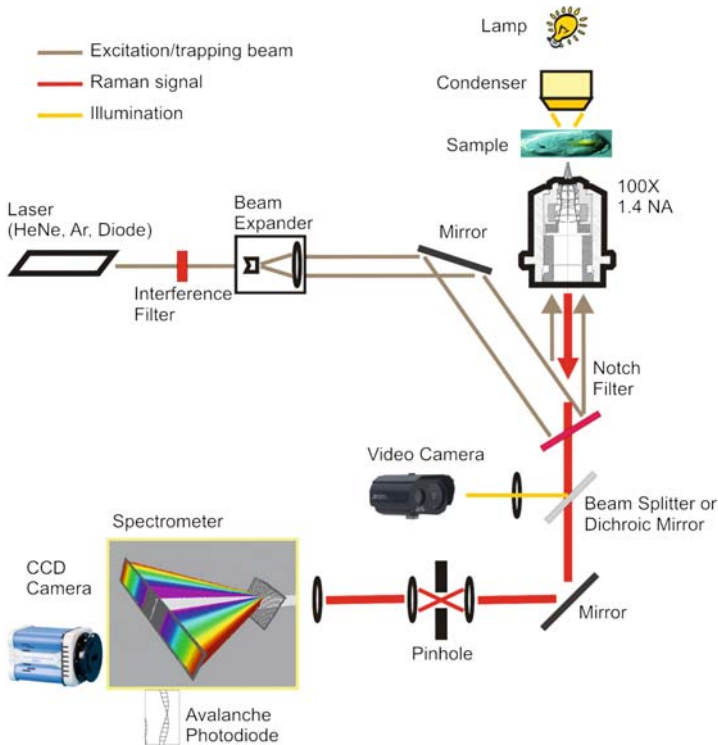


Fig. 2.15. Schematic of a model of Raman tweezers using the same beam to trap and excite the Raman signal (see text for details)

by a lamp or a light-emitting diode whose light is focussed on the sample surface by a condenser. Visual inspection of the trapped particle is accomplished through a video camera. The inelastically scattered radiation, after passing a confocal hole, is focussed on the slits of the spectrometer and then detected by a CCD camera or a single photon counting silicon avalanche photodiode.

2.5.3 Applications

Thalassemia is a blood disease quite diffuse in the Mediterranean Sea region. The fundamental abnormality in thalassemia is an impaired production of either the alpha or beta haemoglobin chain. Unlike alpha thalassemia, beta thalassemia rarely arises from the complete loss of a beta globin gene. The beta globin gene is present, but produces little beta globin protein. The severity of beta thalassemia depends on how badly one or both genes are affected. If both genes are affected, the result is moderate to severe anaemia, that is a lower than normal number of functional red blood cells in blood. A RT system has been employed by the group of A. Sasso to examine the oxygenation capability of β -thalassemic red blood cells [86], that is, their functionality. Figure 2.16 (*top plot*) shows the Raman spectrum of an optically trapped healthy red blood cell in comparison with the one from a heterozygous β -thalassemic cell (*bottom plot*). Raman bands characteristic of oxygenated Hb (oxyHb) are strongly depressed for the β -thalassemic ones (indicated by the *arrows*). Moreover, numerous Raman bands, affected by oxygenation condition, are energy-shifted (*dashed lines*). Both experimental outcomes clearly

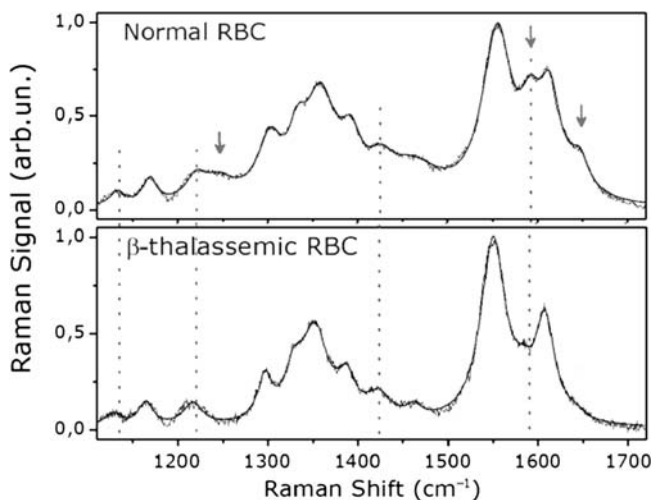


Fig. 2.16. Raman spectra of normal (*top plot*) and β -thalassemic RBC (*bottom plot*). The arrows indicate peaks affected by intensity changes, the dashed lines indicate energy shifts of the peaks positions (reprinted from [85])

demonstrate a lower efficiency of β -thalassemic cell in carrying out their natural role, namely, oxygen transportation from lungs to the entire organism. These results demonstrate that RT may have enormous potential for the monitoring of blood diseases and their response to drug therapies.

SWNTs, simply made by rolling-up a single graphene sheet, are currently at the centre of intense research activity due to their outstanding optical, thermal and mechanical properties. Currently available SWNT synthesis methods yield nanotubes with uncontrolled diameter, chirality and, consequently, electronic properties (metallic/semiconducting, bandgap energy, etc.). A further disadvantage of as-produced SWNTs is their natural tendency to form bundles [87] which almost completely quenches the photoluminescence emission from semiconducting nanotubes. Solutions to these problems take advantage of wet chemistry and biology methods to, first disperse and individualise and, then, sort SWNTs by electronic type. In this context, experimental methods such as RT are of utmost importance as a diagnostic tool to assess the dispersion degree and the chirality homogeneity of individual bundles in liquid, as well as for the selective engineering of nanotubes bundles. Figure 2.17a, b shows, respectively, the optical pictures of a free floating and of an optically trapped nanotube bundle dispersed in a H_2O /Taurodeoxycholate solution [88]. Optical forces align the bundle parallel to the beam propagation direction (B). Trapping and excitation of the Raman signal are carried out at 633 nm. CoMoCAT nanotubes from SouthWest NanoTechnologies have been used (lot SG-000–0007), featuring a narrow diameter distribution around 0.8 nm. The Raman spectrum of the trapped bundle (C) shows the three main spectral features of SWNTs: the radial breathing modes (RBM) in the 100–350 cm^{-1} spectral range, the D and the G bands (1,320 and \sim 1,590 cm^{-1}) and the 2D peak at 2,620 cm^{-1} . The G band shape, in particular, indicates that the trapped bundle is prominently constituted of semiconducting nanotubes (\sim 90%), while the presence of several peaks in the RBM region indicates that the bundle is composed of several nanotubes with different diameter ($\omega_{\text{RBM}} \sim 1/d$). Using real-time acquisition of the RBM Raman scattering, Rodgers et al. [15] were able to monitor the dynamics of HiPCO SWNT solutions in the optical trap over a period of several minutes. As reported in Fig. 2.17d, four main RBM peaks were observed, out of which peak labelled with ‘A’ clearly shows more fluctuation with the time than others together with a steady and marked increase in the amplitude in clear contrast to the neighbouring modes. The authors attribute the independent behaviour of RBMs to the fact that the flux of nanotubes through the focal volume varies between species, implying that the larger SWNTs (corresponding to lower frequency RBMs) experience a different force than the smaller tubes, and that a selective aggregation of a single chirality in the optical trap is taking place, where tubes of only one RBM are seen to be affected by the optical gradient force. This study, in particular, foresees the technologically relevant possibility to produce engineered nanotubes bundles, enriched with a single chirality, via selective aggregation in an

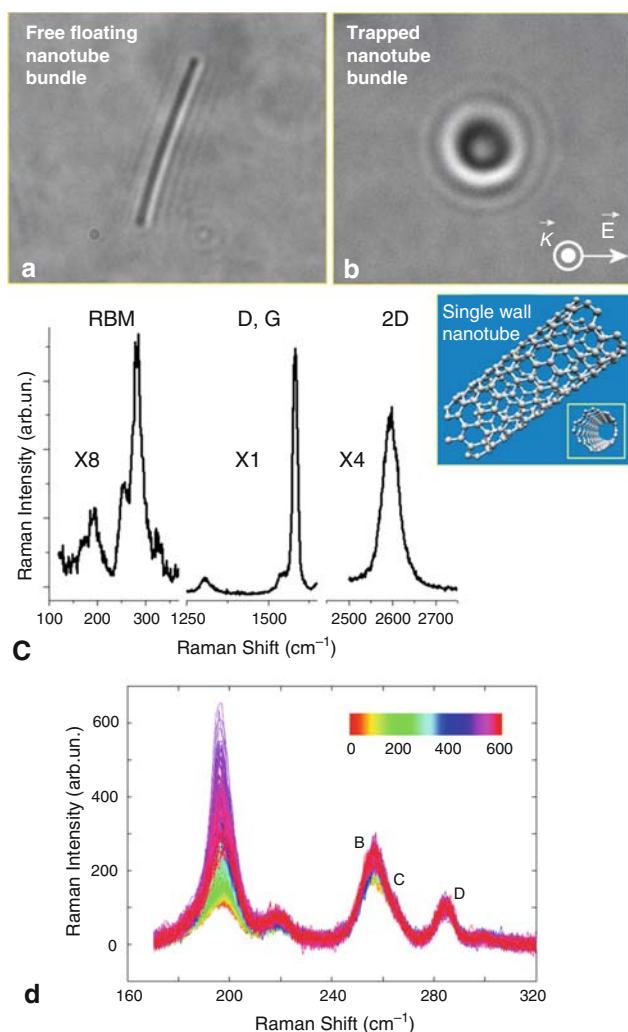


Fig. 2.17. (a, b) Pictures of a free floating and of an optically trapped CoMoCAT nanotube bundle, aligned with the beam propagation direction. (c) Raman spectrum of the trapped bundle showing the chiral heterogeneity and semiconducting prevalence of nanotubes in the bundle. (d) RBM dynamics of trapped HiPCO nanotubes taken at intervals of 2 s, over a period of 10 min. The selective increase of the mode at 196 cm^{-1} indicates selective aggregation of nanotubes with a single chirality in the optical trap (redrawn from [15])

optical trap. The strong wavelength dependence, moreover, can be exploited for selective enhancement of different chiralities.

To extend Raman imaging to nano-materials, extensive effort has been made in order to reduce the excitation laser spot below the diffraction limit by

using near-field optical microscopy techniques (NanoRaman). NanoRaman as well as tip-enhanced Raman scattering allows one to tailor the spatial resolution down to the nanometre scale, through the exploitation of the confined and enhanced fields in proximity of sub-wavelength holes or metallic nanoparticles (can be tips, nanorods, nanospheres) [89]. Near-field microscopy operation in liquid is, however, technologically very difficult, and therefore Kasim et al. [90] have developed a new approach to carry out sub-diffraction near-field Raman imaging in combination with RT, by exploiting the field focussing capabilities of optically trapped latex beads. As shown in Fig. 2.18(a) a single beam RT apparatus has been used with a 532 nm laser beam ($P = 3 \text{ mW}$) focussed via a water immersion objective (60X, NA 1.2) to trap a 3- μm -diameter polystyrene bead in de-ionised water. The bead is pushed down and is in contact with the sample surface which is scanned by a piezo-stage. The back-scattered Raman emission from the sample is collected by the same microsphere-objective

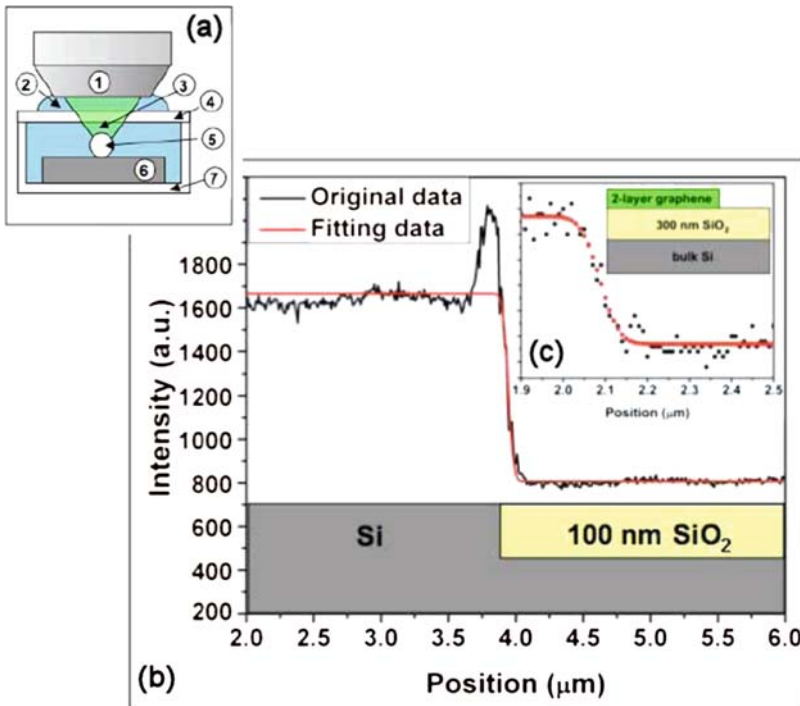


Fig. 2.18. (a) Schematic of the apparatus for Near-Field Raman imaging using latex microspheres. (1) water immersion microscope objective, (2) index matching water, (3) focussed beam, (4) coverslip, (5) trapped latex bead, (6) sample, (7) piezoelectric stage. (B) Silicon Raman intensity versus position with the best fit to determine the effective laser spot size (FWHM = 78 nm). Inset shows the integrated G-band intensity versus position of a 2-layer graphene sheet. The calculated spot size is 84 nm (redrawn from [90])

assembly and detected conventionally. Figure 2.18b shows a scan of the microsphere across a Si/SiO₂ structure with 10 nm step size. The black line reports the line profile of the Si Raman intensity versus position which, after fitting, gives an effective spot-size of 78 nm (FWHM). A further scan carried out on the edge of a very thin 2-layer graphene sheet on SiO₂/Si substrate (C) gives a spot size 84 nm. Although spatial resolution of this method (80 nm) is far from what has been achieved by NanoRaman (10 nm), the simplicity and reproducibility of this approach to achieve high-resolution Raman imaging has the potential to make it attractive to be used for large-scale applications in nano-science and nano-technology.

Nanoparticles composed of noble metals exhibit an enhanced optical interaction with visible light due to the resonant excitation of collective electron oscillations, the so-called localised surface plasmons (LSP's) [91], which are important for the development of novel nanoscopic bio/chemosensors [92]. LSP excitation leads to enhanced optical fields near the nanoparticle surface. This phenomenon is the basis for surface enhanced Raman scattering (SERS), which in some cases allows for vibrational spectroscopy down to the single molecule detection limit [93]. Particle dimers, particularly, enable additional strong field enhancement in the gap regions between the particles [94]. Svedberg et al. [95] have employed an RT setup to manipulate metallic nanoparticles and create 'hot' dimers capable to induce a strong SERS signal of thiophenol molecules dispersed in the liquid at a concentration of 10 μM. Figure 2.19a shows the optical images of two separated silver nanoparticles, one immobilised on the coverslip (I) and one optically trapped (T). Neither of these particles yield any detectable SERS signal [Fig. 2.19c (*blue line*)]. When the trapped particle is brought into near-field contact with the immobilised particle (Fig. 2.19b), a clear SERS spectrum of the thiophenol appears

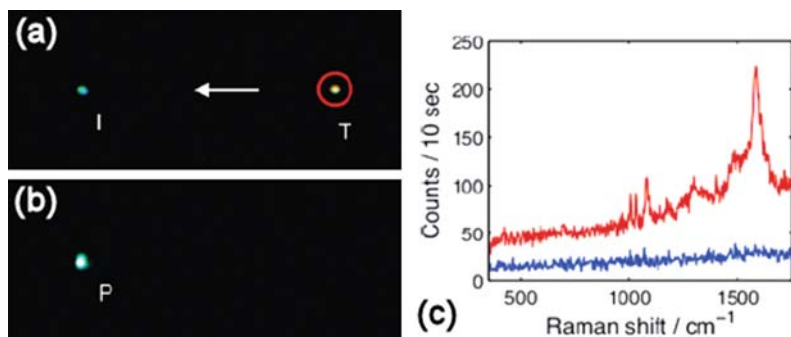


Fig. 2.19. (a) Dark-field images of a trapped (labelled with T) and an immobilised (labelled with I) Ag nanoparticles. (b) When brought into near-field contact, the particle dimer (labelled with P) shows a strongly enhanced SERS signal (c, *red line*) with respect to the signal measured from the individual particles (C, *blue line*) (redrawn from [95])

[Fig. 2.19c (*red line*)], indicating a strong near-field inter-particle interaction with the excitation of a LSP gap mode between the nanoparticles. In particular, this experiment shows that once the optically trapped particle is within 250 nm distance (the typical trap lateral dimension) from the immobilised particle, it has a high probability to jump into the near-field coupling zone (<10 nm), resulting in a strong increase in the SERS intensity, constituting a strong experimental indication of an optically induced aggregation of metal nanoparticles.

2.6 Conclusions

OT have been demonstrated to be a powerful tool for the manipulation of a number of important colloidal, soft matter and biological systems. The combination of optical trapping with fast particle tracking methods has broadened the scope of OT to become a scanning probe *photonic force microscope* capable of imaging with sub-optical wavelength spatial resolution, and force spectroscopy with femtonewton sensitivity. This range of parameters is particularly attractive for applications in the life sciences, and a number of experiments have already proved the utility of PFM in imaging soft systems that are outside the capabilities of AFM.

The outlook for future developments in OT and PFM is particularly exciting if the technique can be combined with other ultra-high-resolution spectroscopic techniques such as micro-, nano- and surface enhanced Raman spectroscopy and near-field optical methods that can reveal information on single biological molecules and engineered nanostructures in situ.

Acknowledgement

A. Camposeo, B. Hoogenboom and G. Memoli are acknowledged for proof-reading and suggestions.

References

1. A. Ashkin, J.M. Dziedzic, J.E. Bjorkholm, S. Chu *Opt. Lett.* **11** 288 (1986)
2. A. Ashkin, J.M. Dziedzic *Science* **235** 1517 (1987)
3. D.G. Grier *Nature* **424** 810 (2003)
4. A. Ashkin *Optical Trapping and Manipulation of Neutral Particles Using Lasers* (World Scientific, Singapore, 2007)
5. S.M. Block, D.F. Blair, H.C. Berg *Nature* **338** 514 (1989)
6. A. Ashkin, K. Schutze, J.M. Dziedzic, U. Euteneuer, M. Schliwa *Nature* **348** 346 (1990)
7. E.L. Florin, A. Pralle, J.K.H. Hörber, E.H.K. Stelzer *J. Struct. Biol.* **119** 202 (1997)

8. R. Agarwal, K. Ladavac, Y. Roichman, G. Yu, C. Lieber, D. Grier *Opt. Express* **13** 8906 (2005)
9. P.J. Pauzauskie, A. Radenovic, E. Trepagnier, H. Shroff, P. Yang, J. Liphardt *Nat. Mater.* **5** 97 (2006)
10. Y. Nakayama, P.J. Pauzauskie, A. Radenovic, R.M. Onorato, R.J. Saykally, J. Liphardt, P. Yang *Nature* **447** 1098 (2007)
11. S. Tan, H.A. Lopez, C.W. Cai, Y. Zhang *Nano Lett.* **4** 1415 (2004)
12. J. Plewa, E. Tanner, D. Mueth, D. Grier *Opt. Express* **12** 1978 (2004)
13. J. Zhang, H.I. Kim, C.H. Oh, X. Sun, H. Lee *Appl. Phys. Lett.* **88** 053123 (2006)
14. O.M. Maragò et al. *Physica E* **40** 2347 (2008)
15. T. Rodgers, S. Shoji, Z. Sekkat, S. Kawata *Phys. Rev. Lett.* **101** 127402 (2008)
16. O.M. Maragò, P.H. Jones, V. Scardaci, F. Bonaccorso, P.G. Gucciardi, A.G. Rhozin, A.C. Ferrari *Nano. Lett.* **8** 3211 (2008)
17. F. Borghese, P. Denti, R. Saija *Scattering from Model Nonspherical Particles*, 2nd edn. (Springer, Berlin, 2007)
18. M.I. Mishchenko, L.D. Travis, A.A. Lacis *Scattering, Absorption and Emission of Light by Small Particles* (Cambridge University Press, New York, 2002)
19. A. Rohrbach *Phys. Rev. Lett.* **95** 168102 (2005)
20. A.R. Zakharian, P. Polynkin, M. Mansuripur, J.V. Moloney *Opt. Express* **14** 3660 (2006)
21. F. Borghese, P. Denti, R. Saija, M.A. Iatì *Opt. Express* **15** 11984 (2007)
22. A. Ashkin *Biophys. J.* **61** 569 (1992)
23. A. Ashkin *Proc. Natl. Acad. Sci. U.S.A.* **94** 4853 (1997)
24. L. Novotny, B. Hecht *Principles of Nano-Optics* (Cambridge University Press, Cambridge, 2006)
25. C.J. Foot *Atomic Physics* (Oxford University Press, Oxford, 2005)
26. B. Richards, E. Wolf *Proc. R. Soc. Lond. B Biol. Sci.* **253** 358 (1959)
27. P.C. Waterman *Phys. Rev.* **D4** 825 (1971)
28. G. Knöner, S. Parkin, T.A. Nieminen, N.R. Heckenberg, H. Rubinsztein-Dunlop *Phys. Rev. Lett.* **97** 157402 (2006)
29. J.D. Jackson *Classical Electrodynamics*, 2nd edn. (Wiley, New York, 1975)
30. R. Saija, M.A. Iatì, F. Borghese, P. Denti, S. Aiello, C. Cecchi-Pestellini *Astrophys. J.* **559** 993 (2001)
31. R. Saija, M.A. Iatì, P. Denti, F. Borghese, A. Giusto, O.I. Sindoni *Appl. Opt.* **42** 2785 (2003)
32. F. Borghese, P. Denti, R. Saija, M.A. Iatì, O.M. Maragò *Phys. Rev. Lett.* **100** 163903 (2008)
33. P. Yang *Nature* **425** 243 (2003)
34. P.H. Jones, O.M. Maragò, E.P.J. Stride *J. Opt. A: Pure Appl. Opt.* **9** S278 (2007)
35. A.H. Mack, M.K. Trias, S.G.J. Mochrie *Rev. Sci. Instrum.* **80** 016101 (2009)
36. F. Gittes, C.F. Schmidt *Opt. Lett.* **23** 7 (1998)
37. A. Pralle, M. Prummer, E.L. Florin, E.H.K. Stelzer, J.K.H. Hörber *Microsc. Res. Tech.* **44** 378 (1999)
38. J.E. Molloy, M.F. Padgett *Contemp. Phys.* **43** 241 (2002)
39. M. Polin, K. Ladavac, S. Lee, Y. Roichman, D.G. Grier *Opt. Express* **13** 5831, (2005)
40. G. Whyte, G. Gibson, J. Leach, M. Padgett *Opt. Express* **14** 12497 (2006)
41. G Gibson et al. *J. Opt. A: Pure Appl. Opt.* **10** 044009 (2008)
42. K.T. Gahagan, G.A. Swartzlander Jr *Opt. Lett.* **21** 827 (1996)

43. K. Ladavac, D.G. Grier *Opt. Express* **12** 1144 (2004)
44. A. Einstein *Ann. Phys.* **17** 549 (1905)
45. Y. Han, A.M. Alsayed, M. Nobili, J. Zhang, T.C. Lubensky, A.G. Yodanis *Science* **314** 626 (2006)
46. A. Rohbach, E.H.K. Stelzer *Appl. Opt.* **41** 2494 (2002)
47. K. Berg-Sorensen, H. Flyvbjerg *Rev. Sci. Instrum.* **75** 594 (2004)
48. Y. Seol, A.E. Carpenter, T.T. Perkins *Opt. Lett.* **31** 2429 (2006)
49. J.C. Meiners, S.R. Quake *Phys. Rev. Lett.* **82** 2211 (1999)
50. S. Henderson, S. Mitchell, P. Bartlett *Colloids Surf A* **190** 81 (2001)
51. W.T. Coffey, Y.P. Kalmykov, J.T. Waldron *The Langevin Equation* 2nd edn. (World Scientific, Singapore, 2004)
52. J. Happel, H. Brenner *Low Reynolds Number Hydrodynamics*. (Springer, Berlin, 1981)
53. S. Broersma *J. Chem. Phys.* **74** 6989 (1981)
54. L.P. Ghisla, W.W. Webb *Opt. Lett.* **18** 1678 (1993)
55. G. Binnig, C.F. Quate, C. Gerber *Phys. Rev. Lett.* **56** 930 (1986)
56. E.L. Florin, J.K.H. Hörber, E.H.K. Stelzer *Appl. Phys. Lett.* **69** 446 (1996)
57. E.L. Florin, A. Pralle, E.H.K. Stelzer, J.K.H. Hörber *Appl. Phys. A* **66** S75 (1998)
58. A. Rohbach, C. Tischer, D. Neumayer, E.L. Florin, E.H.K. Stelzer *Rev. Sci. Instrum.* **75** 2197 (2004)
59. C. Tischer, S. Altmann, S. Fisinger, J.K.H. Hörber, E.H.K. Stelzer, E.L. Florin *Appl. Phys. Lett.* **79** 3878 (2001)
60. A. Pralle, E.L. Florin, E.H.K. Stelzer, J.K.H. Hörber *Appl. Phys. A* **66** S71 (1998)
61. A. Pralle, P. Keller, E.L. Florin, K. Simons, J.K.H. Hörber *J. Cell Biol.* **148** 997 (2000)
62. S.C. Kuo *Traffic* **2** 757 (2001)
63. A.D. Mehta, M. Rief, J.A. Spudis, D.A. Smith, R.M. Simmons *Science* **283** 1689 (1999)
64. R. Merkel, P. Nassoy, A. Leung, K. Ritchie, E. Evans *Nature* **397** 50 (1999)
65. G. Jiang, G. Gianone, D.R. Critchley, E. Fukumoto, M.P. Sheetz *Nature* **424** 334 (2003)
66. O. Thoumine, J.J. Meister *J. Theor. Biol.* **204** 381 (2000)
67. K. Visscher, M.J. Schnitzer, S.M. Block *Nature* **400** 184 (1999)
68. S. Jeney, E.H.K. Stelzer, H. Grubmüller, E.L. Florin *Chem. Phys. Chem.* **5** 1150 (2004)
69. J.R. Ferraro, K. Nakamoto, C.W. Brown (eds.), *Introductory Raman Spectroscopy* (Elsevier, Amsterdam, 2003)
70. G. Turrell, J. Corset (eds.), *Raman Microscopy. Developments and Applications*. (Elsevier, Amsterdam, 1996)
71. E. Abbe *Arch. Mikroskop. Anat.* **9** 413 (1873)
72. C. Xie, M.A. Dinno, Y.Q. Li *Opt. Lett.* **27** 249 (2002)
73. J.W. Chan, A.P. Esposito, C.E. Talley, C.W. Hollars, S.M. Lane, T. Huser *Anal. Chem.* **76** 599 (2004)
74. C. Xie, C. Goodman, M.A. Dinno, Y.Q. Li *Opt. Express* **12**, 6208 (2004)
75. H. Tang, H. Yao, G. Wang, Y. Wang, Y.Q. Li, M. Feng *Opt. Express* **15** 12708 (2007)
76. J.W. Chan, A.P. Esposito, C.E. Talley, C.W. Hollars, S.M. Lane, T. Huser *Anal. Chem.* **76** 599 (2004)

77. K.E. Hamden, B.A. Bryan, P.W. Ford, C. Xie, Y.Q. Li, S.M. Akula *J Virol Methods* **129** 145 (2005)
78. A.C. De Luca, G. Rusciano, R. Ciancia, V. Martinelli, G. Pesce, B. Rotoli, L. Selvaggi, A. Sasso *Opt. Express* **16**(11) 7943 (2008)
79. A.C. Ferrari, J. Robertson (eds.), *Phil Tran R Soc Lond A* 362 2477–2512 (2004)
80. A.C. Ferrari, J.C. Meyer, V. Scardaci, C. Casiraghi, M. Lazzeri, F. Mauri, S. Piscanec, D. Jiang, K.S. Novoselov, S. Roth, A.K. Geim *Phys Rev Lett* **97** 187401 (2006)
81. K.E. Hamden, B.A. Bryan, P.W. Ford, C. Xie, Y.Q. Li, S.M. Akula *J. Virol. Methods* **129** 145 (2005)
82. C. Xie, C. De, Y.Q. Li *Opt. Lett.* **30** 1800 (2005)
83. R. Gessner, C. Winter, P. Rosch, M. Schmitt, R. Petry, W. Kiefer, M. Lankers, J. Popp *Chem. Phys. Chem.* **5** 1159 (2004)
84. K. Ramser, K. Logg, M. Goksr, J. Enger, M. Kll, D. Hanstorp *J. Biomed. Opt.* **9** 593 (2004)
85. P.R.T. Jess, V. Garces-Chavez, D. Smith, M. Mazilu, L. Paterson, A. Riches, C.S. Herrington, W. Sibbett, K. Dholakia *Opt. Express* **14** 5759 (2006)
86. A.C. De Luca, G. Rusciano, R. Ciancia, V. Martinelli, G. Pesce, B. Rotoli, L. Selvaggi, A. Sasso *Opt. Express* **16** 7943 (2006)
87. A. Thess, R. Lee, P. Nikolaev, H. Dai, P. Petit, J. Robert, C. Xu, Y.H. Lee, S.G. Kim, A.G. Rinzler, D.T. Colbert, G.E. Scuseria, D. Tománek, J.E. Fischer, R.E. Smalley *Science* **273** 483 (1996)
88. P.G. Gucciardi et al. to be submitted
89. P.G. Gucciardi, S. Trusso, C. Vasi, S. Patanè, M. Allegrini, in *Applied Scanning Probe Methods V*, ed. by B. Bhushan, C. Fuchs, S. Kawata (Springer, New York, 2007), pp. 287–329
90. J. Kasim, Y. Ting, Y.Y. Meng, L.J. Ping, L.L. Jong, S.Z. Xiang *Opt. Express* **16** 7976 (2008)
91. K.L. Kelly, E. Coronado, L.L. Zhao, G.C.J. Schatz *Phys. Chem. B* **107** 668 (2003)
92. A.J. Haes; R.P.J. Van Duyne *Am. Chem Soc.* **124** 10596 (2002)
93. Kneipp K, Moskovits M, Kneipp H (eds.), *Surface Enhanced Raman Scattering* (Springer. Berlin, 2006)
94. A.M. Michaels; M. Nirmal; L.E.J. Brus *Am. Chem. Soc.* **121** 9932 (1999)
95. F. Svedberg, Z. Li, H. Xu, M. Käll *Nano. Lett.* **6** 2639 (2006)

Polarization-Sensitive Tip-Enhanced Raman Scattering

Pietro Giuseppe Gucciardi, Marc Lamy de La Chapelle, Jean-Christophe Valmalette, Gennaro Picardi, and Razvigor Ossikovski

Summary. Tip-Enhanced Raman Spectroscopy (TERS) is a very promising analytical technique for high sensitivity, high spatial resolution analysis of the physical and chemical properties of nanoscale materials including nanocrystals, biomolecules, carbon-based nanostructures and nanometer-size devices. Polarized TERS, in addition, offers novel opportunities for high contrast spectroscopy and imaging of semiconductor crystals and crystalline nanostructures. This chapter reviews the current state-of-the-art in polarization-sensitive TERS focusing the attention on the experimental implementations of the technique, on the light scattering properties of the metallic probes, on the Raman signal enhancement mechanisms and, finally, on the applications of this technique.

Key words: Scanning near-field optical microscopy, Tip-enhanced Raman spectroscopy, Polarized Raman spectroscopy, Nanoscale imaging.

Abbreviations

AFM	Atomic force microscopy
LSP	Localized surface plasmons
SERS	Surface-enhanced raman scattering
SNOM	Scanning near-field optical microscopy
STM	Scanning tunneling microscopy
TERS	Tip-enhanced raman scattering

3.1 Introduction

The measurement of the Raman spectrum of a crystal is one of the principal methods for obtaining information about its lattice structure. The general explanation of the Raman effect is well known from a long time [1]. Incident light of energy E_i interacts with the crystal to create or destroy one or

more vibration quanta (phonons) of energy E_V . The energy gained or lost by the lattice phonon is compensated by a decrease or increase of the scattered photon energy $E_i \pm E_V$. Measuring the energy of the Raman-scattered photons gives information on the lattice vibration frequencies and, to a final extent, on its arrangement at the atomic scale. Confocal microscopy provides Raman spectroscopy with the capability to perform chemical and structural imaging with a spatial resolution only limited by diffraction, that is, ~ 250 nm in the visible range. This value is far too large for the emerging needs in nanotechnology. A further drawback of Raman spectroscopy is its low efficiency. While fluorescence cross-sections of dye molecules are in the 10^{-14} – 10^{-16} cm² range, the Raman cross section of Rhodamine 6G at 633 nm is around 10^{-27} cm² [2], going down to $\sim 10^{-30}$ cm² for non-resonant molecules such as 2-bromo-2-methylpropane [3].

Tip-enhanced Raman scattering (TERS) [4–7] is a technique bringing together a spectacularly high sensitivity, up to single molecule detection, with few nanometers scale spatial resolution. Using a nanometer-sized metal particle, TERS results in a strong signal enhancement (up to 10^8) and a lateral resolution similar to the dimensions of the particle (down to 10 nm), achieving the best possible control on the characteristics and spatial localization of a single hot-spot producing the enhanced Raman signal [8, 9].

The following chapter focuses on TERS and, in particular, on polarization-sensitive TERS which is particularly important for applications on crystals where the Raman scattering can be strongly polarized [10]. The chapter is organized as follows. Section 3.2 revises the concept of polarized TERS with side illumination optics and describes methods for the production of metal tips. In Sect. 3.3, a model describing polarized Raman scattering from cubic crystals is introduced and in Sect. 3.4 a phenomenological model for the tip scattering properties is given. Section 3.5 affords the issue of local field depolarization induced by metallic tips, while Sect. 3.6 describes the consequences of depolarization in TERS. Finally, Sect. 3.7 reviews the polarized-TERS applications for imaging.

3.2 Tip-Enhanced Raman Spectroscopy

3.2.1 Concept and Advantages

The Tip-Enhanced Raman Scattering is a near-field technique based on the Raman enhancement locally induced by a metallic tip [8, 9, 11–13]. The field enhancement has the same origin than the one observed in Surface Enhanced Raman Scattering (SERS). That is, it exploits the specific optical properties of metallic nanostructures such as localized surface plasmons (LSP) resonances and the lightning rod effect [14–18]. When the metallic tip is illuminated and excited by a light source, a strong electromagnetic field enhancement can be achieved at the extremity of the tip. Such a field is extremely confined. The

confinement is of the same order of magnitude than the tip radius, that is, ~ 10 nm, and comparable to its decay length. By the excitation of the tip, a nanometric light source can be created largely more intense than the incident beam. Moreover, since the Raman signal is proportional to the fourth power of the local field [19], the Raman signal of a very few quantity of matter can be observed. In addition, by scanning the sample surface with the tip, it is possible to map the surface at the nanometric scale and using the Raman spectrum information, the physical and chemical analysis of the surface can be retrieved with nanometer spatial resolution [20, 21].

In TERS, the highest enhancement is obtained for an excitation polarization parallel to the tip axis [22]. Two different illumination configurations have been developed so far: the transmission mode [4, 6, 23] and the side illumination [7, 24–29] or reflection configuration, as shown in Fig. 3.1. The first one (a) uses an inverse microscope to illuminate the tip from below. The main advantage of this configuration is the possibility of using high numerical aperture objectives ($NA = 1.4$). The excitation and collection light spot can be smaller than 300 nm with high power density and less intense background. Unfortunately, the samples studied with this configuration must be deposited on transparent substrates.

In the reflection configuration (Fig. 3.1b), the tip is illuminated from the side. The tip and the microscope objective used for the illumination and the collection are together above the sample. This situation imposes that the objective cannot be very close with the sample surface, requiring the use of long working distance objectives. The limited numerical aperture ($NA < 0.7$) reduces the light collection capabilities compared to transmission configuration and enlarges the illumination spot to diameters larger than 500 nm. The main advantage of the side illumination configuration is the possibility to study any kind of samples: transparent and opaque, of any size and thickness. Furthermore, the polarization of the excitation can be easily controlled and can be rotated by using simple polarizers. We can then perform polarized Raman spectroscopy experiments which are extremely useful for the measure-

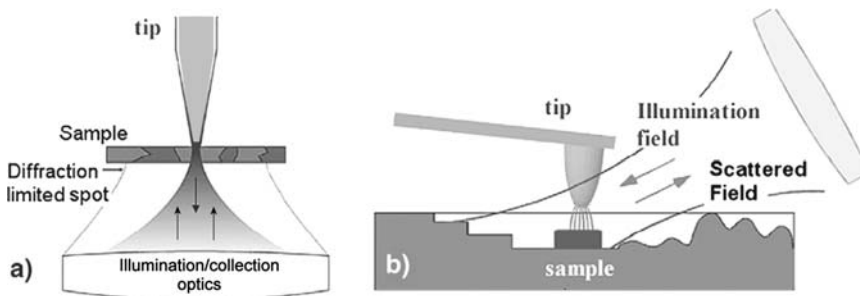


Fig. 3.1. Schematics of (a) the transmission and (b) side illumination TERS configurations. Redrawn from [29]

ment of strain in silicon wafers and the determination of the stress tensor in a Si crystal [30]. Beyond its main drawback (the use of low numerical aperture objectives), the side illumination has, therefore, unique advantages with important application possibilities in the microelectronic field and, more widely, in surface science, bio- and nanotechnologies.

3.2.2 Experimental Implementations of TERS with Side Illumination Optics

A TERS apparatus is nothing else than a Raman spectrometer optically coupled with a scanning probe microscope. The latter can either be a Scanning, Tunneling Microscope [7], an Atomic Force Microscope (AFM) [22] or a Shear-Force Microscope [6,26,31,32]. Nowadays commercial TERS models are available in the market working in both the transmission configuration and the side illumination. The TERS setup used for the measurements described in Sect. 3.6.1 consists (Fig. 3.2a) of a high-resolution Raman spectrometer (HORIBA Jobin Yvon) equipped with a long working distance $50\times$ objective (NA = 0.45) optically coupled in oblique (70°) backscattering geometry to a scanning probe microscope (PSIA) operating either in constant tunneling current mode or in atomic force contact mode. A $\lambda/2$ -waveplate is inserted in the optical path of the excitation laser beam ($\lambda = 514\text{ nm}$ or 633 nm) and an analyzer is set before the spectrometer, thus controlling the polarization states of the incident and Raman-scattered radiations. The settings of the waveplate and the analyzer can be varied manually with a $\pm 5^\circ$ accuracy. The TERS configuration used for measurements reported in Sect. 3.6.3 is depicted in Fig. 3.2b. We have developed a setup consisting of three parts. First, the AFM/SNOM. It is a commercial model (CP-Research from Park

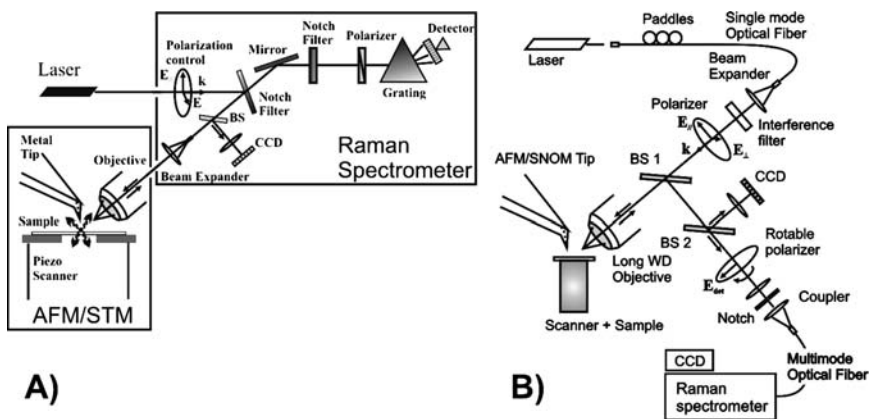


Fig. 3.2. Experimental implementations of TERS with side illumination optics (see text for details)

Scientific), implemented to work in shear-force mode with a quartz tuning fork. The microscope body mounts the sample and controls both the tip-to-sample distance, the sample's scan and the data acquisition. Second, the Raman spectrometer (Jobin Yvon, Labram). It is equipped with a double-notch or edge filters for elastic-light rejection and mounts either a CCD or a cooled PMT as detectors. Third, the optical coupling system (OCS), sketched in detail in Fig. 3.2b. It allows us to focus the laser light onto the tip-sample region and to collect the backscattered light. The system is mounted on an XYZ micrometer translation stage equipped with a further motorized stage for precise focusing on the tip apex region. Maximum flexibility in terms of choice of the excitation source is obtained by injecting the laser beam into a single mode optical fiber, subsequently coupled to the OCS. Here, after expansion (4:1), the beam passes through an interference filter (transmission 80%, bandwidth 3 nm) removing both the Raman emission from the silica and the plasma lines of the laser. Polarization control is accomplished by means of a linear polarizer whose axis is rotated between 0° and 90° , which permits to switch the polarization from p (parallel both to the plane of incidence and to the tip axis) to s (orthogonal both to the plane of incidence and to the tip axis). After passing a first beam splitter (BS1, $T = 40\%$, $R = 60\%$), the light is focused on the tip by means of long working distance microscope objectives (Mitutoyo, 20X, NA = 0.28, WD = 30.5 mm, or Mitutoyo 50X, NA = 0.42, WD = 20.5 mm). The backscattered radiation is reflected by BS1 and, after passing a second beamsplitter (BS2, $T = 90\%$, $R = 10\%$), impinges on a linear polarizer used to select between the p -polarized and the s -polarized radiation to be analyzed. The light passes through a notch filter centred at 633 nm and then is coupled into a multimode optical fiber (core diameter 50 μm) and going to the Raman spectrometer. BS2 partially reflects light to a CCD camera used for visual inspection and fine positioning of the laser spot to the tip apex. The use of a notch filter is very important, since we have observed that the elastic scattering, when coupled into the fiber, can induce a Raman signal (stimulated emission from the silica) that overcomes the signal from the sample that we want to measure (in the 100–600 cm^{-1} range). Coupling with the Raman spectrometer is simply accomplished by positioning the other extremity of the multimode fiber in front of the 10X microscope objective of the spectrometer.

3.2.3 Probes for Tip-Enhanced Raman Spectroscopy

Metallic tips are the core of TERS experiments, playing the dual role of confining the field on length scales comparable to the radius of curvature (5–50 nm) and acting, at the same time, as nanoantennas capable to efficiently collect the near-field radiation close to the sample and scatter it in the far-field. Strong research effort has been put in the 1990s to fabricate metallic tips suitable for STM. Nowadays, research has been focused in developing experimental procedures capable to reliably produce silver, gold and aluminum tips with

apical radii of curvature smaller than 50 nm, recognized to provide strong field enhancement via the excitation of localized plasmon resonances. Metallic tips are produced by either coating AFM tips or electrochemical etching techniques [33, 34]. Silicon nitride contact mode AFM probes (such as Veeco MSCCT) are well suited for TERS, after coating by thermal evaporation with a 10-nm-thick Cr layer followed by a 40-nm-thick Au one. The tip material (silicon nitride) is chosen to avoid overlapping of the *c*-Si sample Raman signal with the Raman response from the tip itself, present when using the more common silicon tips.

Alternatively, gold tips suitable for STM or shear-force TERS implementations can be produced by electrochemical etching [33, 35]. A gold ring of about 8-mm diameter made from 1-mm gold wire is used as counter-electrode, placed on the surface of the etching solution contained in a 10-ml recipient (about 3/4 of the height of the ring is immersed into the liquid). The etching solution used is a 1:1 mixture of concentrated (37%) HCl and ethanol. A gold wire of 0.25-mm diameter is immersed in the center of the ring. The immersed length is normally about 2–3 mm. A DC voltage is applied between the two electrodes (the tip being positive). The etching time is generally 10–12 min. To avoid over-etching of the tip, resulting in a larger apex radius, specific circuit automatically switches off the applied voltage as soon as the lower portion of the wire detaches, detected by a sudden fall in the current monitored by the circuit. This procedure is known to produce tips with a final radius of curvature in the range of 20–50 nm.

Electrochemical techniques suffer some drawbacks, such as the need for high currents (~ 1 A), the necessity that tips must be etched one by one, the oxidation and corrosion of the counter-electrode and finally the bubbling phenomena that, although reduced [34, 36], inevitably affect the final tip shape. Several technical expedients have been introduced to improve the sharpness of metallic tips [35, 37, 38]. Nevertheless, the best radii of curvature obtainable with such low-cost methods are always limited to a few tens of nanometer. Sharper probes require Focused Ion Beam milling, capable to tailor the radius of curvature down to the limit of 10 nm [18, 39].

To control and improve the reliability of the process, two new procedures have been recently introduced. The first one is an electrochemical process able to etch thin gold wires in an HCl/H₂SO₄ mixture (HCl: 0.5 M, H₂SO₄: 0.5 M) [36] which has demonstrated its good performances also when HCl/Ethanol mixtures (50%–50% in volume) are used. A constant DC voltage (12–18 V) is applied between a platinum electrode (the anode, which can be straight or ring-shaped) and the thin gold wire (the cathode) located at a few centimeter from the anode, measuring simultaneously the current (I) flowing through the circuit (see Fig. 3.3a). As soon as the etching reaction starts, currents as high as 2 A start to flow into the circuit (Fig. 3.3b), with a rise time of less than 1 s. The current slowly decreases in the first 20 s ($\Delta I/I \sim -25\%$), then it abruptly goes to zero in the subsequent 5 s. Tips with length in the millimeter range are usually produced. Apical radii of curvature of 30 nm or

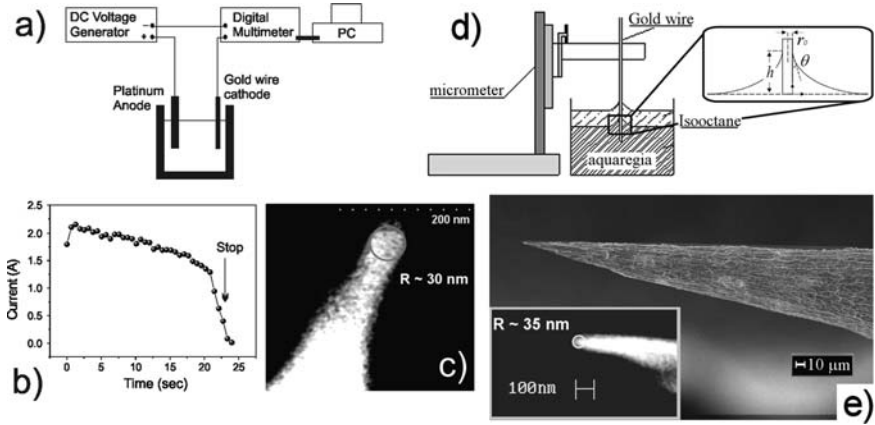


Fig. 3.3. Gold tip fabrication by electrochemical etching in $\text{HCl} + \text{H}_2\text{SO}_4$. (a) Setup. (b) Temporal behaviour of the current during the etching process. (c) SEM picture of the tip apex. Redrawn from [29]. (d) Sketch of the experimental apparatus for chemical etching in aqua regia. The inset shows a diagram of the meniscus at the *aqua regia*–isooctane interface. (e) SEM image of a chemically etched gold tip with a magnification of the apical region (inset). Redrawn from [40]

better (Fig. 3.3c) are obtained, provided that the etching process is stopped in the latest seconds, before the current has drop to zero. The best time to stop the etching process is at the so-called inflection point, where the second derivative vanishes, that is, $d^2I/dt^2 = 0$. It is important, in fact, to avoid the self-termination of the etching process ($I = 0$), stopping the etching before the wire is completely corroded thus preserving the sharp tip that is produced in bulk. In the opposite case, in fact, we end up with tips formed at the air HCl -surface meniscus which are typically blunt due to the strong instability of the meniscus height, caused by bubbling. The second procedure is based on the etching of thin gold wires in *aqua regia* [40]. Since it does not require the application of an electric field, it is easier to implement and can be extended to parallel etching of several wires. As shown in Fig. 3.3d, up to five wires can be etched simultaneously in 40 ml of fresh solution. The *aqua regia* is covered with a 3-mm layer of isooctane that serves for two purposes. First, it prevents relative changes in the aqua regia mixture preventing the loss of its potency. Second, it provides the necessary isolation in order to prevent the detrimental contact of the wire with the *aqua regia* vapors. The wires are immersed in the mixture by means of a micrometric stage attached to a rotation mount (not shown) to control the vertical alignment. Figure 3.3e shows the SEM picture of a typical chemically etched tip. A long taper ($\sim 600 \mu\text{m}$) is observed, ending with a tip featuring a radius of curvature of about 35 nm (inset of Fig. 3.3e). With this method, tips with radius of curvature $\leq 50 \text{ nm}$ are obtained with 40% yield, with the possibility to get tips with sub-30-nm radius of curvature, depending on the grains size of the starting wire. Thermal annealing in a

furnace at 800 °C for 6 h has been demonstrated as viable technique in order to reduce the grain size, yielding electrochemically etched tips featuring radii of curvature smaller than 25 nm with a yield higher than 30% [41].

3.3 Polarized Raman Scattering from Cubic Crystals

In the following section, it is described how to calculate the intensity of the polarized Raman scattering experiment in a backscattering geometry on crystalline materials by taking into account the polarization state of the incident light (Sect. 3.3.1) as well as the crystal symmetry class (and/or the selection rules holding for each of the excited phonon modes) (Sect. 3.3.2). Further refinements can be made by considering also the effects brought by large numerical aperture objectives [42, 43] and the surface morphology [44].

3.3.1 Model for Backscattering Raman Emission in c-Silicon

The Placzek's formalism of normal (i.e., non-resonant) Raman scattering intensity is presented in detail in classic works on optics and material science [45–47]. Here, we use such a formalism to derive a model for the far-field Raman scattering. We shall only need one of its final results, namely, the general expression for the scattered intensity:

$$I \propto \left| \sum_j e_s^T R_j e_i \right|^2 \quad (3.1)$$

in which e_i and e_s are, respectively, the incident and scattered electric field vectors (the superscript T stands for transpose) and R_j are the polarizability tensors of the j Raman-active phonon modes.

In the case of Raman scattering on a *c*-Si sample we are concerned with, (3.1) takes the form [48]

$$I_{ff} \propto \sum_k |v_s^T R'_k v_i|^2 \quad (3.2)$$

where I_{ff} is the far-field intensity, v_i and v_s are the incident and scattered electric field vectors in the laboratory reference frame (xyz) (see Fig. 3.4) and the R'_k , $k = 1, 2, 3$, are the three Raman polarizability tensors of the three (TO₁, TO₂ and LO) Si–Si phonons at 521 cm⁻¹, referred to the same frame:

$$R'_k = F_{10}^T T_y^T(\phi_1) R_k T_y(\phi_1) F_{01} \quad (3.3)$$

In (3.3), R_k are the corresponding tensors referred to the sample reference frame ($x'y'z'$), $T_y(\varphi_1)$ is a rotation matrix transforming ($x'y'z'$) into (xyz), and F_{ij} , $i, j = 0, 1$ are Fresnel matrices accounting for the electric field

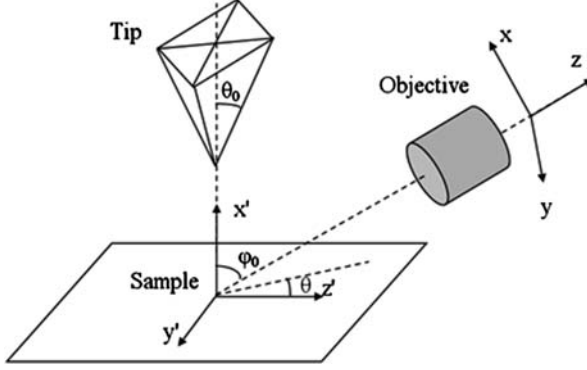


Fig. 3.4. Model geometry used in the calculations, defining the sample reference frame ($x'y'z'$) and the laboratory reference frame (xyz). The x' axis is directed along the sample normal; the z axis is along the direction of propagation of light. The two axes x and x' define the scattering plane. Reproduced with permission from [15]

amplitude changes occurring at the ambient-medium (“01” subscript) and medium-ambient (“10” subscript) interface crossings (the transposition results from the backscattering configuration). The value φ_1 is the incidence angle in the medium, defined by the Snell’s law:

$$n \sin \varphi_1 = n_0 \sin \varphi_0 \quad (3.4)$$

n being the refractive index of the medium (the value of $n = 4.22$ at 514 nm and $n = 3.88$ at 633 nm for c-Si [49] was taken in our case) and n_0 that of the ambient ($n_0 = 1$).

In the crystallographic reference plane, the normalized Raman scattering tensors of the three (TO_1 , TO_2 and LO) F_{2g} phonon modes along the $[100]$, $[010]$, and $[001]$ axes in cubic c-Si have the respective forms

$$R_1^{crystal} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, R_2^{crystal} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ and } R_3^{crystal} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (3.5)$$

In the sample reference plane ($x'y'z'$), the three tensors have identical expressions for a (001)-oriented c-Si. In the case of (111)-oriented c-Si, the above tensors must be transformed according to the relation

$$R_k = T^T R_k^{crystal} T \quad (3.6)$$

where

$$T = \begin{bmatrix} \sqrt{1/3} & -\sqrt{1/6} & -\sqrt{1/2} \\ \sqrt{1/3} & \sqrt{2/3} & 0 \\ \sqrt{1/3} & -\sqrt{1/6} & \sqrt{1/2} \end{bmatrix} \quad (3.7)$$

is the transformation matrix bringing the crystallographic reference frame into the sample one.

If the sample is rotated around its normal at an (azimuth) angle θ , the expressions for the three tensors R_k , $k = 1, 2, 3$, take the form

$$R_k^\theta = T_x^T(\theta)R_kT_x(\theta) \quad (3.8)$$

with the rotation matrix $T_x(\theta)$ being given by

$$T_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{bmatrix}. \quad (3.9)$$

The transformation matrix $T_y(\varphi)$ bringing the sample reference frame into the laboratory one (see Fig. 3.4) has the form of a rotation matrix about the y -axis:

$$T_y(\phi) = \begin{bmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{bmatrix}. \quad (3.10)$$

Finally, the Fresnel matrix F_{01} is given by

$$F_{01} = \begin{bmatrix} t_{01}^p & 0 \\ 0 & t_{01}^s \\ 0 & 0 \end{bmatrix} \quad (3.11)$$

where t_{01}^p and t_{01}^s are the ambient-to-medium (“01” subscript) transmission coefficients for p and s polarizations, respectively. (With the choice of reference frames as in Fig. 3.4, p is directed along the x -axis and s along the y -axis of the laboratory frame.) The transmission coefficients are expressed by the well-known Fresnel relations [50]:

$$t_{01}^p = \frac{2n_0 \cos \phi_0}{n_0 \cos \phi_1 + n_1 \cos \phi_0} \quad \text{and} \quad t_{01}^s = \frac{2n_0 \cos \phi_0}{n_0 \cos \phi_0 + n_1 \cos \phi_1} \quad (3.12)$$

(Neglecting the 1–2-nm-thick native oxide layer present on top of any c-Si wafer results in less than 0.07% error on the transmission coefficients at 514nm.) Note that F_{01} is a rectangular 3×2 matrix since the three-component incident and scattered electric field vectors e_i and e_s pass into the two-component v_i and v_s ones in the laboratory reference frame (the third field component being identically zero in this frame because of the transverse character of the field).

The F_{10} matrix accounting for the medium-to-ambient (“10” subscript) transmission of the scattered field is obtained by simply permuting the subscripts “0” and “1” in (3.11) and (3.12).

Expression (2) – together with (3) – is the one used in modeling the far-field response of the sample.

3.3.2 Selection Rules

In the backscattering configuration, when dealing with linearly polarized light, the general expression for the Raman emission given in (3.1) can be seen as a function of the angle χ between the incident and detection polarization field, that is, $I = f(\chi)$. The specific form of f depends on the crystallographic orientation of the sample and on the geometry of the laboratory reference frame. What is important here is that for specific crystal orientations and excitation/detection polarization arrangements the Raman scattering can be forbidden, that is, $I = 0$. Such a selection rule is, of course, dependent on the symmetry of the mode, that is, on the form of the associated Raman tensor. Modes with different symmetry have different selection rules. To give a few examples, for Si(001) in the [100], [010], and [001] reference frame, if we assume *s*-polarized incident radiation parallel to the [100] or to the [010] crystallographic direction, (3.1) gives $I \propto \sin^2(\chi)$ for the first-order Raman emission at 520 cm^{-1} . This indicates that the emission is totally cross-polarized with respect to the incident polarization or, equivalently, it is forbidden in the polarization state parallel to the incident radiation [27, 48]. An analogous result, that is, cross-polarized Raman emission, is found for *p*-polarization at normal incidence, while if working at oblique incidence the F_{2g} mode is still strongly attenuated [48] when analyzing the *p*-polarized component of the Raman emission. On the other hand, if the sample is rotated by 45° , that is, in the [1–10] [110] [001] reference frame, the Raman scattering turn out to be totally polarized parallel to the incident radiation. Therefore opposite selection rules hold. Similar considerations hold for the second-order Raman modes in the [100] [010] [001] reference frame. Such configurations are particularly important in TERS since, as we will see in the next sections, they allow for a reduction or even a complete suppression of the far-field background.

3.4 Tip-Enhanced Field Modeling

The modeling introduced in the previous chapter to describe the polarized Raman scattering in the far-field is extended here to take into account the influence of the near-field metal probes, with an emphasis on their polarization properties (Sect. 3.4.1). An overview on the estimations provided by numerical methods on the overall field enhancement operating in TERS is given in Sect. 3.4.2.

3.4.1 Phenomenological Model

The tip-enhanced field modeling is based on the introduction of a “tip-amplification tensor” A expressing the well-established – both theoretically [51, 52] and experimentally [22] – fact that the electric field component parallel to the tip axis is preferentially amplified compared with that perpendicular

to it. Assuming the tip to be normal to the sample surface, the tensor A has the following diagonal representation in the sample reference frame

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{bmatrix} \quad (3.13)$$

where a and b (with $a > b$) are phenomenological tip-amplification factors. The tip geometry (i.e., form, radius of curvature, aperture angle, etc.) and material dielectric constant determine the actual values of the amplification factors. It is interesting to note that a similar tensor has been proposed and successfully used by J. A. Creighton in modeling SERS [53, 54] and found by Bouhelier et al. to be valid also in transmission-mode scattering-SNOM [55]. When transformed to the laboratory reference frame, the tensor A takes the form:

$$A' = F_{00}^T T_y^T(\phi_0) A T_y(\phi_0) F_{00} \quad (3.14)$$

in which the T - and F -type matrices play the same role as those entering (3.3) above with now φ_0 being the incidence angle with respect to the sample normal ($\varphi_0 = 70^\circ$ in our case) and the F_{00} matrix being simply the 3×2 identity matrix:

$$F_{00} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (3.15)$$

as can be readily arrived at by substituting two identical indices in (3.11) and (3.12). The transmission coefficients along the main diagonal are both equal to one since the complete tip-radiation interaction, which results in the field amplification and in the modification of the field polarization state, has already been accounted for in the amplification tensor.

The tip-enhanced scattered field intensity is given by:

$$I_{tef} \propto \sum_k |v_s^T R_k'' v_i|^2 \quad (3.16)$$

where the R_k'' are the “effective” Raman scattering tensors in the presence of the tip:

$$R_k'' = A'^T R_k' A' \quad (3.17)$$

Note that the pre- and post-multiplication of the Raman tensors R_k' by the tip-amplification tensor A' in the above expression expresses the action of the tip on the incident and on the scattered field, respectively. The tensor R_k'' can therefore be looked upon as being the effective scattering tensor of the tip-sample system.

It should be noted that the tip-amplification factors are generally expected to be wavelength-dependent because of the resonant nature of the TERS process. Thus, in the case of strongly wavelength-dependent field-enhancement

mechanism such as a LSP-type excitation, two distinct tip-amplification tensors, A_{inc} and A_{sc} , for the incident and scattered radiation, respectively, may have to be considered as a consequence of the wavelength shift (up to several tens of nanometer) inherent to the Raman scattering. Such a distinction may prove to be relevant for a tip-coating material such as Ag exhibiting very sharp resonances [56, 57]; Au usually displays broader (and less intense) resonances [58, 59]. In all our simulations, we have set $A_{inc} = A_{sc} \equiv A$.

Equation (3.16) represents, in principle, the expression for the tip-enhanced field intensity for an infinitely thin tip perpendicular to the sample. However, real tips exhibit finite aperture apex angles and moreover, may be tilted from the sample normal. Therefore, (3.16) must be averaged over the solid aperture angle of the tip apex while taking into account the tip geometry (conical, pyramidal, etc.), as well as the tip position with respect to the sample normal. Since a closed-form result is impossible to be obtained in the general case, we have numerically averaged (3.16) by simply calculating the mean intensity for four different positions $m = 1, 2, 3, 4$ of the tip expressed by four matrix representations A_m of the tip-amplification tensor A ,

$$A_m = T_m^T(\theta_0) A T_m(\theta_0) \quad (3.18)$$

where the four ‘‘tilt’’ matrices $T_m(\theta_0)$ with $m = 1, 2, 3, 4$ are given by

$$T_{1,2}(\theta_0) = \begin{bmatrix} \cos \theta_0 & 0 & \pm \sin \theta_0 \\ 0 & 1 & 0 \\ \pm \sin \theta_0 & 0 & \cos \theta_0 \end{bmatrix}, T_{3,4}(\theta_0) = \begin{bmatrix} \cos \theta_0 & \pm \sin \theta_0 & 0 \\ \pm \sin \theta_0 & \cos \theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3.19)$$

and represent rotations (or ‘‘tilts’’) of the tip position, initially supposed normal to the sample, at an angle $\pm\theta_0$ about the y' and z' axes, respectively. In our case, the four positions were chosen to coincide with the four edges of a pyramid with an aperture angle $2\theta_0$; see also the tip model in Fig. 3.4. (It should be noted that the angle θ_0 is a model parameter that depends on the averaging procedure adopted and therefore, although being directly related to and certainly close to the real tip apex-aperture angle, cannot be expected to be numerically identical to the latter.) Of course, other positions and a different number of points may be used in the averaging procedure. After bringing the matrices A_m to the laboratory frame in total analogy with (3.14), one has

$$A'_m = F_{00}^T T_y^T(\phi_0) A_m T_y(\phi_0) F_{00}. \quad (3.20)$$

Consequently, the averaged tip-enhanced scattered intensity \bar{I}_{tef} takes the form

$$\bar{I}_{tef} \propto 1/N \sum_m \sum_k \left| v_s^T R_k^{(m)} v_i \right|^2 \quad (3.21)$$

in which N is the number of tip positions used in the averaging. The ‘‘effective’’ Raman scattering tensors $R_k^{(m)}$ are calculated for each tip position m :

$$R_k^{(m)} = A_m'^T R'_k A'_m \quad (3.22)$$

Finally, the total-field scattered intensity is given by the incoherent sum of the far-field (3.2), and the tip-enhanced-field (3.11) contributions,

$$I_{tf} = \alpha I_{ff} + \bar{I}_{tef} \quad (3.23)$$

where the empirical factor α accounts for the possible presence of a “tip-shadowing effect” in the experiment. In all our calculations, α was set equal to one.

The diagonal tensor in (3.13) provides a phenomenological model of the scattering properties of metal tips. In particular, it is successful to calculate the TERS intensity from dielectric samples in which the interaction of the tip with the sample is not expected to significantly alter the tensor components a , b . When dealing with probe samples (e.g., single molecular layers) deposited metallic surfaces we expect a further field enhancement due to the optical tip-sample interaction with the metallic surface, for example, of the resonant excitation of tip-sample gap modes. In particular, the a component, describing the scattering of the field component parallel to the tip axis (and orthogonal to the sample surface), is expected to be strongly dependent from the tip-sample distance, from the dielectric constant of the metallic surface, from the incidence angle and from the excitation energy. A simplified model of the enhanced tip-sample interaction can be derived using the dipole-image-dipole electrostatic model [13, 60, 61] where the tip is approximated by a polarizable sphere which interacts with its image dipole, yielding an effective polarizability which depends non-linearly from the tip-sample distance.

3.4.2 Numerical Models and Results

Numerical methods are based on discrete solving of the Maxwell’s equation over a meshed volume containing the metal tip apex, approximated as a half-sphere or ellipsoid, the – generally flat – substrate and an intermediate medium (air, water) gap, by appropriate setting of boundary conditions [17, 62–68]. Very often the metal is treated as having a free electron (Drude) response, thus neglecting inter-band transition in the wavelength region of interest. These methods are employed to estimate the local (below the tip) electric field enhancement and its spatial extension, that is, the ultimate optical lateral resolution. In the following, the influence of the polarization state of the exciting plane wave, as predicted in some of these simulations, is reported.

Figure 3.5 reproduces in color code the field-enhancement factor as calculated with the finite element method by A. Downes et al.[69] for an in-plane (p -polarized) and out-of-plane (s -polarized) incident light. The local enhancement achieved in the former case (up to 62) exceeds the value obtained in the latter case (up to 10). The incident field component parallel to the tip axis drives more efficiently the surface charge oscillations near the end of the elongated and cylindrically symmetric tip, generating a larger vertical dipole [70].

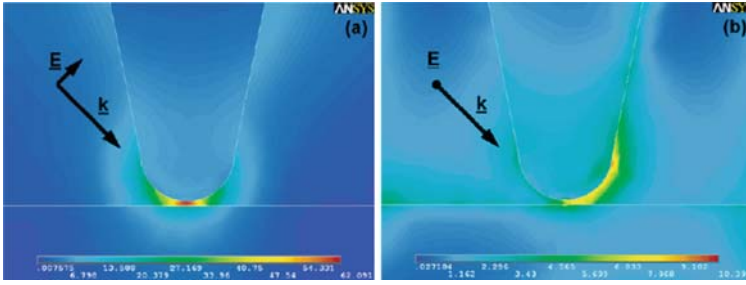


Fig. 3.5. Numerical simulation of the electric field enhancement on the substrate plane for a 20-nm-radius gold tip above a mica substrate. In (a) the incident field vector is oriented parallel to the plane of incidence (p -polarization); in (b) the incident field vector is oriented normal to the plane of incidence (s -polarization). The angle of incidence is 45° . $\lambda_{\text{exc}} = 533$ nm. Reproduced with permission from [69]

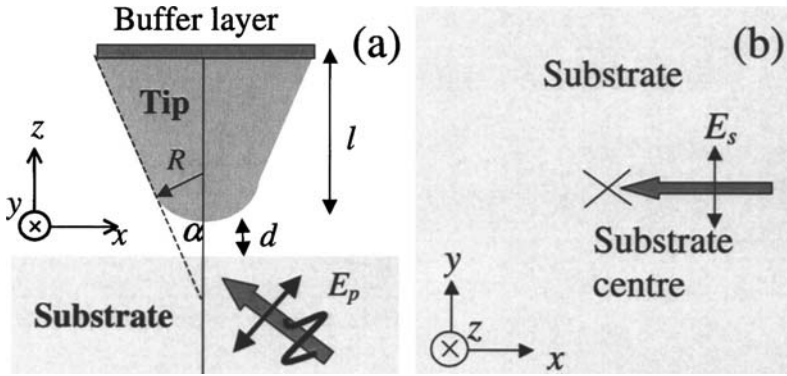


Fig. 3.6. The problem geometry and reference system employed by Demming et al. in the numerical simulation of the electric field enhancement for an illuminated silver tip above a glass substrate. $R = 20$ nm is the tip radius of curvature, $\alpha = 30^\circ$ its aperture angle, $d = 5$ nm is the tip to substrate distance, θ the angle of incidence. In (a), the plane of incidence is reported and the electric field vector is oriented parallel (p -polarization) to the plane. In (b), the substrate plane is reported and the electric field vector is oriented parallel (s -polarization) to the plane. Reproduced with permission from [51]

Moreover, in the presence of a metallic substrate, image charges are created giving rise (with p -polarized incident light) to an even larger vertical dipole than for the isolated tip [71]. This contribution is lacking with less reflecting substrates.

P. Royer et al. [72] and Demming et al. [51] studied in more detail the effects of polarization on the enhancement of the different field components parallel and perpendicular to the substrate (Fig. 3.6); this is a particularly relevant issue to correctly interpret the coupling of the (enhanced) electric field with the molecular (lattice) vibrations of the probed sample as dictated

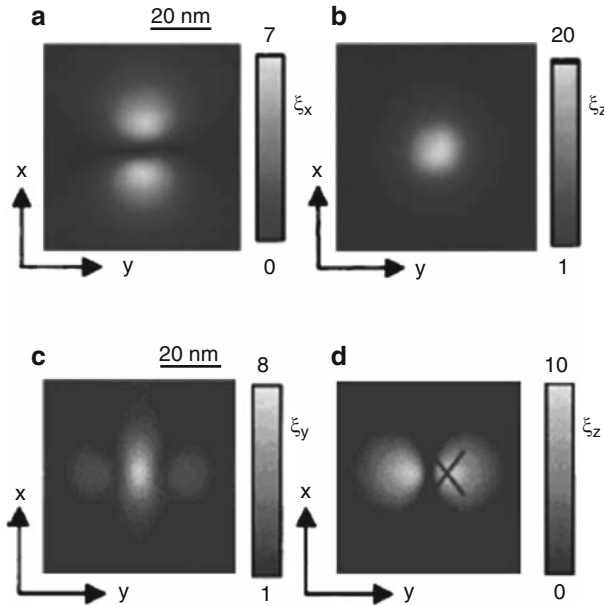


Fig. 3.7. (a) The x and (b) z components of the field-enhancement distribution on the substrate plane for p -polarized excitation; (c) y and (d) z components of the field enhancements on the substrate surface for s -polarized excitation. Reproduced with permission from [51]

selection rules. The calculations performed for a silver probe using the finite difference time domain method under p -polarized illumination show a dominating enhancement of the z field component (parallel to the tip axis) with a lower yet still appreciable enhancement of the x component also (in the substrate plane), see Fig. 3.7a, b.

With the incident polarization perpendicular to the incidence plane (s -polarization), it is again noted that the global field enhancement is minor than what is achieved with p -polarization, see Fig. 3.7c, d. Noticeably, the charge is differently distributed in the two configurations: under p -polarization, the maximal z -component enhancement is found exactly below the apex, while for s -polarization two lobes exist at either side of the tip. The authors point out at possible optical artefacts (features doubling) that may arise when imaging under such configurations. Regarding the enhanced field component in the substrate plane, this one is slightly greater under s -polarized illumination.

A relevant conclusion is that when sending linearly polarized light with the electric field vector lying in the plane of incidence (with a large component parallel to the tip axis), an out-of-plane field component (normal to the tip axis) is found in the proximity of the tip. Note that in the oblique incident configuration, there is still an initial field component normal to the tip axis also for p -polarization. Conversely and more surprising, an in-plane component

(parallel to the tip axis) is found with s -polarized excitation. Thus, in the near-field of an illuminated metal tip, or more generally, an elongated metallic nano-structure (e.g., nano-rods), a cross-polarization effect from s - to p - takes place, despite this being recognized as minor.

3.5 Depolarization of Light Scattered by Metallic Tips

Metallic near-field probes behave as scattering objects of the incident fields. In some cases, they can strongly perturb the local field, for example, by enhancing it, and acting as nanoscopic light sources. The local field will depend on the tip shape [73, 74], as well as on the illumination geometry and excitation wavelength, especially when plasmonic resonances are involved [73–77]. It is widely recognized that efficient field enhancement is obtained when the metallic tip is excited with a field polarized parallel to its axis, that is, parallel to the scattering plane (p -polarization), due to the excitation of an LSPs underneath the tip apex [51, 52, 75]. On the other hand, light scattered by non-spherical particles is in general only partially polarized [76], that is, there is a nonzero intensity component with polarization orthogonal to the incident one (the excitation polarization is assumed to be linear). In particular as we have seen in the previous section, numerical calculations [51, 52] show that when the tip is excited with both p - and s -polarized light, a cross-polarized field occurs at short distances (10 nm) from the tip apex, that is, the field is locally depolarized.

A fast way to quantify this effect, is to measure the cross-polarized component of the scattered light I_{\perp} and define a depolarization factor [77] $\Delta I_{\perp,x} = I_{\perp}/I_{//}$ by normalizing I_{\perp} to the intensity of the scattered light which is polarized parallel to the incident polarization, $I_{//}$. Here $x = s$ or p indicates the direction of the incident polarization. In order to highlight any modification to the incident polarization state induced by the scattering from an a-SNOM tip, we illuminate the tip (located far from the sample) with s - or p -polarized light and analyze the intensity $I_{s,p}(\theta)$ of the *backscattered* radiation as a function of the polarization angle θ , defined with respect to the horizontal direction (orthogonal to the scattering plane, 0° indicates s -polarization) [78], as shown in Fig. 3.8.

Depolarization measurements have been carried out on commercially available AFM gold-coated tips from Micromasch (Nsc16–Cr–Au) and Olympus (AC160TS) silicon tip coated in our laboratories with 6 nm of Cr and 30 nm of Au. In both cases, we measure apical radii of ~ 50 nm [27, 29]. At a visual inspection, strong light scattering is observed when the spot is focused on the tip apex, for both p - and s -incident polarizations (Fig. 3.9a, b, respectively), with similar intensities for both polarizations. The tip is located several microns far from the surface; therefore, this suggests that the intensity of far-field scattering is basically independent from the incident polarization

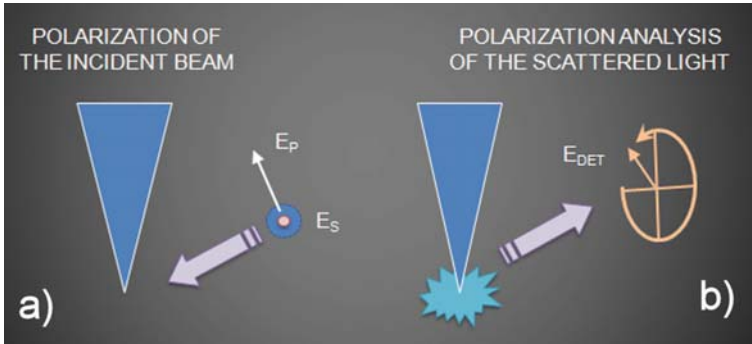


Fig. 3.8. Schematics of the experiment to measure the *s*-to-*p* and *p*-to-*s* polarization conversion in TERS tips

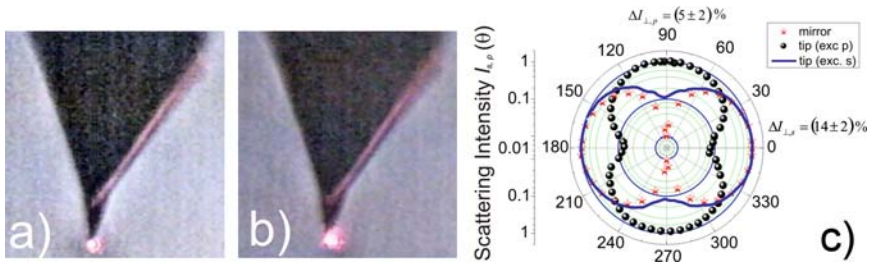


Fig. 3.9. Optical images showing the scattering from a gold-coated tip when light is focused on the apex region for *p*- (a) and *s*- (b) incident polarization. Reprinted from [27]. (c) Polar plot of the scattered intensity from a gold-coated tip as a function of the polarization angle (0° is the horizontal, *s*-direction). When the tip is illuminated with a *p*-polarized beam, we observe (black circles) a cross-polarized scattering in the order of 5% of the parallel polarized scattering. When exciting with a *s*-polarized beam (blue line), the depolarization grows up to 14%. No depolarization is observed, within our sensitivity limit of 1%, when light is focused on a flat mirror surface (red stars). Reprinted from [29]

state. In Fig. 3.9c, we show the polar plot of the backscattered intensity $I_{s,p}(\theta)$ for *s*- and *p*-incident polarization.

The intensity is normalized to $I_{//}$. Measurements have been carried out at 633 nm, using a long working distance microscope objective (Mitutoyo WD 30.5 mm, NA 0.28), inclined at 45° with respect to the vertical axis of the probe to focus the beam on the tip. A rotating Polaroid sheet is used as the analyzer. The red stars refer to a control experiment in which the tip has been substituted by a mirror, where the expected value $I_{\perp} = 0$ (mirrors do not depolarize light) is actually observed (we measure $\Delta I_{\perp,s} < 0.01$, due to the limited extinction ratio of the analyzer). Equivalent results are obtained when focusing light on the lateral cantilever edge. When light is focused on the apex of the Olympus gold coated tip, we clearly observe the appearance of

a component of light polarized orthogonally to the incident polarization. For p -polarized incident radiation (Fig. 3.9c, black circles), the average depolarization factor is $\Delta I_{\perp,p} = (5 \pm 2)\%$. For s -polarized incident light (Fig. 3.9c, blue line), the depolarization increases to $\Delta I_{\perp,s} = (14 \pm 2)\%$. Higher depolarization factors, up to 30%, have been measured on defective tips as well as on bulk metallic tips [27]. These results, therefore, indicate that light scattered from metallic tips is only partially polarized, and that a non-negligible part of the radiation impinging on the sample is expected to be polarized orthogonally to the incident state.

3.6 Polarized Tip-Enhanced Raman Spectroscopy of Silicon Crystals

TERS on silicon crystals has recently received new momentum due to the unique opportunity for background suppression in the side-illumination configuration due to the well-defined polarization properties of the Raman emission [10, 48]. In particular, the strong and highly reproducible signal enhancement observed for the 1-phonon mode of c-Si in the specific configurations has attracted much interest due to promising applications in high-resolution imaging of stress in silicon-based devices. In fact, Zhu et al. [79] demonstrated TERS imaging on strained Si films with 250-nm spatial resolution, while Sokolov's group exploited the depolarization configuration for high-resolution Raman imaging on Si nanostructures claiming 20-nm spatial resolution [25]. In this section, we show how, playing with polarization, it is possible to suppress the far-field background (Sect. 3.6.1), and to selectively enhance specific Raman modes (Sect. 3.6.2).

3.6.1 Background Suppression

Bulk crystalline silicon is known to exhibit relatively low total-field-to-far-field contrast under “unanalyzed” (i.e., with no analyzer in the scattered beam) TERS conditions. The near-field Raman signal is stemming from a nanoscale volume below the externally illuminated tip and is generally much weaker than the far-field background from the spot-illuminated volume, despite the near-field contribution undergoing the huge exaltation due to the plasmonic tip-enhancement. Indeed, in the oblique incidence configuration, suitable to study opaque samples, the illuminated (elliptical) area may reach up to $\sim 4\text{--}7\ \mu\text{m}^2$ (considering the NA of the most commonly employed long working distance objectives), while the laser light penetration depth is up to $\sim 3.2\ \mu\text{m}$ ($0.8\ \mu\text{m}$) at 633 nm (514 nm). On the other side, the tip enhancement extends laterally over an area that is in a first approximation limited by the apex size ($R_t = 30\text{--}50\ \text{nm}$) and stretching vertically up to 10–30 nm below the surface, due to its evanescent nature [80, 81]. With the contrast C being defined as the ratio

of the total-field scattered light intensity (“tip in contact”) to the far-field (“tip withdrawn”) one,

$$C = I_{tf}/I_{ff}, \quad (3.24)$$

this value is reported to reach at most between 1.15 and 1.35 for the Si phonon band at 521 cm^{-1} , with an incident polarization set parallel to the scattering plane (and to the tip axis) to achieve maximum TERS enhancement [82, 83].

The introduction of an analyzer in the scattered beam – set at 90° azimuth (axis perpendicular to the scattering plane or *s* polarization) – may already increase the contrast value up to 1.85 [28]. This reveals the potential capabilities of the “polarized” TERS configuration for contrast increase in (bulk) crystalline samples, as first suggested by Poborchii et al. [84]. By controlling the polarization states of the incident and scattered radiations even higher contrast values can be achieved by preferential suppression of the far-field signal, while still letting through a part of the near-field signal [85]. The efficiency of the polarization approach in the contrast increase is due to the fact that far-field and near-field contributions have different polarizations. Indeed, the tip modifies the states of polarization of both the incident and the scattered beams, while enhancing mainly the electric field component parallel to its axis.

The samples studied here are square [triangular for (111)-orientation] pieces of *c*-Si wafers (ITME, Poland) with two different crystallographic orientations: (001) and (111). For STM measurements employing electrochemically sharpened Au tips, highly *n*-doped ($0.05 - 0.01\ \Omega\cdot\text{cm}$) wafers were used. After ultrasonic cleaning in acetone, the oxide-free surface was prepared by performing two cycles of oxidation in a 3:1 concentrated $\text{H}_2\text{SO}_4/30\%\ \text{H}_2\text{O}_2$ and etching in 40% NH_4F . Tunneling experiments were performed in air over a period never exceeding 2 h after sample preparation. The sample holder could be manually rotated to an arbitrary azimuth angle with $\pm 5^\circ$ accuracy.

Figure 3.10a, b presents, respectively, the two scattered-field intensities at 521 cm^{-1} – total and far – as well as the resulting contrast on a (001)-oriented *c*-Si obtained with an Au-coated AFM tip illuminated with 514 nm wavelength, as functions of the incident polarization with the analyzer set at 90° . The phonon peak intensity is clearly higher with the tip in contact. The curves are calculated with the models presented in Sect. 3.3 for the far-field response and in Sect. 3.4 for the total-field response. In the numerical simulation, the incident angle and linear polarization of the incoming radiation, the polarization state of the detected backscattered radiation as defined by the analyzer setting, the azimuth angle of the crystalline sample and its orientation in the laboratory reference frame are all fixed according to the experimental configuration. With these parameters defined, the Raman-scattered radiation in the far-field as a function of the polarization of the incident light can be uniquely reproduced. The curve of the “absolute” intensity has to be adjusted to “reproduce” the actual instrumental sensitivity. Thus, a “calibration” factor is also used when simulating the total-field curve. In this second step, the

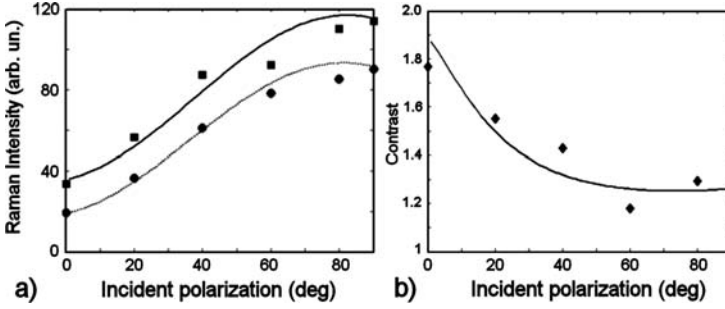


Fig. 3.10. (a) Experimental (*points*) and calculated (*lines*) peak Raman intensity of the 521 cm^{-1} phonon band of (001)-oriented c-Si against incident polarization with the tip in contact (*squares and solid line*) and the tip withdrawn (*circles and dotted line*). (b) Experimental (*circles*) and calculated (*line*) contrast ratios from (a). The analyzer was set at 90° , the sample azimuth was 45° with respect to the [001] crystallographic axis. Reproduced with permission from [15]

tip-in-contact signal is simulated. To properly weight the near-field contribution to the total signal, the smaller number of scatterers excited by the tip enhanced field as compared to the much bigger number contributing to the far-field component must be taken into account. Since the exact field distribution in the silicon sample in the near-field (but also in the far-field case) can be only very roughly estimated, a signal “attenuation” factor is introduced. This one is defined as the product of the absolute value of one of the field components amplification factor, b , by the unknown ratio between the two volumes involved in the Raman process. In computing the near-field signal contribution, the volume ratio gets cancelled if considering only the relative ratio a to b as one of the fit parameter. By adopting this procedure, we can focus on the polarizing properties of TERS basing our simulation and subsequent discussion basically only on the recorded data. Figure 3.11 shows the results obtained on a (111)-oriented c-Si sample at two different azimuth angles when employing etched Au STM probes illuminated with 633-nm wavelength. The model curves shown in Figs. 3.10 and 3.12a (AFM tip) were obtained with a ratio $a : b = 6.1 : 1$; while for Fig. 3.11 (STM tips), the ratios were 5.5:1 and 1.6:1. The attenuation factor was always in the order of 10^{-4} and the aperture angle θ_0 was varied between 60° and 70° .

So far, the incident polarization was varied while the analyzer azimuth was kept fixed. However, as mentioned above, the total- and far-field intensities and therefore, the contrast, should depend on the polarization states of both the incident and the scattered light beams. This is illustrated in Fig. 3.12a where the contrast ratios obtained at different analyzer azimuth settings are plotted against the incident polarization. Higher contrast values can be thus obtained by combining appropriate incident and scattered polarization azimuths. This is further shown by the data plotted in Fig. 3.12b collected on a (001)-oriented c-Si sample by setting the analyzer azimuth so

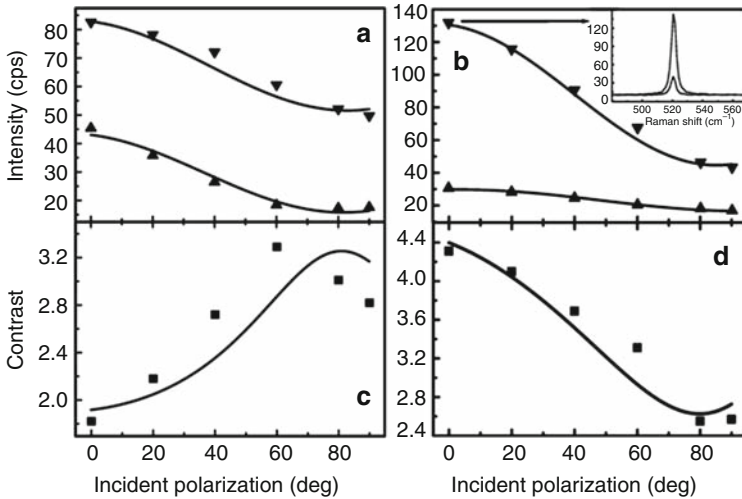


Fig. 3.11. (a), (b): Raman intensities of the 521 cm^{-1} phonon band of a bulk Si(111) sample as a function of the polarization of the incident light for two different tips, (a) and (b): tip in tunneling contact (*downward pointing triangles*), tip retracted (*upward pointing triangles*). (c), (d): contrasts corresponding to the intensities (a) and (b), respectively. Sample is rotated about its normal with respect to the incident light by 20° for (a,c) and 70° for (b,d). The analyzer was set at 90° . The solid curves are model fits. In the inset, the original Raman spectra with and without the gold tip for the *p*-polarized incident light case are reported. Reproduced with permission from [86]

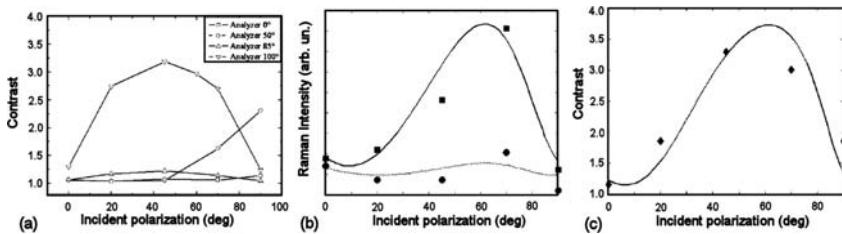


Fig. 3.12. (a) Contrast ratio of the tip-in-contact to tip-withdrawn peak Raman intensities of the 521 cm^{-1} Raman band of (001)-oriented c-Si against incident polarization for different analyzer settings. Reproduced with permission from [15]. (b) Experimental (*points*) and calculated (*lines*) peak Raman intensity of the 521 cm^{-1} phonon band of (001)-oriented c-Si against incident polarization with the tip in contact (*squares and solid line*) and the tip withdrawn (*circles and dotted line*). (c) Experimental (*circles*) and calculated (*line*) contrast ratios from (b). For each incident polarization, the analyzer was set so as to minimize the far-field scattered intensity. Reproduced with permission from [15]

as to minimize the far field for each incident polarization value and therefore, to maximize the contrast [28]. Indeed, the contrast obtained at “optimal” analyzer azimuth settings – Fig. 3.12c – generally exceeds that obtained with a “fixed” analyzer – Fig. 3.10b – except at small incident polarization azimuths. Unlike in the unanalyzed configuration, the maximum contrast is not found at p -incident polarization because at p -polarization the near-field and far-field have close polarizations and are almost equally suppressed by the analyzer. Nor does the maximum contrast occur at s -polarization since the near field is much less enhanced at this polarization [87]. Instead, the highest contrast is observed at an intermediate incident polarization. The polarization state of the incoming radiation is modified by the tip, as is also the polarization of the Raman-scattered radiation. The twice-interacting tip affects the polarized nature of the tip-in-contact Raman signal, differentiating it from the tip-retracted (far-field) signal. Hence, the ability to minimize the latter with the help of the analyzer.

Far-field suppression can be further optimized through sample rotation [28]. In the oblique incidence configuration, rotation of the crystalline sample around its normal modifies the relative excitation probability of the three degenerate phonon modes. As a result, the polarization state of the scattered radiation changes with the sample orientation and this already influences the recorded intensity in the far-field, especially if an analyzer is present. This can be appreciated by comparing the lower (far-field) curves in Fig. 3.11a, b. Moreover, considering the near-field case, the scattered radiation interacts with the tip-enhanced field more or less strongly depending on its polarization state, exactly as is the case with the incident radiation. Rotation of the sample changes the polarization state of the scattered radiation; the near-field component of the tip-in-tunneling signal is accordingly modified by sample rotation, but to a different extent than the far-field one, because these two components are already differentiated in their polarization state by the first interaction of the incoming radiation with the tip. The very different slopes observed in Fig. 3.11c, d also reflect a change in the sample orientation. In the general case for crystalline materials, higher contrast values [e.g., those observed in Fig. 3.11d for the second tip with respect to those observed in (c) for the first] do not necessarily correspond (only) to a stronger near-field excitation: the influence of the sample orientation and the far-field (preferential) suppression by the analyzer should not be overlooked. More specifically, a high contrast value may be the result of a suitable choice of sample orientation, incident polarization and analyzer setting, even for a moderately enhancing tip.

The absolute value of the near-field Raman enhancement cannot be directly derived by the phenomenological model. It rather focuses on the polarization properties of the used probes. From the a to b ratios evaluated from the fitted data reported above, it results that the amplification of the field component parallel to the tip axis is greater than that perpendicular to the tip axis, as expected. Nonetheless, the field perpendicular to the tip axis is also enhanced. As discussed in Sect. 3.4.2, numerical simulations of the

field-enhancement factors of aperture-less tips performed at p - and s -incident polarizations also predict the presence of a cross-polarization phenomenon (an s -to- p polarization conversion and vice versa) in the near field. The tip-enhancement model implicitly takes into account these cross-polarization properties by averaging the effective tensor responses over the finite aperture angle of the tip. In particular, it can be seen from (3.21) and (3.22) that the averaging procedure generates cross-polarization while the response of the effective tensor alone does not.¹ However, one must be particularly cautious in a possible attempt to physically interpret the tip-amplification factors since the model, being purely phenomenological, provides no direct information on the nature of the mechanisms involved (LSP, electric field gradient, tip-sample interaction, etc.). This is in total contrast with Creighton's model of a SERS-exhibiting particle in which the enhancement tensor is readily expressible in terms of the particle dielectric constant within the frame of the Lorentz local field theory.

3.6.2 Selective Enhancement of the Raman Modes Induced by Depolarization

The physical origin of the signal increase in the maximum-contrast configuration has been discussed for long. The observed signal increase is usually a priori associated to an enhancement of the electromagnetic field, neglecting other effects (e.g., induced by light depolarization) that may also contribute to the signal contrast and which are not discriminated. Also, a portion of the incoming beam may be deflected by the tip shaft and generate an additional stray field contribution [26,88] that may more than compensate the losses due to the shadowing of the original focus. Understanding the physical origin of these contributions is peculiar since it gives insight on the spatial extension of the effective volume probed by global field and, ultimately on the spatial resolution that can be attained. Electromagnetic field enhancement, in fact, is strongly localized on the 10-nm scale, while it is not yet clear whether this applies also to depolarization or not. Stray fields extend on micron scale areas.

Our groups have recently shed light on the origin of the signal increase showing how this is due to the interplay of both electromagnetic field enhancement and depolarization induced by the tip [27, 29, 89, 90]. Better insight on this phenomenon has been obtained by a careful analysis of the enhancement factors on both the 1-phonon mode at 520 cm^{-1} and the 2-phonon modes at 300 and 980 cm^{-1} in Si(001) crystals exciting the tip-sample assembly with s -polarized radiation, for which minimum field-enhanced is expected. Such modes feature different symmetries and therefore can be used to discriminate between pure field enhancement (i.e., increase of the local excitation power) and depolarization (i.e., increase of one specific component of the local electric field).

¹ To explicitly account for cross-polarization effects, non-zero off-diagonal elements should be allowed in the representation of the tip amplification tensor A .

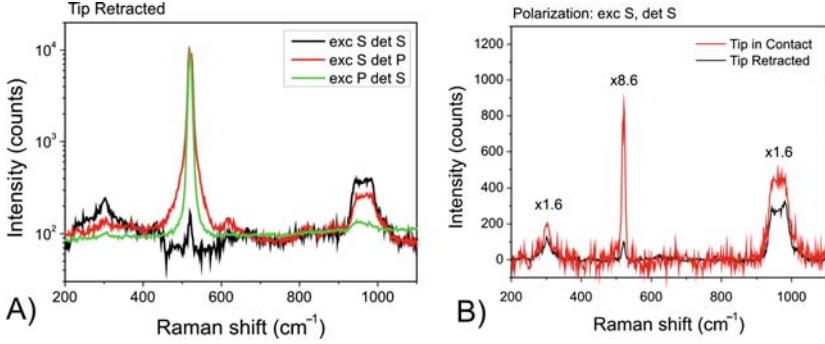


Fig. 3.13. (a) Far-Field (tip retracted) Raman emission of c-Si(100) as a function of the detection polarization state, for *s*-polarized excitation parallel to the [100] crystallographic direction. The 1-phonon mode at 520 cm^{-1} is strongly *p*-polarized (red line). The 2-phonon mode at 300 cm^{-1} is strongly *s*-polarized (black line), while the one at 980 cm^{-1} is more intense for *s*-polarization. Vice versa, for *p*-polarized excitation (green line), the Raman emission at 520 cm^{-1} is strongly *s*-polarized, while the 2-phonon modes are almost forbidden in the *s*-polarization state. (b) Background-subtracted Raman spectra of c-Si(100) for *s*-polarized excitation and *s*-polarized detection with tip retracted (black line) and tip in contact (red line). In the latter case, we observe a similar enhancement ($\times 1.6$) of the 2-phonon Raman modes and a much stronger enhancement ($\times 8.6$) of the 1-phonon mode. The Raman signal enhancement, observed when the tip is in contact with the surface, is always accompanied by a strong continuous background (red line in the inset) due to inelastic scattering from the tip metal coating. Laser wavelength 514.5 nm , power 3 mW , integration time 60 s . Reproduced from [90]

Figure 3.13a shows the far-field Raman spectra of a c-Si(100) sample excited at 514.5 nm with *s*-polarized (black and red lines) and *p*-polarized (green line) laser radiation. The tip is retracted. For *s*-polarized excitation, the field is parallel to the [100] crystallographic direction. As expected, in this configuration, the 1-phonon mode at 520 cm^{-1} is totally *p*-polarized (red line) with a residual *s*-polarized signal of the order of 1.1% (black line), probably due to the limited extinction power of the Polaroid sheet ($\sim 1\%$). We measure, in fact, $I_{SP}(520)/I_{SS}(520) \approx 90$ where $I_{XY}(\Delta E)$ is the signal measured with *X*-polarized excitation, analyzing the *Y*-polarization component of the Raman emission, for a Raman shift of $\Delta E\text{ cm}^{-1}$. Similarly, when the incident polarization is switched to *p* (i.e., parallel to the scattering plane), the Raman emission at 520 cm^{-1} gets strongly *s*-polarized (Fig. 3.13a, green line). This is in agreement with the polarization- and incidence angle-dependent selection rules discussed in Sect. 3.3.2. The behaviour of the 2-phonons modes at 300 cm^{-1} and 980 cm^{-1} is opposite. For *s*-polarized excitation (Fig. 3.13a, black and red lines), maximum Raman emission is observed in the *s*-polarization state [$I_{SS}(980)/I_{SP}(980) \approx 2$ and $I_{SS}(300)/I_{SP}(300) \approx 5$]. Conversely, for *p*-polarized excitation, the 2-phonons emission in the *s*-polarization state is strongly attenuated [$I_{PS}(980)/I_{SS}(980) \approx 1/10$].

The spectra in Fig. 3.13a represent the far-field background in TERS experiments and show that the best configuration to suppress the 520 cm^{-1} emission in c-Si(100) is the *SS* one (*s*-polarized excitation parallel to the [100] or [010] crystallographic directions, followed by detection of the *s*-component of the Raman scattering). The ratio between the 1-phonon and the 2-phonons emission intensity $\eta_{XY} = I_{XY}(520)/I_{XY}(980)$ is a good parameter to assess the background suppression achieved. In particular, it is independent from the local excitation power, that is, from the global field enhancement. η is typically of the order of 10^2 for those configurations in which the 520 cm^{-1} mode is allowed, and goes down by more than two orders of magnitude in the *SS* configuration in which the mode is forbidden ($\eta_{SS} \approx 0.5$). *We remark that the SS configuration does not suppress the overall far-field Raman scattering. It selectively minimizes the background at 520 cm^{-1} . The emission at 300 cm^{-1} and 980 cm^{-1} is, on the contrary, maximized.* A further important conclusion that can be drawn here is that, by setting the detection analyzer to *s*, the signal measured at 520 cm^{-1} becomes extremely sensitive to any *p*-polarized component present in the excitation field (compare black and green lines in Fig. 3.13a). On the contrary, the signals at 300 cm^{-1} and 980 cm^{-1} are almost insensitive *p*-polarized excitation [Fig. 3.13a (green line)]. We can therefore conclude that any *p*-polarized component in the excitation field will selectively enhance the 520 cm^{-1} emission with respect to the one at 980 cm^{-1} .

Figure 3.13b (red line) shows the TERS spectrum of c-Si(100) in the *SS* configuration ($\lambda = 514.5\text{ nm}$, $P = 3\text{ mW}$, integration time 60 s). The tip (silicon nitride gold coated) is in contact with the surface, the gold fluorescence background has been subtracted for clarity. We observe a considerable enhancement ($\times 8.6$) of the 1-phonon mode at 520 cm^{-1} with respect to the far-field spectrum [Fig. 3.13b (black line), tip retracted by several microns], whereas the 2-phonons modes are subject to a much smaller signal increase ($\times 1.6$). To quantify these effects, we use the contrast function defined by (3.24). In this case, the contrast function depends on the excitation/detection polarization states *XY* and on the energy of the Raman mode considered ΔE , $C = C_{XY}(\Delta E)$. For the *SS* configuration reported in Fig. 3.13b, we get a contrast $C_{SS}(520) \approx 7.6$, much stronger than $C_{SS}(980) \approx 0.6$. The $\eta_{SS} = I_{SS}(520)/I_{SS}(980)$ parameter increases from ~ 0.5 in the far-field Raman spectrum [Fig. 3.13b (black line)] to ~ 2 in the TERS spectrum [Fig. 3.13b (red line)]. Such a difference is not compatible with a simple interpretation in terms of local field enhancement. Field enhancement is expected to amplify all the vibrational peaks by similar factors, and therefore leave the η_{SS} parameter almost unchanged. On the contrary, here we observe a selective enhancement of the 520 cm^{-1} peak. As seen above, this can be interpreted in terms of depolarization of the incident *s*-polarized field, yielding a local field partially *p*-polarized, which in turn produces a selective enhancement of the T_{2g} mode signal at 520 cm^{-1} . The enhancement of the 2-phonons modes, on the contrary, cannot be assigned to depolarization of the incident *s*-polarized field since any *p*-polarized component would give negligible contributions to

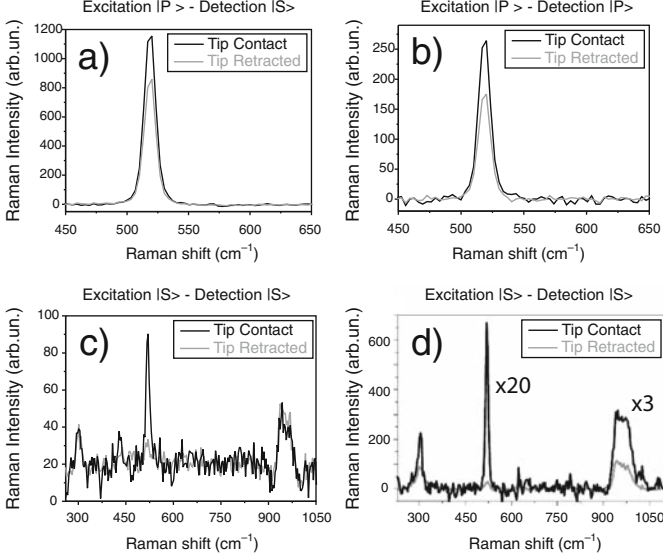


Fig. 3.14. Raman spectra acquired in the PS (a) and PP (b), SS (c and d), excitation/detection configurations, with the tip in contact with the surface (*black lines*) or withdrawn by a few microns from the surface (*grey lines*). Redrawn from [27, 29]

the *s*-polarized emission at 300 and 980 cm^{-1} [see Fig. 3.13a (green line)]. We therefore interpret such a signal increase as a tip-induced signal enhancement.

Let us focus now on TERS experiments done on c-Si(001) with *p*-polarized excitation using 632.8-nm excitation (Fig. 3.14, for more experimental details see [27]). A bulk gold tip has been used for the measurements in shear-force mode.

For *p*-polarized excitation and *s*-polarized detection, we observe an increase of 37.5% of the 520 cm^{-1} mode intensity (Fig. 3.14a). This increase can only be due to tip-enhancement since the depolarization has no effect in this configuration. In fact, any *s*-polarized field due to depolarization would give no contribution to *s*-polarized Raman scattering. Let us now consider the *p*-polarized scattering at 520 cm^{-1} (Fig. 3.14b). We observe an increase of 57% of the signal when the tip is in contact. In this case, in order to account for the global increase of the Raman spectrum, we need to consider both the field enhancement and the depolarization effect. Indeed, in the PP configuration, any *s*-polarized field due to depolarization is expected to contribute to the *p*-polarized scattering (I_{SP} for the 520 cm^{-1} mode is different to 0 and is largely higher than the I_{PP} , $I_{SP} = 4 \times I_{PP}$). Thus, the observed mode intensity I_{TERS} includes two contributions:

$$I_{TERS} = I_{PP}(\text{enhanced signal}) + I_{SP}(\text{depolarized signal}) \quad (3.25)$$

To determine the actual influence of the depolarization effect on the signal enhancement, we propose to estimate the depolarization factor from all the observed values. If we consider that for the same excitation energy $I_{SP} = 4 \times I_{PP}$, and that the enhancement factor is 1.375 (+37.5%) and that intensity observed with tip is 57% higher than the one observed without tip in PP configuration, we can write the following equation:

$$I_{TERS} = 1.57 \times I_{bkg} = 1.57 \times I_{PP} \quad (3.26)$$

with I_{TERS} the intensity with the tip and I_{bkg} , the Raman intensity without the tip.

$$\begin{aligned} I_{TERS} &= 1.375 I_{PP} \cdot (1 - \Delta) + I_{SP} \cdot \Delta = 1.375 I_{PP} (1 - \Delta) \\ &+ 4 I_{PP} \cdot \Delta = I_{PP} (1.375 + 2.65 \Delta) \end{aligned} \quad (3.27)$$

with Δ , the depolarization factor.

From (3.27) the depolarization factor can be estimated to be $\Delta = 7.4\%$. This factor is very close to the one observed experimentally with P excitation state ($\Delta = 6 \pm 1.5\%$) [27]

3.6.3 Evaluation of the Field Enhancement Factor

The enhancement factor (EF) can be estimated using the following formula [13]:

$$EF = (I_{TERS}/I_{FF} - 1) \cdot V_{FF}/V_{NF} \quad (3.28)$$

with V_{NF} the probed volume in Near Field and V_{FF} the probed volumes in Far Field.

In presence of the tip, the field confinement is supposed to be on a scale defined by the tip radius of curvature (~ 20 nm), that is, $V_{NF} \sim 10^{-5} \mu\text{m}^3$. In the case of the far field, the laser spot is close to $1 \mu\text{m}$ in diameter and therefore $V_{FF} \sim 1 \mu\text{m}^3$. Thus $V_{FF}/V_{NF} \sim 10^5$. If we calculate the enhancement factor in the PS configuration, we obtain $EF = 2.6 \times 10^6$ whereas in the PP configuration we get $EF = 4 \times 10^6$. Thus, we measure two different enhancement factors for the PP and PS polarization configurations whereas the enhancement could not be changed since the excitation has not been modified. The signal variation can be simply explained by a small depolarization ratio at the tip extremity, indicating how important it is to take into account depolarization in the enhancement factor calculation.

In the SS configuration (Fig. 3.14c), in particular, large mistakes can be made by attributing the total signal enhancement to electromagnetic field enhancement. Indeed, the signal of the 520 cm^{-1} mode is very low in this configuration without tip whereas its intensity increases about ten times when the tip is approached to the sample. This effect is only due to depolarization and although we observe a signal increase at 520 cm^{-1} , there is actually no

enhancement since the 2-phonon mode is unaltered. A different case occurs in Fig. 3.14d. Here, similarly to what observed in Fig. 3.13b, we have a different signal amplification for each mode. The Raman signal is 3 times higher for the overtones and 20 times higher for the 520 cm^{-1} mode. One more time, we get two different enhancement factors: 1.6×10^7 for the overtones and 11.4×10^7 for the 520 cm^{-1} mode. The second enhancement factor is nearly one order of magnitude higher than the first one, which has no physical meaning unless we consider the contribution of depolarization to the enhancement of the 520 cm^{-1} signal.

With all these examples we have shown that the global signal increase observed in TERS could not be only due to field enhancement but can have other origin such as light depolarization at the tip extremity. It is then of real importance to take into account all these effects to correctly estimate the enhancement factors.

3.7 Conclusions

When performing Tip Enhanced Raman Spectroscopy on crystals, the polarization of the fields plays an important role. Raman scattering on crystals is strongly polarized, depends on the mode symmetry and on the experimental geometry in a manner that can be exactly predicted by calculations. This is of invaluable help to find specific configurations capable to null the far-field background which is typically orders of magnitude more intense than the near-field scattering due to the tip. Crystals, in turn, can be viewed as probes of the actual field scattered by the metallic tip and help understanding the intriguing optical properties of these fascinating nanostructures. In this chapter, after reviewing the principal setups for TERS with side-illumination, we have analyzed the theoretical model for Raman scattering in crystalline silicon and introduced a phenomenological model describing the tip scattering. We have then shown how, based on such models, we can find specific configurations maximizing the near-field contrast and highlighted that depolarization induced by sharp metallic tips plays a crucial role in the selective mode enhancement observed experimentally. New studies are ongoing in order to better understand the scattering properties from metallic tips, and in particular to assess the degree of polarization of the scattered radiation as a function of the tip geometry, excitation wavelength and illumination angle.

Acknowledgements

D. Barchiesi and A. Hartschuh are acknowledged for careful reading of the manuscript. PGG acknowledges the Université de Technologie de Troyes and the Université du Sud Toulon-Var for financial support.

References

1. D.A. Long (ed.) *The Raman effect* (Wiley, Chichester, 2002)
2. E. Blackie, E.C. Le Ru, M. Meyer, M. Timmer, B. Burkett, P. Northcote, P.G. Etchegoin, *Phys Chem Chem Phys* **10**, 4147 (2008)
3. E.C. Le Ru, E. Blackie, P.G. Meyer M Etchegoin, *J Phys Chem C* **111**, 13794 (2007)
4. C. Stöckle, Y.S. Suh, V. Deckert, R. Zenobi, *Chem Phys Lett* **318**, 131 (2000)
5. N. Hayazawa, Y. Inouye, Z. Sekkat, S. Kawata, *Opt Commun* **183**, 2000 (2000)
6. A. Hartschuh, E.J. Sanchez, S. Xie, L. Novotny, *Phys Rev Lett* **90**, 095503 (2003)
7. B. Pettinger, B. Ren, G. Picardi, R. Schuster, G. Ertl, *Phys Rev Lett* **92**, 096101 (2004)
8. E. Bailo, V. Deckert, *Chem Soc Rev* **37**, 921 (2008)
9. A. Hartschuh, *Angew Chem Intl Ed* **47**, 8178 (2008)
10. R. Loudon, *Adv Phys* **50**, 813 (2001)
11. E.J. Sanchez, L. Novotny, X.S. Xie, *Phys Rev Lett* **82**, 4014 (1999)
12. S. Kawata, V.M. Shalaev (eds.) *Tip Enhancement* (Elsevier, Amsterdam, 2007)
13. S. Patanè, P.G. Gucciardi, M. Labardi, M. Allegrini, *Riv Nuovo Cimento* **27**, 1 (2004)
14. D. Richards, R.G. Milner, F. Huang, F. Festy, *J. Raman Spectrosc* **34**, 663 (2003)
15. R. Ossikowski, Q. Nguyen, G. Picardi, *Phys Rev B* **75**, 045412 (2007)
16. O.J.F. Martin, M. Paulus, *J Microsc* **205**, 147 (2002)
17. I. Notinger, A. Elfick, *J Phys Chem B* **109**, 15699 (2005)
18. L. Novotny, S. Stranick, *Annu Rev Phys Chem* **57**, 303 (2006)
19. M. Moskovits, *J Raman Spectrosc* **36**, 485 (2005)
20. H. Qian, C. Georgi, N. Anderson, A.A. Green, M.C. Hersam, L. Novotny, A. Hartschuh, *Nano Lett* **8**, 1363 (2008)
21. T. Ichimura, N. Hayazawa, M. Hashimoto, Y. Inouye, S. Kawata, *Phys Rev Lett* **92**, 220801 (2004)
22. L. Aigouy, A. Lahrech, S. Gresillon, H. Cory, A.C. Boccarda, J.C. Rivoal, *Opt Lett* **24**, 187 (1999)
23. S.S. Kharintsev, G.G. Hoffmann, P.S. Dorozhkin, G. de With, J. Loos, *Nanotechnol* **18**, 315502 (2007)
24. X. Wang, Z. Liu, M.D. Zhuang, H.M. Zhang, X. Wang, Z.X. Xie, D.Y. Wu, B. Ren, Z.Q. Tian, *Appl Phys Lett* **91**, 101105 (2007)
25. N. Hayazawa, M. Motohashi, Y. Saito, H. Ighitobi, A. Ono, T. Ichimura, P. Verma, S. Kawata, *J Raman Spectrosc* **38**, 684 (2007)
26. N. Lee, R.D. Hartschuh, D. Mehtani, A. Kisliuk, J.F. Maguire, M. Green, M.D. Foster, A.P. Sokolov, *J Raman Spectrosc* **38**, 789 (2007)
27. P.G. Gucciardi, M. Lopes, R. Deturche, C. Julien, D. Barchiesi, M. Lamy de La Chapelle, *Nanotechnol* **19**, 215702 (2008)
28. Q. Nguyen, J. Schreiber, R. Ossikovski, *Opt Commun* **274**, 231 (2007)
29. P. G. Gucciardi, F. Bonaccorso, M. Lopes, L. Billot, M. Lamy de La Chapelle, *Thin Sol Films* 516, 8064 (2008)
30. R. Ossikovski, Q. Nguyen, G. Picardi, J. Schreiber, *J Appl Phys* **103**, 093525 (2008)
31. E. Betzig, P.L. Finn, J.S. Weiner, *Appl Phys Lett* **60**, 2484 (1992)
32. K. Karrai, R.D. Grober, *Appl Phys Lett* **66**, 1842 (1995)

33. M.C. Baykul, *Mat Sci Eng B – Solid* **74**, 229 (2000)
34. B. Ren, G. Picardi and B. Pettinger, *Rev Sci Instrum* **75**, 841 (2004)
35. X. Wang, Z. Liu, M. Zhuang, H. Zhang, X. Wang, Z. Xie, D. Wu, B. Ren, Z. Tian, *Appl Phys Lett* **91**, 101105 (2007)
36. L. Lioubille, Y. Houbion, J.M. Gilles, *J Vac Sci Technol B* **13**, 1325 (1995)
37. L. Billot, L. Berguiga, M.L. de la Chapelle, Y. Gilbert, R. Bachelot, *Eur Phys J Appl Phys* **31**, 139 (2005)
38. K. Dickmann, F. Demming, J. Jersch, *Rev Sci Instrum* **67**, 3 (1996)
39. Blow up, Images from the nanoworld, <http://www.s3.infm.it/blowup/catalogo.html>
40. F. Bonaccorso, G. Calogero, G. Di Marco, O.M. Maragò, P.G. Gucciardi, U. Giorgianni, K. Channon, G. Sabatino, *Rev Sci Instrum* **78**, 103702 (2007)
41. C. Williams, D. Roy, *J Vac Sci Technol B* **26**, 1761 (2008)
42. E. Bonera, M. Fanciulli, D.N. Batchelder, *J Appl Phys* **94**, 2729 (2003)
43. R. Ossikovski, Q. Nguyen, G. Picardi, J. Schreiber, P. Morin, *J Raman Spectrosc* **39**, 661 (2008)
44. K. Mizoguchi, S. Nakashima, *J Appl Phys* **65**, 2584 (1989)
45. R. Loudon, *The Quantum Theory of Light* (Clarendon, Oxford, 1984)
46. Cardona M, in *Light Scattering in Solids II*, ed. by M. Cardona, G. Guntherodt (Springer, Berlin, 1981)
47. G. Turrell, in *Practical Raman Spectroscopy*, ed. by D.J. Gardiner, P.R. Graves (Springer, Berlin, 1989)
48. E. Anastassakis, Y.S. Raptis, *J Appl Phys* **57**, 920 (1984)
49. D.F. Edwards, in *Handbook of Optical Constants of Solids*, ed. by E.D. Palik (Academic, Orlando, 1985)
50. M. Born, E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, 2005)
51. A.L. Demming, F. Festy, D. Richards, *J Chem Phys* **122**, 184716 (2005)
52. A.L. Demming, F. Festy, F. Huang, D. Richards, *J Kor Phys Soc* **47**, S1 (2005)
53. J.A. Creighton, *Surf Sci* **124**, 209 (1983)
54. J.A. Creighton, *Surf Sci* **158**, 211 (1985)
55. A. Bouhelier, M.R. Beversluis, L. Novotny, *Appl Phys Lett* **82**, 4596 (2001)
56. F. Demming, J. Jersch, K. Dickmann, P.I. Geshev, *Appl. Phys. B* **66**, 593 (1998)
57. P.I. Geshev, S. Klein, T. Witting, K. Dickmann, M. Hietschold, *Phys. Rev. B* **70**, 075402 (2004)
58. C. Soennichsen, S. Geier, N.E. Hecker, G. von Plessen, J. Feldmann, H. Ditlbacher, B. Lamprecht, J.R. Krenn, F.R. Aussenegg, V.Z. Chan, J.P. Spatz, M. Moeller, *Appl. Phys. Lett.* **77**, 2949 (2000)
59. C.C. Neacsu, G.A. Steudle, M.B. Raschke, *Appl Phys B* **80**, 295 (2005)
60. B. Knoll, F. Keilmann, *Opt Commun* **182**, 321 (2000)
61. B. Pettinger, K.F. Domke, G. Picardi, D. Zhang, G. Picardi, R. Schuster, *Surf Sci* **603**, 1335 (2009)
62. F. Festy, A. Demming, D. Richards, *Ultramicrosc* **100**, 437 (2004)
63. P.I. Geshev, S. Klein, T. Witting, K. Dickmann, M. Hietschold, *Phys Rev B* **70**, 075402 (2004)
64. F. Demming, J. Jersch, K. Dickmann, P.I. Geshev, *Appl Phys A* **66**, 593 (1998)
65. M. Micic, N. Klymshyn, Y.D. Suh, H.P. Lu, *J Phys Chem B* **107**, 1574 (2003)
66. A.V. Goncharenko, H.C. Chang, J.K. Wang, *Ultramicrosc* **107**, 151 (2007)
67. I.I.J.T. Krug, E.J. Sanchez, X.S. Xie, *J Chem Phys* **116**, 10895 (2002)

68. K.B. Crozier, A. Sundaramurthy, G.S. Kino, C.F. Quate, *J Appl Phys* **94**, 4632 (2003)
69. A. Downes, D. Salter, A. Elfick, *J Chem Phys B* **110**, 6692 (2006)
70. L. Novotny, R.X. Bian, X.S. Xie, *Phys Rev Lett* **79**, 645 (1997)
71. A. Cvitkovic, N. Ocelic, J. Aizpurua, R. Guckenberger, R. Hillenbrand, *Phys Rev Lett* **97**, 060801 (2006)
72. P. Royer, D. Barchiesi, G. Lerondel, R. Bachelot, *Phil Trans R Soc Lond A* **362**, 821 (2004)
73. H.C. van de Hulst, *Light scattering by small particles* (Wiley, New York, 1957)
74. F. Borghese, P. Denti, R. Saija, *Scattering from Model Nonspherical Particles* (Springer, Berlin, 2003)
75. A. Bouhelier, J. Renger, M.R. Beversluis, L. Novotny, *J Microsc* **210**, 220 (2003)
76. C.F. Bohren, D.R. Huffman, *Absorption and scattering of light by small particles* (Wiley, New York, 1983)
77. N.G. Khlebtsov, A.G. Mel'nikov, V.A. Bogatyrev, A.V. Alekseeva, B.N. Khlebtsov, *Opt Spectrosc* **100**, 448 (2006)
78. P.G. Gucciardi, R. Micheletto, Y. Kawakami, M. Allegrini, in *Applied Scanning Probe Methods II*, ed. by B. Bhushan, H. Fuchs, (Springer, Berlin, 2006), p. 321
79. L. Zhu, C. Georgi, M. Hecker, J. Rinderknecht, A. Mai, Y. Ritz, E. Zschech, *J Appl Phys* **101**, 104305 (2007)
80. M. Sackrow, C. Stanciu, M.A. Lieb, A.J. Meixner, *ChemPhysChem* **9**, 316 (2008)
81. B. Pettinger, K.F. Domke, D. Zhang, R. Schuster, G. Ertl, *Phys Rev B* **76**, 113409 (2007)
82. W.X. Sun, Z.X. Shen, *Ultramicrosc* **94**, 237 (2003)
83. R.E. Geer, N. Meyendorf, G.Y. Baaklini, B. Michel, *Proc. SPIE* **5766**, 134 (2005)
84. V. Poborchii, T. Tada, T. Kanayama, *J Journ Appl Phys* **44**, L202 (2005)
85. D. Mehtani, N. Lee, R.D. Hartschuh, A. Kisliuk, M.D. Foster, A.P. Sokolov, J.F. Maguire, *J Raman Spectrosc* **36**, 1068 (2005)
86. G. Picardi, Q. Nguyen, R. Ossikovski, J. Schreiber, *Appl Spectrosc* **61**, 1301 (2007)
87. N. Nayazawa, A. Tarun, Y. Inouye, S. Kawata, *J Appl Phys* **92**, 6983 (2002)
88. C. Georgi, M. Hecker, E. Zschech, *Appl Phys Lett* **90**, 171102 (2007)
89. M. Motohashi, N. Hayazawa, A. Tarun, S. Kawata, *J Appl Phys* **103**, 034309 (2008)
90. A. Merlen, J.C. Valmalette, P.G. Gucciardi, M. Lamy de La Chapelle, A. Frigout, R. Ossikovski, *J Raman Spectrosc* **40**, 1361 (2009)

Electrostatic Force Microscopy and Kelvin Force Microscopy as a Probe of the Electrostatic and Electronic Properties of Carbon Nanotubes

Thierry Mélin, Mariusz Zdrojek, and David Brunel

Summary. This chapter addresses recent experimental studies on carbon nanotubes and nanotube devices using electrical techniques derived from atomic force microscopy. Electrostatic force microscopy (EFM), Kelvin force microscope (KFM), and their variants are introduced. We show how EFM-related techniques are used to image the electrostatic and electronic properties of individual carbon nanotubes on insulators, to manipulate their charge state, and to measure field-emission and band-structure properties of individual nanotubes. We then describe how KFM-related techniques can bring insight into the operation of electronic devices based on carbon nanotubes. We focus here on the case of field effect transistors, and describe how KFM techniques can be used to study charge transfers at the nanotube–contact interfaces, to assess the transport properties in carbon nanotubes, and, finally, to characterize carbon nanotube devices under operation.

Key words: Scanning probe microscopy, electrostatic force microscopy, Kelvin force microscopy, charge detection, carbon nanotubes, carbon nanotube field effect transistor, electrostatics, surface potential measurement.

4.1 Introduction

This chapter consists of a review of recent experimental studies on carbon nanotubes and nanotube devices using electrical techniques derived from atomic force microscopy (AFM). It falls in the conjunction of two recently developed research fields: (1) on the one hand, the field of carbon nanotubes since their discovery by Iijima et al. [1] and the strong interest in the field of nanoelectronics following the first carbon nanotube field effect transistor (CNTFET) [2, 3], and (2) on the other hand, the invention of the scanning tunneling microscope [4] followed by the AFM [5], allowing electrostatic or electronic measurements on insulators at the nanometer scale, either by direct measurement of electrostatic forces [6–9] or by the measurement of local surface potentials [10]. The aim of this chapter is to review the experimental studies

carried out in the intersection of these fruitful domains. This chapter starts with a general introduction about electrostatic force microscopy (EFM) and Kelvin force microscope (KFM) experimental techniques as well as their variants (Sect. 4.2). We describe how these techniques have been used to image the electrostatic properties of carbon nanotubes on insulators (Sect. 4.3), and then how more sophisticated electrostatic experiments have been developed at the scale of single nanotubes, in order to manipulate the charge state of individual nanotubes on insulators (Sect.4.4), or to measure their local electronic and band-structure properties (including semiconducting/metallic characters) (Sect. 4.5). In Sect. 4.6, we address the case of CNTFETs, and especially the study of charge transfers at the nanotube–contact interfaces, the assessment of diffusive or ballistic character of transport along single-walled carbon nanotubes (SWCNTs), and finally, more recent studies performed on carbon nanotube devices under operation.

4.2 Electrostatic Measurements at the Nanometer Scale

This section consists of a review of the main electrical techniques derived from AFM [5], which have been developed to perform electrostatic measurements at the nanometer scale [6–10]. Although all techniques presented here are based on the measurement of electrostatic forces, they strongly differ in their implementations, output signals, advantages, and drawbacks. No standard terminology for these techniques has been clearly used, so that, for instance, in the case of surface potential measurements, nearly identical techniques can be named either Kelvin force microscopy (KFM), surface potential probe microscopy, ac-EFM, or scanning Kelvin force microscopy. The aim of this section is to establish a comprehensive description of these techniques that can be used as a guide to understand the literature and recent developments related to the measurements of the local electrostatic properties of carbon nanotubes and devices based on carbon nanotubes.

4.2.1 Electrostatic Force Microscopy

Principle

EFM is a technique based on the measurement of electrostatic forces gradients. Basically, the force $F(z)$ felt by the cantilever oscillating at a mean distance z_0 from the substrate can be written in a first-order approximation as: $F(z) = F(z_0) + \partial F / \partial z(z_0) \cdot (z - z_0)$. In this expression, $\partial F / \partial z(z_0)$ is the derivative with respect to z of the vertical force on the cantilever tip, also referred to as the “force gradient” acting on the tip (expressed in N m^{-1}). It is equivalent to a modification $-\Delta k$ of the cantilever spring constant, which shifts the resonance eigenfrequency f_0 . Using $f_0 = 1/2\pi(k/m)^{1/2}$ (where k is the cantilever spring and m the effective mass of the cantilever), the basic relation between the

cantilever frequency shift Δf and the force gradient $\partial F/\partial z(z_0)$ can be readily obtained:

$$\frac{\Delta f}{f_0} = -\frac{1}{2} \frac{\partial F/\partial z(z_0)}{k} \quad (4.1)$$

Although the minimum detectable force gradient depends on several parameters such as the measurement bandwidth, cantilever temperature, and cantilever quality factor Q_{cant} , and, in some situations, may be limited by the AFM itself (e.g., by the photodiode detector), a practical value for the detectable frequency shifts in typical scanning conditions at atmospheric pressure is of the order of a fraction of 1 Hz, corresponding to $\Delta f/f_0$ as low as a few 10^{-6} , and thus to force gradients less than 10^{-5} N m^{-1} for cantilevers with spring constant of a few N m^{-1} , and resonance eigenfrequency f_0 in the 100–300 kHz range. This sensitivity is sufficient to detect a few tens of elementary charges stored in nanostructures, and, in specific cases, enables to observe single charge events at room temperature and atmospheric pressure [11]. The lateral resolution of EFM is, however, in general poorer than the AFM resolution in topography mode because electrostatic forces need to be dissociated from short range surface forces such as repulsive or capillary forces. Hence, in ambient air conditions, the tip is in general lifted from the substrate so as to perform EFM measurements. This limits the lateral resolution down to a few tens of nanometers in ambient air at most, which is, however, sufficient to address and map the electrostatic properties of single nanostructures or nanodevices such as carbon nanotubes or CNTFETs.

Phase Shifts Versus Frequency Shifts

Measurements are generally carried out in a two-step procedure in which each EFM scan line is interleaved between topography scan lines. This enables to define a tip pass in the EFM mode at a controlled distance z over to the substrate plane (*see* Fig. 4.1, *right*). EFM data are recorded with the tip mechanically oscillated in the vicinity of its resonance frequency f_0 , and biased at a detection voltage V_{EFM} . Two implementations can then be used. First, the cantilever can be mechanically excited at fixed frequency f'_0 close to f_0 , and the recorded EFM signal is the phase shift $\Delta\phi$ of the cantilever oscillation. Alternatively, a feedback loop can be introduced that maintains the phase shift $\Delta\phi$ as constant, by adjusting the mechanical excitation frequency f'_0 . The EFM signal consists then in the shift Δf of the excitation frequency f'_0 necessary to maintain $\Delta\phi$ as constant. Both procedures lead to qualitatively similar images. The phase shift $\Delta\phi$ can be related to Δf at the resonance frequency using $\Delta\phi/2Q_{\text{cant}} = \Delta f/f_0$. However, because of the nonlinearity of the cantilever phase with respect to frequency, this relation no longer holds for large-frequency shifts (i.e., $\Delta f > f_0/2Q_{\text{cant}}$). Frequency shift measurements, which are directly proportional to force gradients, are thus more appropriate than phase shift measurements for quantitative charge imaging purposes.

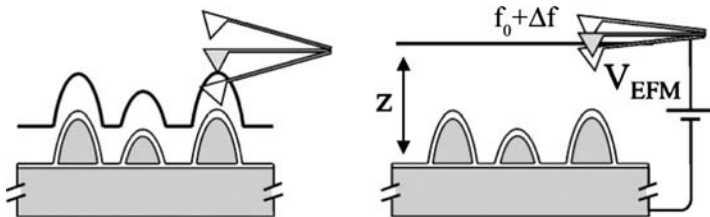


Fig. 4.1. *Left:* Schematics of the topography measurement (here in tapping mode). *Right:* EFM measurement, in which the cantilever is lifted at a distance z above the substrate plane. The cantilever is mechanically excited close to its resonance frequency f_0 , and the EFM signal consists in the cantilever eigenfrequency shift Δf (or in the cantilever phase shift $\Delta\phi$, see text)

Capacitive Versus Charge EFM Signals

We now treat the case of electrostatic forces involved in EFM signals, by describing the tip-substrate as an ideal capacitance $C(z)$ and introducing a surface potential V_S accounting, for example, for the tip-surface work function difference and surface states. Assuming no charge inside the tip-substrate capacitance, the force gradient on the tip is readily expressed as:

$$\frac{\partial F}{\partial z}(z_0) = \frac{1}{2} \frac{\partial^2 C}{\partial z^2} (V_{\text{EFM}} - V_S)^2 \quad (4.2)$$

Capacitive forces thus lead to negative frequency shifts according to (4.1), varying quadratically as a function of $(V_{\text{EFM}} - V_S)$. When a charge Q is now introduced in the tip-substrate capacitance, an effective surface potential V_Q proportional to the charge Q builds in,¹ and the total force gradient can now be developed as follows:

$$\frac{\partial F}{\partial z}(z_0) = \frac{1}{2} \frac{\partial^2 C}{\partial z^2} [(V_{\text{EFM}} - V_S)^2 - 2(V_{\text{EFM}} - V_S)V_Q + V_Q^2] \quad (4.3)$$

Two charge contributions thus appear. The first one corresponds to the interaction between charge Q and the charge at the EFM tip apex, and leads to a frequency shift proportional to $(V_{\text{EFM}} - V_S) \cdot V_Q$. The main advantage of this signal is that it enables a determination of the sign of the surface charge Q . The second charge contribution – proportional to Q^2 – corresponds to image charge effects. This component therefore always corresponds to negative frequency shifts (attractive force gradients) and is amplified when scanning over insulating layers [12, 27]. The three electrostatic contributions of EFM signals can be dissociated experimentally by a spectroscopic analysis of EFM signals [13].

¹ V_Q depends strictly speaking on the tip surface distance z , while V_S would not depend on z , if only related to the tip-substrate work function difference. The determination of V_Q as a function Q will not be described here.

Modulated ($1\omega/2\omega$) EFM/FM-KFM

We now describe the case of modulated EFM, also called $1\omega/2\omega$ EFM, in which the EFM static bias is replaced by an ac + dc voltage: $V_{\text{EFM}}(t) = V_{\text{dc}} + V_{\text{ac}} \cos \omega t$, where $\omega \ll 2\pi f_0$ (quasistatic approximation). The electrostatic force gradient $\partial F/\partial z(z_0)$ now reads: $\partial F/\partial z(z_0) = 1/2\partial^2 C/\partial z^2 (V_{\text{dc}} + V_{\text{ac}} \cos \omega t - V_{\text{S}} - V_{\text{Q}})^2$ and hence exhibits one static component and two oscillating components at 1ω and 2ω :

$$\begin{cases} \partial F/\partial z(z_0)|_{0\omega} = 1/2\partial^2 C/\partial z^2 [V_{\text{dc}}^2 + V_{\text{ac}}^2/2] \\ \partial F/\partial z(z_0)|_{1\omega} = \partial^2 C/\partial z^2 (V_{\text{dc}} - V_{\text{S}}) \cdot V_{\text{ac}} \cos \omega t - \partial^2 C/\partial z^2 (V_{\text{Q}} \cdot V_{\text{ac}} \cos \omega t) \\ \partial F/\partial z(z_0)|_{2\omega} = 1/4\partial^2 C/\partial z^2 V_{\text{ac}}^2 \cos 2\omega t \end{cases} \quad (4.4)$$

This shows that provided V_{dc} can be adjusted to V_{S} (e.g., with the tip positioned over the surface), the force gradient component at 1ω measures the interaction between charge Q and ac charges at the EFM tip, whereas the component at 2ω measures capacitive effects. This technique can be employed to record separately charge and capacitive (dielectric) force gradient images, respectively [11]. Alternatively, the introduction of a feedback loop that sets to zero (or “nullifies”) the force gradient at 1ω enables to measure the local contact potential difference $V_{\text{S}} + V_{\text{Q}}$, which contains the information about the surface potential and local charge Q . Although this technique is derived from EFM (i.e. based on force gradients), it is also called frequency modulation Kelvin force microscopy (FM-KFM) in noncontact AFM, in analogy with Kelvin force microscopy described hereafter, which enables to measure local surface potentials.

4.2.2 Kelvin Force Microscopy

Principle of Amplitude Modulation Kelvin Force Microscopy

KFM (for a historical reference, *see* [10]) is in its usual implementation on atmospheric pressure AFM, a technique based on the electrostatic excitation of cantilevers – in contrast with EFM in which the cantilever is mechanically excited – and on the nullification of its oscillation amplitude. This technique is thus here based on forces, rather than on force gradients. It is now also called amplitude modulation Kelvin force microscopy, in comparison with FM-KFM based on force gradients (*see* Sect. 4.2.1).

More precisely, the electrostatic excitation is generated by an ac + dc voltage that can be applied either to the substrate, a device, or directly to the EFM tip. In this later option, the ac + dc voltage leads to an electrostatic force $F(z)$ at a distance z from the surface: $F(z) = 1/2\partial C/\partial z (V_{\text{dc}} + V_{\text{ac}} \cos \omega t - V_{\text{S}} - V_{\text{Q}})^2$. Following the development of (4.4), this force splits into one static component and two oscillating components at 1ω and 2ω :

$$\begin{cases} F(z)|_{0\omega} = 1/2\partial C/\partial z[V_{\text{dc}}^2 + V_{\text{dc}}^2/2] \\ F(z)|_{1\omega} = \partial C/\partial z(V_{\text{dc}} - V_{\text{S}}) \cdot V_{\text{ac}} \cos \omega t - \partial C/\partial z V_{\text{Q}} \cdot V_{\text{ac}} \cos \omega t \\ F(z)|_{2\omega} = 1/4\partial C/\partial z(V_{\text{ac}}^2 \cos 2\omega t) \end{cases} \quad (4.5)$$

Here, however, the ac angular frequency is generally set close to the eigenfrequency of the cantilever: $\omega \sim 2\pi f_0$, so that $F(z)|_{0\omega}$ and $F(z)|_{2\omega}$ fall out of resonance with the cantilever, whereas $F(z)|_{1\omega}$ leads to an oscillation amplified by the cantilever quality factor Q_{cant} (typically $Q_{\text{cant}} \sim 200$ at atmospheric pressure). A feedback loop is then introduced, so as to nullify the cantilever oscillation amplitude at 1ω , by adjusting the dc component of the electrostatic excitation. From (4.5), this condition is obtained provided $V_{\text{dc}} = V_{\text{S}} + V_{\text{Q}}$, that is, when the dc bias on the tip exactly matches the local surface potential $V_{\text{S}} + V_{\text{Q}}$: KFM therefore enables the real-time measurement and mapping of local surface potentials.

Open-Loop KFM or ac-EFM

This section briefly discusses the case of open-loop KFM measurements, which consist in mapping the cantilever oscillation amplitude at 1ω , when electrostatically excited, but in absence of KFM regulation loop. This technique is used so as to generate an image of the force at 1ω at the tip, proportional to $\partial C/\partial z(V_{\text{dc}} - V_{\text{S}}) \cdot V_{\text{ac}}$ as seen in (4.5). It should be noted that this mode has been applied to carbon nanotubes and carbon nanotube devices [14]. It was then referred to as “ac-EFM,” due to the absence of feedback loop measuring the local surface potential, whereas, however, the “ac-EFM” signal is the one of an open-loop KFM mode, that is based on forces, and not on force gradients like in EFM techniques.

4.2.3 Lateral Resolution in EFM and KFM

Side Capacitance Effects

We describe here a difference between EFM and KFM as for quantitative charge and/or surface potential imaging. In both techniques, the actual detection is performed locally, that is not only through a capacitance $C(z)$ between the tip apex and surface, but also through a series of side capacitances $C_{1\text{t}}(z)$, $C_{2\text{t}}(z)$, \dots , $C_{i\text{t}}(z)$, “seeing” effective surface potentials Φ_i , as described in Fig. 4.2.

In the case of KFM, the nullification of the total force at 1ω at the cantilever in presence of side capacitances leads to the following equation:

$$\partial C/\partial z(V_{\text{dc}} - V_{\text{S}} - V_{\text{Q}}) \cdot V_{\text{ac}} + \sum \partial C_{i\text{t}}/\partial z(V_{\text{dc}} - F_i) = 0$$

so that the measured dc component regulated by the KFM loop becomes an average between $(V_{\text{S}} + V_{\text{Q}})$ and the potentials Φ_i seen by side capacitances

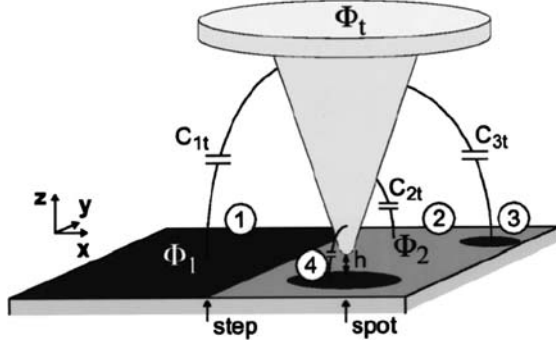


Fig. 4.2. Schematics of the side capacitances C_{1t} , C_{2t} , ..., C_{it} between the cantilever tip and the surface with inhomogeneous local surface potentials Φ_1 , ..., Φ_i . Adapted from [15]

C_{it} [15], the relative weight of each of the surface potential Φ_i below the tip being determined by the derivatives of the local capacitances $\partial C/\partial z$ or $\partial C_{it}/\partial z$.

$$V_{dc} = \frac{\partial C/\partial z'(V_S + V_Q) + \sum \partial C_{it}/\partial z'F_i}{\partial C/\partial z + \sum \partial C_{it}/\partial z} \quad (4.6)$$

The assessment of side-capacitance effects is of great importance, for example, for quantitative voltage drop measurements along carbon nanotubes, described later in this chapter. Additionally, one might notice that in the case of EFM or FM-KFM (*see* Sect. 4.2.1), the weights in (4.6) are to be replaced with the second derivatives of the capacitances. This reduces the effects of side-capacitances, due to the increased weight of $\partial^2 C/\partial z^2$ with respect to the sum of the $\partial^2 C_{it}/\partial z^2$ in techniques based on force gradients (i.e., on phase or frequency shifts), as compared with the weight of $\partial C/\partial z$ with respect to the sum of the $\partial C_{it}/\partial z$ in techniques based on the measurement of forces (i.e., on amplitudes).

This lateral resolution effect is illustrated in Fig. 4.3. It shows the topography (Fig. 4.3a) of a connected SWCNT on a SiO₂ surface (*top*), as well as the corresponding EFM image (*middle*, negative EFM tip bias, 3 Hz color scale) and the corresponding KFM image (*bottom*, 200 mV color scale) acquired in similar detection conditions. The nanotube charge environment (here negative) is evidenced by the bright contrast along the nanotube in the EFM image (repulsive interaction between the negative charge in the nanotube environment and the negative charge at the EFM tip) and a dark contrast in the KFM image (negative potential along the nanotube). In addition, a positive local spot of charges has been introduced in the SiO₂ layer in the vicinity of the nanotube by a charge injection technique (to be described later in this chapter), which is monitored as a dark spot in the EFM image (attractive interaction between the positive spot of charges and the negatively biased EFM tip) and as a positive spot in the KFM surface potential image. The

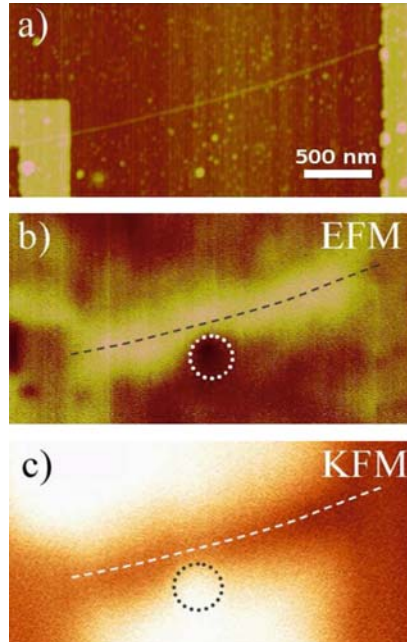


Fig. 4.3. (a) Topography image of carbon nanotube field effect transistor (CNTFET) (500 nm scale bar). (b) EFM image (tip-substrate distance $z = 80$ nm, $V_{\text{EFM}} = -2$ V, 3 Hz color scale) showing (1) the negatively charged environment of the CNTFET channel (bright contrast of the nanotube) and (2) a local spot of positive charge injected in the SiO_2 layer in the vicinity of the nanotube (*dark spot* in the EFM image), serving here as a test bench for the lateral resolution of electrostatic measurements. (c) Corresponding KFM image ($z = 70$ nm, 200 mV color scale), in which the nanotube charged environment corresponds to a negative potential (*dark contrast*) and the local spot of positive charge as a positive potential (*bright contrast*). The difference in lateral resolution in EFM and KFM is visible from (b) and (c) (*see text*)

difference in lateral resolution between the EFM and the KFM images is obvious from Fig. 4.3b, c, both along the nanotube and for the local spot of charges. The lower lateral resolution in the KFM image illustrates the bigger weight of side capacitance in force techniques compared with force gradient techniques, as explained previously.

Carbon Nanotube Tip Probes

Finally, we also mention the use of carbon nanotube probes for electrostatic (EFM or KFM) measurements. An example of such a probe [16] is given in Fig. 4.4, showing a multiwalled carbon nanotube tip mounted on a standard AFM cantilever tip.

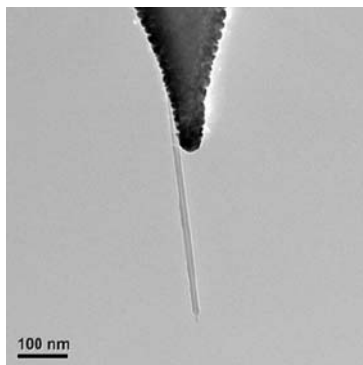


Fig. 4.4. Transmission electron microscopy image of a cantilever tip modified with a glued multiwalled carbon nanotube. Adapted from [16]

As can be obviously seen from Sect. 4.2.3, such a high-aspect ratio tip is ideal to reduce parasitic effects associated with side capacitances, leading to increased lateral resolution, which has been demonstrated in the case of dopant profiling [16], and on contact potential measurements on biased Al/Al₂O₃/Al junctions [17] as well as on bundles of nanotubes [17] or individual MWCNT [18]. This effect is illustrated in the Fig. 4.5, in which the same nanotube has been imaged with a carbon nanotube probe with 200 nm length and ~ 20 nm apex diameter (Fig. 4.5, *left*) and with a conventional Au-plated silicon tip probe (Fig. 4.5, *right*). Although the topography shows a comparable lateral resolution in both cases, the associated KFM images demonstrate the improvement of the lateral resolution using the carbon nanotube tip probe.

4.3 Electrostatic Imaging of Carbon Nanotubes

This section describes the work done on imaging the dielectric properties and native charge in carbon nanotubes using EFM. A large interest has been shown recently on such experiments because EFM can be used to image nanotubes either deposited or directly grown on insulating layers [19–22] or buried inside thin insulating layers [23]. The main physical property imaged in this work is the conductive nature of nanotubes with respect to their insulating environment, giving rise to highly contrasted EFM images, with the possibility to even “reveal” carbon nanotubes not visible in topography images. Hence, EFM had been termed “scanning conductance microscopy” – although no conductance is truly measured, but rather capacitive forces, as explained in Sect. 4.2. This section provides a comprehensive review of these experiments.

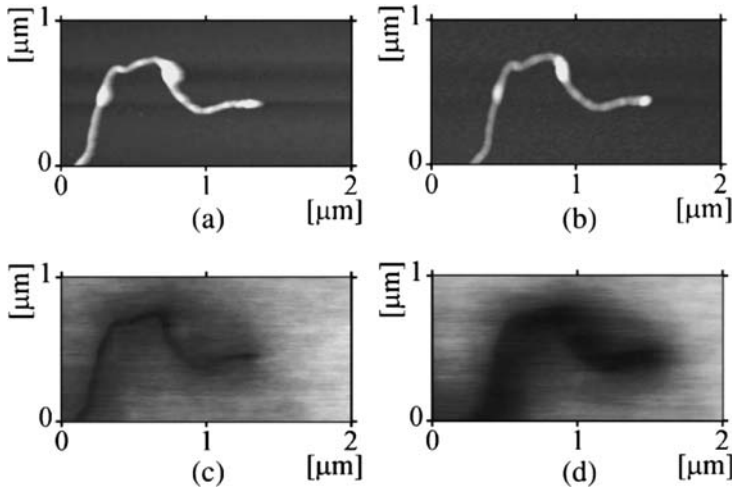


Fig. 4.5. *Top:* Topography of a multiwalled carbon nanotube (diameter ~ 20 nm) taken with a multiwalled carbon nanotube tip (*left*) and an Au-plated silicon cantilever tip (*right*). *Bottom:* Corresponding KFM images (gray scale ~ 100 mV), showing an enhanced lateral resolution with the carbon nanotube tip probe. Adapted from [18]

4.3.1 Capacitive Imaging of Carbon Nanotubes in Insulating Layers

The mapping of individual nanotubes buried in thin poly-(methyl methacrylate) (PMMA) films has been reported by Jespersen et al. [23], in a way that illustrates the capacitive detection performed by EFM. Figure 4.6a shows a schematics of the composite sample made of SWCNTs dispersed in PMMA and spin-coated as an ~ 60 nm-thin film on a SiO_2 -covered doped silicon substrate. In addition, it shows the principle of EFM detection, performed in a similar way as described in Sect. 4.2, here, however, using the cantilever phase shift to track electrostatic force gradients. The topography image of the sample is shown in Fig. 4.6b. As nanotubes are buried into the PMMA film, they cannot be revealed from this image that only shows the presence of small particles with height ~ 10 – 20 nm, stemming from the carbon nanotube powder dispersed in PMMA. This contrasts with the EFM image (Fig. 4.6c) acquired simultaneously with topography, where single or agglomerated carbon nanotubes can be revealed and imaged as dark lines in the EFM image. The EFM signal stems from the much higher conductance of SWCNT compared with the PMMA matrix, leading to a significant increase of the tip-sample capacitance because of the presence of the carbon nanotube in this capacitance, and, thus, to a cantilever negative phase shift (*see* Sect. 4.2 for details).

The contrast of the EFM image is further analyzed, and shows strong variations in the phase shift amplitude (*see* Fig. 4.7, *left*, for illustration) even

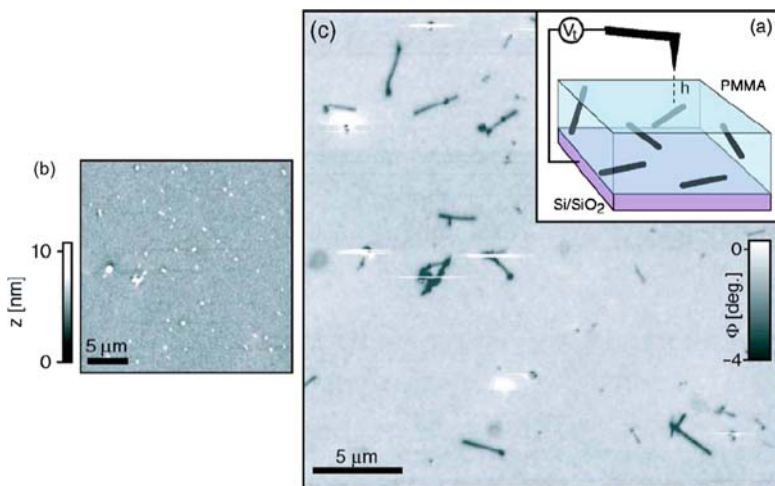


Fig. 4.6. (a) Schematic illustration of the polymer/SWCNT sample and EFM operation. (b) Topography image of the 60 nm-thick film of PMMA/SWCNT composite. Because of the polymer, the tubes cannot be observed. (c) Corresponding EFM image (tip-substrate distance $h = 35$ nm, tip biased at +7 V), in which individual SWCNTs are clearly seen as dark lines (negative phase shifts). Adapted from [23]

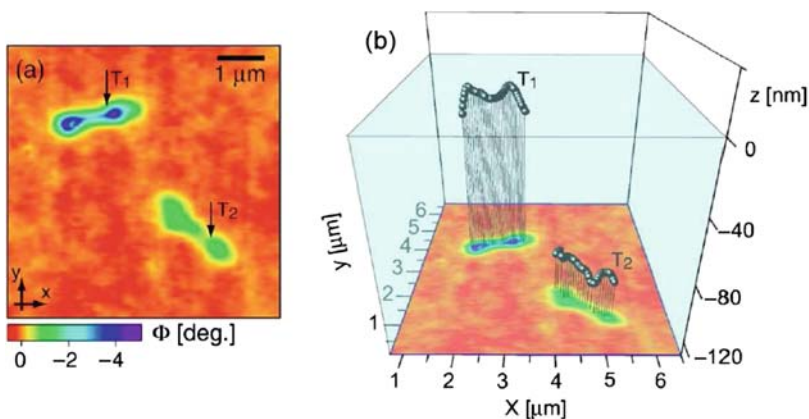


Fig. 4.7. *Left:* EFM image showing two SWCNTs embedded in an ~ 170 nm-thick film of SWCNT/PMMA composite (5° phase shift color scale). *Right:* Projection view of the three dimensional map of the two nanotubes as inferred from the modeling of EFM signals. The blue region illustrates the PMMA matrix. Adapted from [23]

for nanotubes of similar length. This is explained by the fact that the nanotube can be incorporated in the PMMA matrix either (1) close to the substrate (*see*, e.g., the nanotube labeled T2 in Fig. 4.7) or (2) close to the PMMA layer surface (*see* nanotube T1). This situation is depicted in Fig. 4.7 (*right*), and was quantitatively linked with the EFM image phase shift contrasts associated with the nanotube [23], showing the EFM can both reveal and locate carbon nanotubes in the PMMA matrix, although not visible in topographic images.

4.3.2 EFM Imaging of Carbon Nanotubes and DNA

Another use of EFM to assess the dielectric properties of one-dimensional nanostructures on surfaces is handled in [19], in which the case of SWCNTs and DNA single strands has been debated. The EFM detection is first discussed in the case of SWCNTs on oxide surface with different lengths, as shown in Fig. 4.8. Images with similar contrast as in Fig. 4.6 are obtained, here showing also a dependence on the EFM contrast as a function of the nanotube length L . This is interpreted as stemming from the variation of the nanotube-substrate plane capacitance as a function of the L , leading to the following dependence for the cantilever phase shift $\Delta\phi$: $(\Delta\phi)^{-1/2} = a + bL^{-1}$ (*see* [19] for details), in agreement with experimental data (*see* Fig. 4.8, *lower inset*).

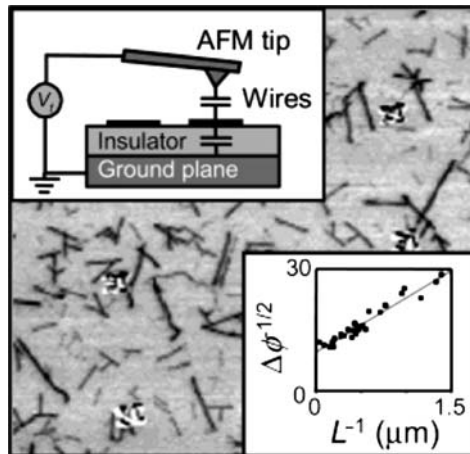


Fig. 4.8. *Upper inset:* experimental setup. An AFM cantilever is driven near its resonance frequency. A “scanned conductance image” is obtained by producing a gray scale plot of the phase shift ϕ of the cantilever as a function of tip position. (*Main panel*) Scanned conductance image for SWCNTs. Dark lines indicate negative phase shifts $\Delta\phi$ occurring when the tip is over a tube. (*Lower inset*) Plot of $\Delta\phi^{-1/2}$ as a function of the inverse tube length L^{-1} for 26 individual tubes. The data follow a straight line (*see* text). Adapted from [19]

The case of SWCNTs is then compared with the case of λ -DNA strands deposited on the same surface. Since both molecules show the same height (~ 1 nm) in the topography images, the identification of carbon nanotubes and λ -DNA molecules has been carried out by imaging first the topography of a nanotube dispersion on the substrate (data not shown) and then imaging the same surface after λ -DNA deposition using recognition markers predefined by lithography (*see* yellow feature in Fig. 4.9a), enabling an unambiguous identification of the λ -DNA strands.

The topography of the surface is illustrated in Fig. 4.9a, where the λ -DNA molecules are green, and the carbon nanotubes appear in red. A high-resolution $4.5 \times 4.5 \mu\text{m}^2$ zoom of the blue box is shown in Fig. 4.9b, together with the corresponding EFM image in Fig. 4.9c. It is striking in Fig. 4.9c that only the capacitive features associated with the carbon nanotubes can be

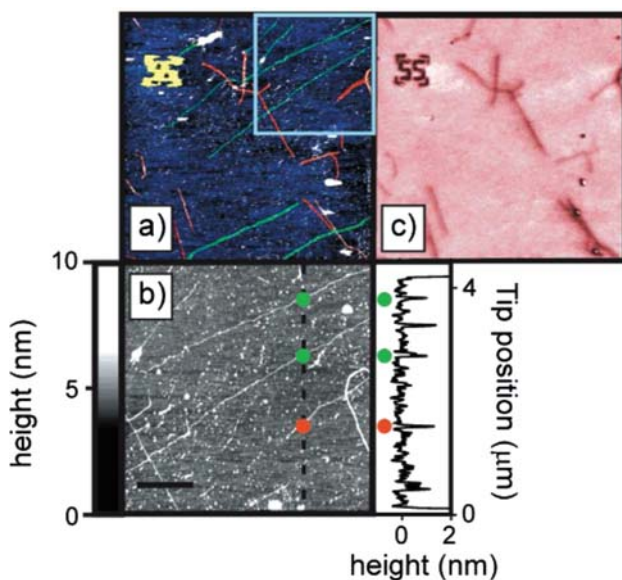


Fig. 4.9. (a) Colorized topographic image of SWCNT bundles and single tubes, colored red, and λ -DNA molecules, colored green. A $1 \mu\text{m}$ alignment mark is shown in yellow. (b) $4.5 \times 4.5 \mu\text{m}^2$ topographic scan of the region enclosed by the *blue square* in (a). The higher resolution scan shows the topographic features more clearly than it is possible with the larger area scan in (a). The gray level scale is shown to the immediate left of the image. The scale bar is $1 \mu\text{m}$. A SWCNT is marked by a *red dot*, whereas two λ -DNA molecules are marked by *green dots*. The graph to the immediate right of the panel shows a line trace through the topographic image along the *dotted line*, with a constant offset height subtracted from the data. Both the SWCNT and λ -DNA molecules are ~ 1 nm in height. (c) EFM image of the region shown in part (a). Although the SWCNT and the λ -DNA are similar in height, no signal appears in the EFM image when the tip is over the λ -DNA molecules. Adapted from [19]

seen in the EFM image, whereas λ -DNA strands cannot be resolved. This was interpreted as a proof that λ -DNA molecules consist in an electrical insulator with an extremely low conductivity, in contrast with SWCNTs² [19].

4.3.3 Imaging of Native Charges in Carbon Nanotube Loops

Native charges have been observed in the environment of SWCNTs directly grown by chemical vapor deposition on SiO₂ surfaces, and more precisely when the nanotubes are forming coils on the surface [21]. This is illustrated in Fig. 4.10a, for which an EFM image is displayed with a negative bias V_S applied to the sample. This image shows the usual capacitive footprint of the nanotube topography [here the topography is not shown, *see* [21] for details], appearing as a negative cantilever phase shift. The striking feature here is the appearance of a homogeneous negative phase shift signals associated with some of the carbon nanotube coils, such as those selected in Fig. 4.10. This signal corresponds to an attractive force, and thus, due to $V_S < 0$, to a negative charge stored in the coil, with density in the range of a few $-10^{-8} \text{ C m}^{-2}$, that is $\sim 10^3$ electrons stored in a loop with $\sim 1 \mu\text{m}$ radius. Those charges were

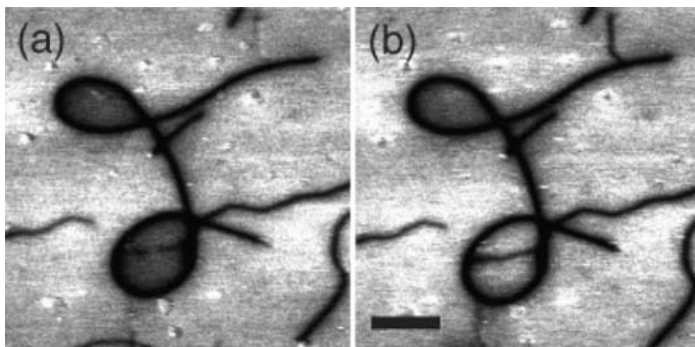


Fig. 4.10. (a) EFM image of two loops formed on the same single-walled nanotube, each trapping charges as seen by the phase difference measured in their interior. (b) EFM image of same area as (a) after touching the interior of the lower loop by the grounded AFM-tip. All measurements were performed with $z = 60 \text{ nm}$ and negative substrate bias $V_S = -6 \text{ V}$. The scale bar is $2 \mu\text{m}$. Adapted from [21]

² Although this work brought a spectacular scanning-probe insight in the debate about the DNA conductivity, it should be nonetheless mentioned that (1) EFM experiments do not rigorously bring information about conductance, but rather about the polarisability (dielectric constant) of the molecules; and (2) the above analysis should be only valid when the charge environment of the nanotube or DNA molecules are evaluated, since EFM is also sensitive to local charges, as shown in Sect. 4.2. This is especially relevant for DNA deposited on insulators, for which a positive charge is found in the surrounding of the DNA ropes [25]. This point is also likely to strongly depend on preparation protocols.

attributed to a native charging of the sample silicon dioxide inside the coils after growth. Moreover, it was shown possible to remove the coil charge by bringing the grounded tip in contact with the inside of the coil (*see* Fig. 4.10a, b for a comparison). Since the charges were uniformly distributed within the loops, this also suggests that the charges are relatively mobile on the surface and not trapped deeply in the substrate, in such type of samples.

4.4 Charge Injection Experiments in Carbon Nanotubes

The results described earlier demonstrate the relevance of electrical techniques derived from AFM to establish the local electrostatic properties of individual nanotubes on insulating layers, and already put forward the interplay between their dielectric and charging properties. It is the aim of this section to develop these aspects, based on experiments that manipulate the charge state of individual carbon nanotubes, that is charge injection experiments. First, we describe the charge injection technique, then we illustrate EFM signals, and finally, we review the properties of charged carbon nanotubes: charging of individual carbon nanotubes and nanotube networks, field-emission properties from charged nanotubes, charging mechanisms and inner-shell charging in the case of MWCNT, and finally, charge enhancements at the nanotube caps, associated with Coulomb repulsion.

4.4.1 Charge Injection and Detection Techniques

Charge injection and detection experiments are typically conducted as follows, using standard AFMs with EFM modes in ambient air (here Multimode/Nanoscope IIIA or Dimension/Nanoscope IV microscopes from Veeco Instruments, Santa Barbara, US). First, the microscope is placed under a dry nitrogen atmosphere in order to avoid (1) charge leakage due to adsorbed surface water and (2) anodic oxidation during the charge injection experiments. Commercial Pt-Ir-coated cantilevers with ~ 60 kHz resonance frequency and with low ($1\text{--}3\text{ N m}^{-1}$) spring constants are used, so as to address single nanotubes in contact mode, but also to facilitate EFM detection (i.e., cantilever phase or frequency shifts).

Charging experiments are achieved by first localizing individual CNTs from the sample topography acquired in intermittent contact mode AFM. The simplest implementation for charge injection into a given nanotube is to use local approach-retract curves with the EFM cantilever biased at an injection voltage V_{inj} (typically $V_{\text{inj}} = -2$ up to -12 V) with respect to the sample substrate [12, 25, 27]. In this process, the distance z between the biased EFM tip and the substrate is periodically swept (~ 1 Hz sweep rate), and its minimum value is tuned in order to reach a permanent contact regime between the EFM tip and the CNT (*see* Fig. 4.11) during the z -sweep. In this regime, the cantilever oscillation amplitude becomes zero, and a cantilever quasistatic

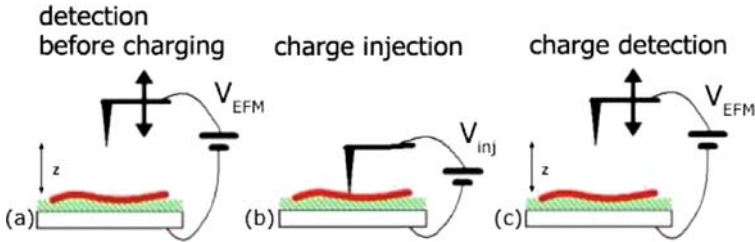


Fig. 4.11. Schematics of the charge injection and detection technique. (a) EFM data acquired before charge injection. The recorded EFM signal is the cantilever resonance frequency shift when the tip is biased at V_{EFM} and scanned at a constant distance z above the sample. (b) Charge injection with the tip in contact with the nanotube and biased at V_{inj} with respect to the substrate. (c) EFM data acquisition after injection. Adapted from [12]

deflection is monitored. This enables to adjust the tip-CNT contact force (here a few nano Newtons, as measured from the cantilever static deflection) and the contact duration (typically 100 ms for tip-substrate distance sweeps with 1 s period). Charging times given here after (typically 1–4 min) always refer to the total duration of the approach-retract sweeps, whereas, the cumulated duration of the permanent contacts between the EFM tip and the CNT during the approach-retract sweeps, therefore corresponds to approximately one tenth of the charging process duration.

The global procedure for charge injection and detection is described in Fig. 4.11. First, EFM and topography images are acquired in an interleave scheme, such as described in Sect. 4.2.1, with the tip biased at a detection voltage V_{EFM} and lifted at a distance z above the substrate during the EFM pass. Then, a charge Q is injected in the nanotube using a tip voltage V_{inj} during the charge injection procedure. Finally, EFM data are acquired again after charge injection, with the tip bias set at V_{EFM} .

4.4.2 Experimental Illustration of EFM Signals

First of all, we present charging experiments [25] performed on a 19 nm diameter MWCNT (topography image in Fig. 4.12a) deposited on a 200-nm-thick SiO_2 layer grown on a doped silicon substrate. The EFM image acquired with $V_{\text{EFM}} = -3$ V before charge injection is shown in Fig. 4.12b that reveals a dark footprint of the nanotube. The EFM signal is here the cantilever frequency shift Δf required to maintain a constant cantilever phase in the EFM pass (*see* Sect. 4.2.1). In absence of charge, the frequency shift Δf is due to the increase of the tip-substrate capacitance during the linear pass of the tip over the CNT. This capacitive signal is thus proportional to $(V_{\text{EFM}} - V_{\text{S}})^2$ (V_{S} being here negligible) and corresponds to a negative frequency shift (dark feature associated with the nanotube in Fig. 4.12b), which physically corre-

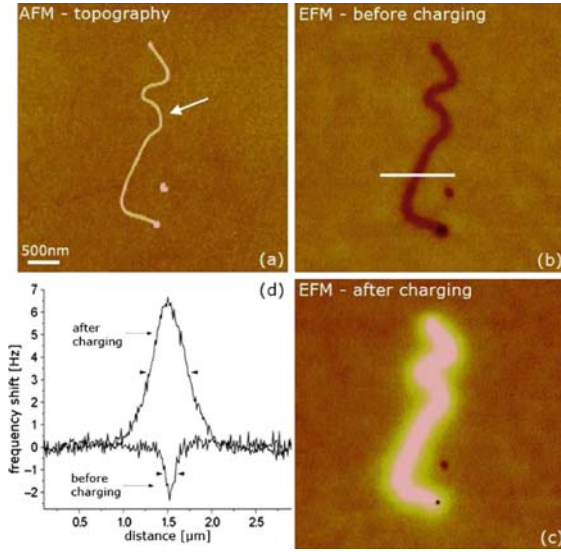


Fig. 4.12. (a) Atomic force microscopy image of MWCNT of ~ 19 nm diameter on a 200 nm silicon dioxide surface. (b) EFM image ($V_{\text{EFM}} = -3$ V) acquired before charging (10-Hz color scale) with a tip-substrate distance $z \sim 100$ nm. (c) EFM image taken in the same conditions, after a charge injection performed at the point indicated by an *arrow* in (a), using $V_{\text{inj}} = -5$ V for 2 min. (d) Cross-sections of the EFM scans for the case (b) and (c). The position of the cross-section is shown by the horizontal bar in (b). The full width at half minimum (FWHM) of the EFM signals before and after charging (resp. ~ 100 and 440 nm) is marked by the *triangles* as a guide to the eye. Adapted from [12]

sponds to a local increase of the attractive capacitive force on the cantilever when the tip is moved above the nanotube (*see* Fig. 4.1 in Sect. 4.2).

This behavior, however, no longer holds when the CNTs are charged because the total electric force also includes a component proportional to the nanotubes charge Q and to V_{EFM} (*see* Sect. 4.2). This is illustrated in Fig. 4.12c, showing the EFM signals after a charge injection experiment using $V_{\text{inj}} = -5$ V during 2 min at the injection point indicated by the *arrow* in Fig. 4.12a. The delocalization of the injected charge along the MWCNT is visible from the EFM image of Fig. 4.12c (taken after the injection), in which the nanotubes appear homogeneously bright. This contrast is due to $V_{\text{inj}} < 0$ (the nanotubes are negatively charged), and to the negative V_{EFM} values used for charge detection ($V_{\text{EFM}} = -3$ V in Fig. 4.12). The force component proportional to $Q \times V_{\text{EFM}}$ leads therefore to a positive frequency shift imaged as a bright feature in the EFM image (*see* Fig. 4.12c). It physically corresponds to the (here repulsive) interaction between the stored charge and capacitive charges at the EFM tip. An estimation based on numerical calculations gives here a nanotube linear charge density of $\sim 160|e|/\mu\text{m}$ [12, 30].

However, it should be mentioned that charge, capacitive and image charge force gradients compete altogether, so that even for $V_{\text{EFM}} \times V_{\text{inj}} > 0$, charged nanotubes can only be imaged as bright features provided charge signals predominate (1) over capacitive EFM signals proportional to $(V_{\text{EFM}} - V_{\text{S}})^2$ and (2) over the image charge contribution, which is proportional to Q^2 (*see* Sect. 4.2), which is enhanced when scanning on insulators. Such imaging conditions can be reached either for sufficiently charged CNTs or by lowering the detection bias V_{EFM} . These scanning conditions have been set for all EFM images presented in this section, so that dark and bright features will always correspond – unless explicitly stated – to capacitive and charge effects, respectively.

We now discuss the width of the cross-sections of EFM signals that are displayed in Fig. 4.12d, and have been taken at the CNT position indicated by the horizontal bar in Fig. 4.12b. The full width at half maximum (FWHM) of the EFM signal of the uncharged nanotubes equals ~ 100 nm, and is already larger than the apparent CNT diameter in the topography image, as can be seen from the AFM and EFM images in Fig. 4.12a, b, respectively. This point simply illustrates the lower spatial resolution of EFM compared with topography measurements, which is due to long-range electrostatic forces and the large (80–100 nm) tip-surface distance during EFM data acquisition. A more striking feature is the broadening of the EFM signals when the nanotube is charged (440 nm FWHM), as seen from Fig. 4.12d and from the comparison between Fig. 4.12b, c. This effect does not occur in EFM experiments performed in nanostructures deposited on conductors [13], where the width of EFM signals remains unchanged after charging. These two situations can be explained by the difference in the electrostatic screening induced by the substrate supporting the nanostructures [*see* [12, 25] for more details]. This change of FWHM for the charged nanotube is in practice a demonstration that the charge carried by the nanotube prior to the charging experiment is negligible, in contrast, for instance, with the case of DNA ropes deposited on similar substrates [24].

Finally, the delocalization of the injected charge has been found to exceed a few tens of microns along the CNTs [12, 26], using charge injection experiments on carbon nanotube networks (*see* Fig. 4.13). The EFM image of the ~ 40 μm as-grown nanotube network is presented in Fig. 4.13a (EFM image), showing the dark capacitive footprint of uncharged nanotubes. The black arrow indicates the injection point, and the injected charge is found to be delocalized along the whole network, as visible in Fig. 4.13b recorded after charge injection.

Abrupt Discharging Processes in Carbon Nanotubes

This section focuses on the phenomena associated with discharge processes in carbon nanotubes. First, we discuss abrupt discharging at single or multiple points along the CNT and at the CNT caps [25], which unambiguously demonstrate charge storage *in* carbon nanotubes. Next, we discuss the field emission

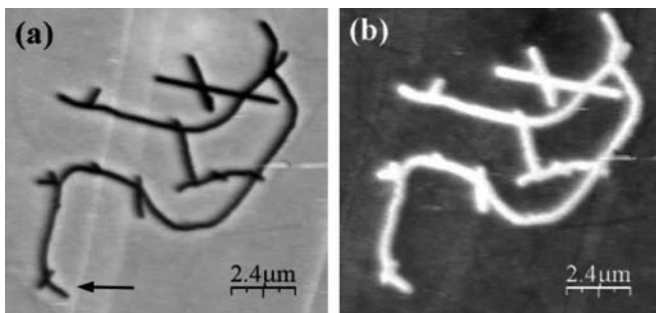


Fig. 4.13. (a) EFM image (frequency shift) of long SWCNTs network, performed before charge injection. The *black arrow* indicates the injection point. The dark nanotubes footprint corresponds to the uncharged network. (b) EFM scan after charging. The bright feature corresponds to a nearly uniform delocalized charge injected from one of the network extremity. Adapted from [26]

of charges to the oxide layer supporting the CNTs, and its enhancement at the CNT caps.

A sudden discharge behavior of the charge injected in a MWCNT with ~ 18 nm diameter and $\sim 2.3 \mu\text{m}$ [12, 27] is presented in Fig 4.14. The point used for charge injection experiments is indicated by the arrow in the topography image (Fig. 4.14a). The EFM signal before charging (Fig. 4.14b) shows a negative frequency shift equal to 1.2 Hz. After the injection ($V_{\text{inj}} = -6$ during 3 min), the EFM scan is started from the bottom to the top (Fig. 4.14c). The nanotube appears first as homogeneously charged (bright feature with maximum frequency shift $+8.2$ Hz), but then undergoes three abrupt discharges at the points labeled d1, d2, and d3, each discharge occurring within one scan line. The CNT EFM image was then found stable on further scanning (no more sudden discharges), indicating that the bright halo surrounding the CNT capacitive footprint after the third discharge in Fig. 4.14c is due to charge emitted by the CNT and trapped in the oxide surface [25]. This effect will be discussed in the following section. The CNT frequency shifts before and after d1 are compared in Fig. 4.14d, and show that the capacitive footprint of the CNT topography can be already clearly identified after the first discharge, whereas the EFM signal was dominated by the CNT charge just after injection.

To associate the discharge locations at specific nanotubes points, an injection experiment has been made again at the same nanotube point using $V_{\text{inj}} = -7$ V. The resulting EFM image scanned from top to bottom is shown in Fig. 4.14e, and was acquired ~ 2 h after the scan in Fig. 4.14c. Two abrupt discharges are evidenced in Fig. 4.14e, and correspond to the discharge points labeled d3 and d2 in Fig. 4.14c. This demonstrates that abrupt discharges are induced by specific points (likely due to the defects) along the MWCNT. At the discharge points, the CNT charge is field-emitted back to the EFM tip

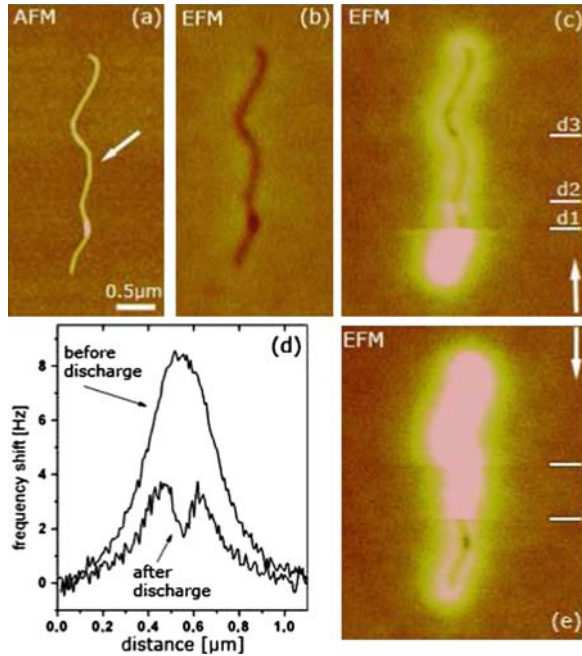


Fig. 4.14. (a) AFM topography picture of $\sim 2.3\text{-}\mu\text{m}$ -long MWCNT with $\sim 18\text{ nm}$ diameter. (b) EFM image acquired before charging. The capacitive footprint of the nanotube topography is visible and corresponds to a 1.2 Hz negative frequency shift. (c) EFM scan after charge injection with the tip biased at $V_{\text{inj}} = -6\text{ V}$ for 3 min at the injection point marked in (a). The direction of data acquisition is indicated by the *vertical arrow*. Three discharge points (d1, d2, and d3 labeled in the figure) successively occur during the CNT scanning. After the discharge, the CNT capacitive footprint is clearly visible in the EFM image, however, surrounded by a bright halo. This EFM image proved stable on further scanning (no observable discharge). (d) EFM signal taken across the CNT before and after the sudden discharge labeled d1 in (c). (e) EFM scan after injection with the tip biased at $V_{\text{inj}} = -7\text{ V}$. The image is scanned from top to the bottom (*see vertical arrow*). The two abrupt frequency shift drops that are visible in this scan correspond to the position of the discharges labeled d3 and d2 in (c). Adapted from [12]

apex that is grounded and intermittently brought into contact with the CNTs during the topography pass.

Charge Emission to the Oxide

The EFM images of Fig. 4.14c, e (after discharge) also display a residual negative charge along the nanotubes in the form of a bright halo around the nanotubes. The actual issue is to determine whether this negative charge corresponds to electrons still stored *in* the CNT, or to charge emitted by the CNT

and trapped on the oxide surface. If not discussed, the issue of surface charge emission can lead to potential misinterpretations of EFM images [12, 26–28]. In Fig. 4.14c, e, the bright halos observed after charging can be unambiguously attributed to charges trapped on the oxide surface for the following two reasons: (1) the residual charge pattern cannot be removed by discharging the nanotubes (i.e., performing a charge injection experiment on the CNT using $V_{\text{inj}} = 0 \text{ V}$) and (2) the residual charge pattern decays very slowly with time (order of a few hours), so that no evolution can be observed in the first subsequent EFM images acquired after the abrupt discharge phenomena. This behavior strongly contrasts with that of charge injected *in* the CNTs, which can be either removed by a local charge injection using $V_{\text{inj}} = 0 \text{ V}$ or abruptly emitted back to the EFM tip as described in the previous section. Moreover, the decay times of oxide charges stored in the vicinity of the nanotubes, which are found to be much longer than decay times on the bare oxide ($\sim 40 \text{ min}$, *see* [12, 27]). The decay of charge injected in bare thin oxide layers has been proposed to result from the attraction between the oxide charge and their substrate images [29], leading to a gradual move of oxide charges toward the conductive substrate electrode. Here, the oxide charge near the CNTs is also subjected to an attractive image force from the CNTs themselves. This attractive interaction is consistent with the much longer retention times observed for oxide charge along the CNTs [25].

Continuous Discharge Processes

Interestingly, SWCNTs do not exhibit the abrupt discharging phenomena as in the case of the nanotube with bigger diameter [12, 27]. Here we discuss the fast continuous discharging of SWCNTs giving as an example the 1 nm diameter nanotube shown in Fig. 4.15a. The EFM scan taken before injection is shown in Fig. 4.15b. The nanotubes and metal markers appear here as dark features, corresponding to uncharged objects. The EFM scan after injection is displayed in Fig. 4.15c and shows the delocalization of the injected charge along the whole SWCNT length ($\sim 30 \mu\text{m}$), whereas the square metal markers remain uncharged. Clearly, no abrupt discharge occurs in Fig. 4.15c, while the CNT has been continuously scanned as shown in Fig. 4.15d and e, respectively, the time between two successive EFM images being $\sim 40 \text{ min}$. A discharge – though not abrupt here – is evidenced by comparison of the SWCNT EFM images of Fig. 4.15c, d, whereas no significant change is observed on further scanning, as seen by comparison of Fig. 4.15d, e, and from the following EFM scans (not shown here). It is thus reasonable to conclude that the SWCNT in Fig. 4.15 undergoes a continuous discharge process.

Similar continuous discharging processes can be observed for MWCNTs, but, in contrast to SWCNTs, the discharge process shows a much longer timescale. An example is presented in Fig. 4.16. An $\sim 30\text{-nm}$ -diameter MWCNT shows only a weak decay of the total charge signal 13 h after charge injection (*see* Fig. 4.16a, b). In this case, a blurring of the MWCNT charge

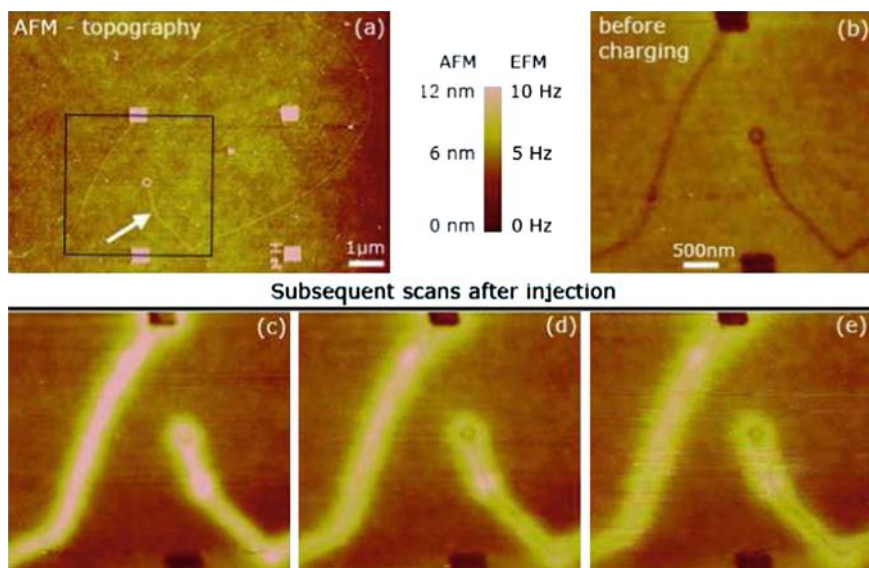


Fig. 4.15. (a) AFM topography image of a SWCNT with 1 nm diameter. $500 \times 500 \text{ nm}^2$ metal squares separated by a $5 \mu\text{m}$ pitch are visible in the image. One of the caps of the SWCNT appears to be ring-shaped. The length of the nanotubes is $\sim 32 \mu\text{m}$. The arrow indicates the point used for charge injection experiments. (b) EFM image of the area indicated by the *black square* in (a), acquired before injection ($V_{\text{EFM}} = -3 \text{ V}$). (Inset) Color scale bars for the AFM topography and the EFM scans. (c) EFM image after a charge injection experiment with $V_{\text{inj}} = -5 \text{ V}$ for 3 min. The delocalization of the charge along whole nanotubes is visible. A *bright spot* can also be seen at the injection point, showing that charges were also injected locally into the dielectric surface. (d, e) Successive EFM scans acquired after the injection shown in (c). The SWCNT discharge is visible by comparison of (c) with (or to) (d), while no further discharge is observable afterward, as seen by comparison of (d) to (e). Adapted from [12]

signal can be observed in Fig. 4.16b. This is seen more clearly on the frequency shift cross-section in Fig. 4.16c, also showing the resurgence of the MWCNT capacitive frequency shift dip associated with the larger FWHM of the EFM signal 13 h after charge injection. Since the total EFM signal is approximately unchanged between Fig. 4.16a, b, it is assumed that the observed blurring of the EFM image corresponds to a slow emission of the MWCNT charge to the oxide, suggesting that the injected charges has been here stored in the nanotube for hours.

Nanotube Charge Versus Oxide Charge

Here, we illustrate the need for correct interpretations of EFM signals to demonstrate charge storage in SWCNTs and MWCNTs with respect to charge

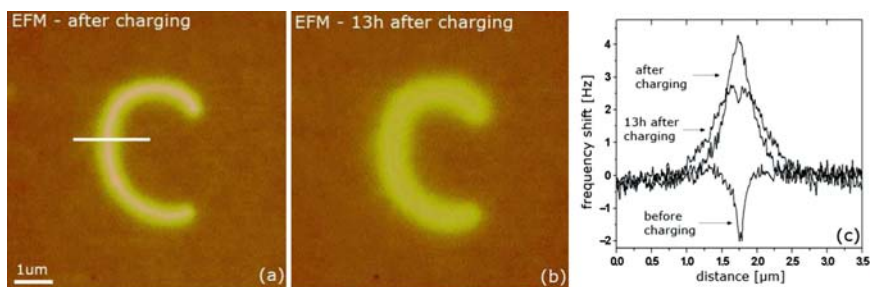


Fig. 4.16. (a) EFM image of a multiwalled carbon nanotube with 30 nm diameter just after charge injection. The horizontal bar across the nanotube indicates the place where the frequency shift cross-section has been taken. (b) EFM scan of the same nanotube 13 h after the charge injection. (c) Corresponding frequency shift cross-sections, together with the data acquired prior to charging. Adapted from [12]

storage in the oxide layer along the CNTs. While demonstrating the charging of a nanotube is relatively easy in the case of MWCNTs exhibiting abrupt discharging processes, it, however, needs to be assessed carefully in the case of SWCNTs, where the oxide charge can predominate over the CNT charge. To show this, comparative charge injection and EFM experiments performed on MWCNTs and SWCNTs need to be discussed. This is done in Fig. 4.17, showing the EFM data of a MWCNT (also used in Fig. 4.14) and a SWCNT (~ 3 nm diameter) acquired in the linear mode using $V_{\text{EFM}} = -3$ V and lift heights $z = 80$ and 60 nm, respectively. The negative EFM frequency shifts observed in Fig. 4.17a, c prior to charging experiments are the capacitive footprints of the nanotubes topography. After charging (with $V_{\text{inj}} = -6$ V for 2 min), the EFM signal of the MWCNT exhibits a pronounced positive frequency shift (*see* Fig. 4.17a), predominating over the capacitive signal and corresponding to a negative stored charge. A first evidence for charge storage in the MWCNT is given by the occurrence of an abrupt discharge while scanning, as discussed previously. Similar experiments have been performed on the SWCNT. Immediately after charging, the nanotubes height rises to 4.8 nm, and falls to 3.5 nm within the first EFM scan indicating that the SWCNT also undergoes a discharge, though not abrupt. The discharge is further evidenced by comparison of the EFM data of the two consecutive scans acquired after charging, showing the resurgence of the SWCNT capacitive frequency dip after the discharge (*see* Fig. 4.17c). These features show that some charge has indeed been injected in the SWCNT, but in contrast with the MWCNT, it only represents here a small fraction of the trapped oxide charge visible after the discharge, as seen from Fig. 4.17c,d. The issue of oxide charging thus needs to be taken in consideration in order to investigate the electrostatics of CNTs from EFM and charge injection experiments.

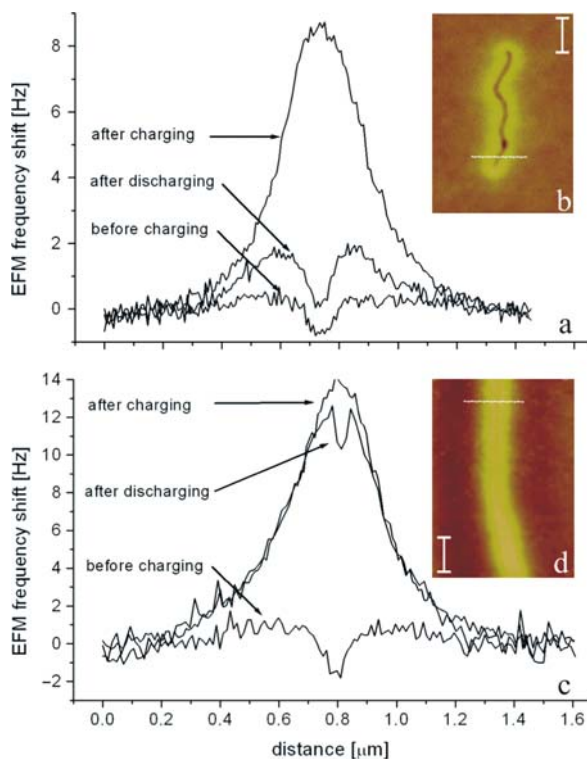


Fig. 4.17. (a) EFM frequency shifts ($V_{\text{EFM}} = -3 \text{ V}$) acquired across a 15 nm multi-walled carbon nanotube (see horizontal line in (b)) before charging, after charging ($V_{\text{inj}} = -6 \text{ V}$ for 2 min), and after the MWCNT discharge. (b) EFM image (10 Hz scale) of the discharged MWCNT. The scale bar is 500 nm. (c) Similar EFM signals acquired on a 3.0 nm SWCNT. (d) EFM image (20 Hz scale) of the discharged SWCNT. The scale bar is 300 nm. Adapted from [13]

4.4.3 Inner-Shell Charging of CNTs

The detailed charge storage mechanisms of SWCNTs and MWCNTs and their relation with classical electrostatics – that is, do CNTs charge as capacitors? – have been established from EFM experiments [30], but require a truly quantitative analysis of the CNT charge states from EFM experiments. In this section, we discuss whether nanotubes behave like cylindrical metal capacitors – and thus only exhibit an outer-shell charging following a classical cylinder-plane capacitance, or whether they exhibit an internal response to the biased tip probe, allowing an inner-shell charging of the nanotubes. In order to find an answer to this question, one needs to keep in mind that on local charging of CNTs, there are three distinct behaviors: either the injected charge is mostly stored *in* the CNT, or mostly emitted by the CNT to the

substrate and trapped in the insulating SiO_2 layer, or finally stored both in the CNT and oxide traps.

The crucial issue to understand the electrostatics of CNTs from charge injection and EFM measurements is to perform a quantitative analysis of injected charge densities from EFM signals. Models derived from plane capacitor geometries have been used in early work on SWCNTs [26] but in order to get truly quantitative charge measurements, the EFM data are analyzed by comparing the CNT capacitive frequency shift Δf_c with the additional frequency shift Δf_λ observed after charging the CNT with a linear charge density λ , and compute the ratio $R = \Delta f_\lambda / \Delta f_c$ [31]. The quantities Δf_c and Δf_λ are in practice measured from the EFM images recorded prior to charging and after charge injection (*see, e.g., Fig. 4.12*). The advantage of this normalization procedure is that the ratio R does not depend on the spring constant of the cantilever, which enables direct comparisons between charging experiments on different CNTs. It gives the possibility to derive the linear charge densities (λ) of CNTs [30]. The relation of R and linear charge densities can be performed without any adjustable parameter from the z -component of electrostatic force gradients at the EFM tip and the ratio R between charge and capacitive frequency shifts.

The linear charge densities corresponding to experimental EFM data are presented in Fig. 4.18. EFM experiments have been performed on 25 nanotubes: 18 MWCNTs (*full symbols*) and 7 SWCNTs (*open symbols*). Linear charge densities have then been normalized by V_{inj} , and finally, plotted as a function of the CNT diameter d_{CNT} . For each CNT, it has been determined whether the charge was detected in the CNT (*squares*), trapped on the SiO_2 surface (*circles*), or in part in the CNT and in oxide traps (*diamonds*). For the purpose of comparison, classical electrostatic predictions for metal cylinders are also shown in Fig. 4.18 (*top curves*). The curve (a) corresponds to a cylinder-plane capacitance model, in a medium with an average dielectric constant accounting for the geometry of the CNT deposited on the oxide layer. Points shown in (b) have been obtained directly from the three-dimensional Poisson calculation with the EFM tip in contact with CNTs as in Fig. 4.11b.

One striking conclusion can be already drawn from Fig. 4.18: the experimental linear charge densities are found to deviate from classical electrostatic (curves a and b) predictions by more than one order of magnitude, already showing that CNTs (either MWCNTs or SWCNTs) *do not* charge like cylindrical capacitors. Within experimental accuracy, the experimental charge densities also exhibit a common variation law as a function of the CNT diameter d_{CNT} , irrespective of the actual charge location, that is mostly injected in the CNTs, in the oxide surface, or in part in CNTs and in oxide traps.

The reason for this effect is evidenced in Fig. 4.19, showing an energy diagram from the nanotube with the EFM biased at V_{inj} in contact with the nanotube, but prior to charge transfer to the nanotube or surface. Because of the finite transverse polarizability of the nanotube, the electric field generated at the EFM tip apex is seen to penetrate the MWCNT shells (Fig. 4.19,

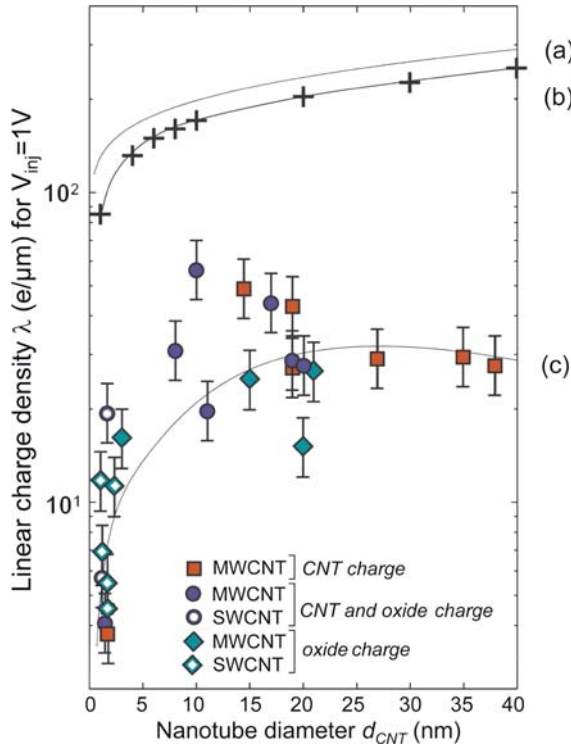


Fig. 4.18. Plot of the CNT linear charge density (in $e/\mu\text{m}$) as a function of the CNT diameter d_{CNT} for an injection voltage $V_{\text{inj}} = 1 \text{ V}$. The *top curves* correspond to densities expected for metal cylinders. The *curve (a)* is derived from the cylinder-plane capacitance. *Points in (b)* are obtained from three-dimensional Poisson calculations with the EFM tip in contact with the CNT as in Fig. 4.1a. The *line in (b)* is a guide to the eye. *Curve (c)* corresponds to linear charge densities expected from the inner-shell charging of nanotubes, as a response to the electric field generated by the EFM tip. Adapted from [30]

left), leading to a voltage drop across the MWCNT prior to charging. The charging of the MWCNT will then correspond to an equilibrium situation that takes place when the electrostatic energy of a nanotube will equal ΔV . Since $\Delta V < V_{\text{inj}}$, this accounts for the fact that the experimental linear charge densities in Fig. 4.18 are found below the capacitive charging expectations. This process also accounts quantitatively for experimental observations (see line c in Fig. 4.18). The MWCNT inner-shell charging is finally explained by the fact that electrons are stored in the nanotube outermost (but possibly subsurface) metallic shell, in agreement with qualitative observations of charge storing in MWNCTs as seen from the abrupt discharging mechanisms shown in Fig. 4.14.

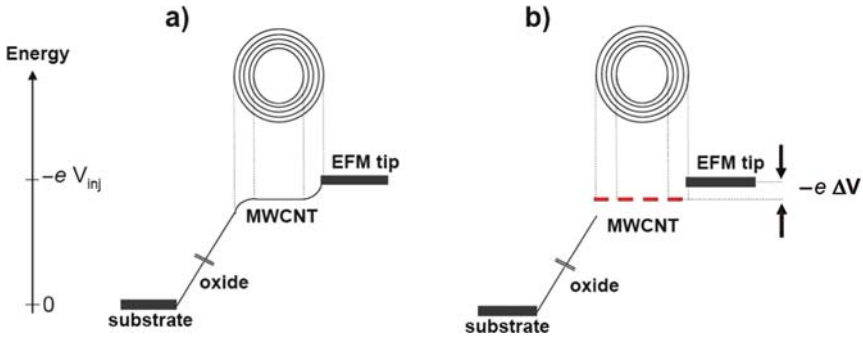


Fig. 4.19. *Left:* Schematics of the potential energy drop across the EFM tip/nanotube/oxide during the charge injection process (no charge transferred to the nanotube). *Right:* average of the potential energy for each shell, showing the voltage drop to be compensated by the nanotube charging is $\Delta V < V_{inj}$.

4.4.4 Electrostatic Interactions in SWCNTs

The above examples show that the EFM technique is useful to study the static distribution of electric charges in carbon nanotubes. It is interesting to point out that the theoretical calculation of charge distribution on SWCNTs had predicted U-like shapes of the charge distribution [32], with a charge accumulation at the tube ends stemming from Coulomb repulsion. These predictions, however, have been proposed for short CNTs (< 100 nm). The EFM technique has been applied to verify if these theoretical findings are suitable, however, for longer CNTs.

Wang et al. [33] proposed study of the charge enhancement localized at the tube ends and simultaneously uniform charge density along the tube taking into account the oxide surface on which nanotube is deposited. They demonstrated that the previously predicted U-shape charge distribution in CNTs is present, but strongly weakened in conditions that are usually accessed by experiments (i.e., for tubes with μm length), in which a $\sim 20\%$ charge enhancement can be expected.

The linear charge distributions observed from EFM charging experiments of SWCNTs deposited on a SiO_2 substrate is shown in Fig. 4.20. In order to evidence the weak charge enhancement at the nanotube cap, a nonlinear color scale has been used in this figure. Note that the charge enhancement is localized only within 200 nm at the nanotube cap, and corresponds here [33] to the expected value for charge enhancement. This weak charge enhancement clearly shows that the charge distribution expected for short nanotubes is strongly modified for longer objects.

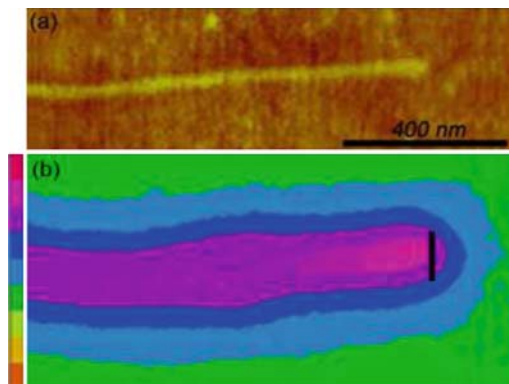


Fig. 4.20. (a) AFM topography image of single-walled nanotube with diameter 1.6 nm and length 2 μm deposited on SiO_2 . (b) EFM scan of the same tube made after charge injection. A nonlinear color scale has been used in order to clearly show the weak enhancement at the tube end. The black line is a “guide to the eye” for the physical end of the tube. Adapted from [33]

4.5 Probing the Band Structure of Nanotubes on Insulators

Various measurements of the electronic properties of carbon nanotubes have been carried out at single molecular level since the discovery of nanotubes. Scanning tunneling spectroscopy has been mostly used for this purpose, providing for instance very valuable information about Van Hove singularities of the nanotube density of states [34]. In these experiments, nanotubes need to be deposited over a metallic substrate to allow electrons to tunnel to/from a scanning tunneling spectroscopy tip. However, because of the nanotube–substrate interaction, many of their properties could not be investigated. EFM techniques give the possibility to circumvent this problem allowing the characterization of the electronic properties of nanotube on insulators. The subject of this section is to give short overview on recent studies carried out *on insulators*, on the assessment of the metallic/semiconducting character and of the local density of states of SWCNTs.

4.5.1 Imaging the Semiconductor/Metal Character of Carbon Nanotubes

The electrostatic response of charged carbon nanotubes probed by EFM technique can be used to investigate the metallic/semiconducting character of the carbon nanotubes. Barboza et al. [35] have recently reported on a semiconductor–metal transition for SWCNTs monitored from EFM experiments.

Charging and EFM experiments used in this work were similar to technique described previously, except that no bias is applied between the tip and the

sample ($V_{\text{EFM}} = 0$) during the second pass of EFM imaging, so that the EFM signal consists solely in the image charge interactions (interaction with the nanotube charge together with their electrostatic image with respect to the metal-plated tip). This corresponds to the $(V_Q)^2$ term in (4.3), leading to an always negative frequency shift (force gradient associated with the attractive image charge interaction).

Semiconducting and metallic tubes used in the EFM experiments were first preselected by Raman spectroscopy studies in order to determine their semiconducting/metallic character and chirality. Then, each tube was subjected to charge injection, and EFM data were recorded as a function of tip injecting bias (V_{inj}), using a minimal and constant ($\sim 0.3 \text{ N/m}$) tip-nanotube force during the injection process.

Barboza et al. measured the linear charge densities as a function of V_{inj} for metallic and semiconducting nanotubes (see Fig. 4.21). For a metallic nanotube, the charge-bias plot (black squares) exhibits a nicely linear symmetric behavior with a minimum bias of approximately $\pm 2 \text{ V}$ necessary for detection of extra charge at the nanotube. The charge-bias plot for semiconducting nanotubes (in red) exhibits an important difference compared with the metallic tube: the plot is asymmetrical with respect to the injection bias (see the deviation at $V_{\text{inj}} = -3 \text{ V}$), and the charge density is also smaller than that for metallic tubes for a given V_{inj} .

In the second part of the charging experiment, the injection bias (V_{inj}) and injection time were kept constant. The linear charge densities were measured as a function of the compressive tip-tube force (per unit length) during the injection (see Fig. 4.21b). For the metallic tube (dark squares), charge densities exhibit a rather weak dependence on the compressive force. Semiconducting nanotubes, however, show striking features on the increase of

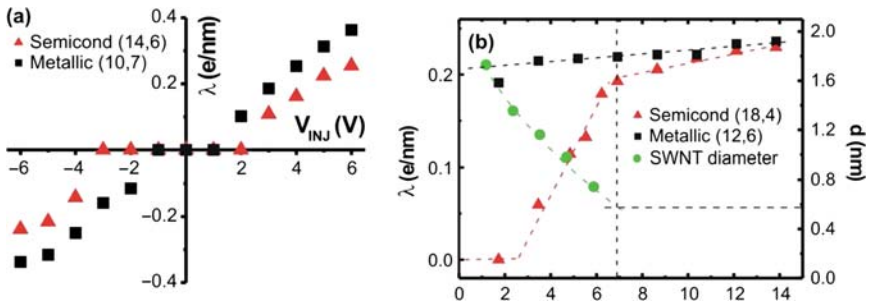


Fig. 4.21. (a) Plot of the charge density (λ) versus injection bias (V_{inj}) for metallic (black squares) and semiconducting tube (red triangles). (b) Plot of the charge density (λ) as a function of the force (per unit of length) applied during nanotube charging for metallic and semiconducting tube, black squares and red triangles, respectively. The evolution of the semiconducting nanotube as the compressive force is applied is represented by green circles. The dashed lines are guides for the eye. Adapted from [35]

the compressive force. For a small EFM tip force during the charge injection, no injected charge is detected whereas for bigger force ($> 3\text{ N/m}$), a steep increase of charge density is detected until it reaches the value measured for metallic tube. This effect was attributed by Barboza et al. to a semiconducting-to-metal transition on an applied external compressive force.

4.5.2 Imaging the Density of States of Carbon Nanotubes

Probing of local density of states in carbon nanotubes is certainly another remarkable capability of the EFM technique. In order to show this, Heo et al. [36] proposed an experiment that is carried out in a similar way as described earlier. However, the slow scan axis is disabled, so the tip always scans the same line across the sample. A nanotube lying on the oxide layer is scanned in two modes: standard tapping mode (topography) and EFM mode. In the latter one, the tip is set at very low tip-sample height (typically a few nanometers). During each second pass (EFM mode), the tip is biased at voltage V_{EFM} and phase shift ($\Delta\varphi$) between the mechanical drive and the cantilever motion is recorded.

The biased EFM tip (V_{EFM}) is used to vary the Fermi level tube (by changing V_{EFM}). The local density of states drastically increases whenever the Fermi level is swept through a van Hove singularity in the density of states. This changes the cantilever-tube coupling, and thus the mechanical oscillation of the tip, which is detected during the experiment and V_{EFM} sweep. The upper panel of Fig. 4.22 shows a color scale plot of the cantilever phase shift $\Delta\varphi$ versus ΔV and tip position Δx along a 0.9 nm diameter SWCNT. ΔV is here proportional to V_{EFM} (see [36] for more details). The dashed line indicates the nanotube position. The lower panel of Fig. 4.22 shows the plot of $\Delta\varphi$ versus ΔV when the tip is directly over the nanotube. The plot follows the trend $\Delta\varphi \sim -\Delta V^2$; however, some kinks are seen for certain voltages (indicated by the *arrows*), which are clearly visible in the when the data background is subtracted (Fig. 4.22b): now the plot follows approximately the trend $-d^2(\Delta\varphi)/\Delta V^2$ and it shows series of peaks as ΔV varies. These peaks occur at the same voltage position of the kinks presented in Fig. 4.22a. According to [36] the peaks in the $\Delta\varphi$ directly reflect the one-dimensional van Hove singularities of the tube. The voltage position of the peaks is nanotube dependent, as expected for tubes with different electronic structure and thus allowing distinguishing between individual metallic and semiconducting tubes.

The possibility to measure the dielectric polarization of individual SWCNTs has also been proposed recently by [37]. According to them, the longitudinal dielectric constants (probed by EFM) can exhibit different responses for metallic and semiconducting tubes, thus allowing the assessment of the electronic character of the nanotubes (see [37] for more details).

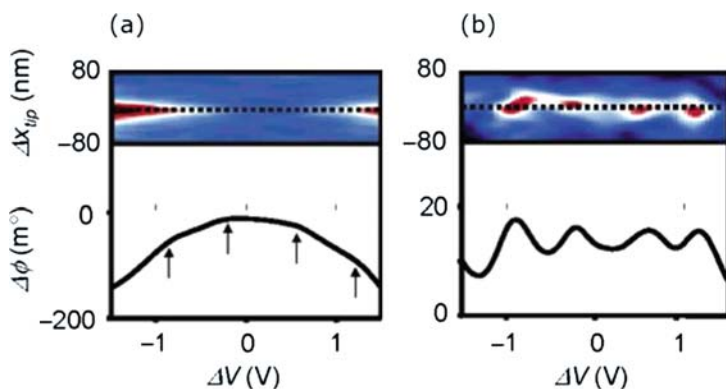


Fig. 4.22. (a) *Upper panel:* color plot of the phase shift ($\Delta\phi$) as a function of tip position (Δx) and voltage ΔV , where ΔV is proportional to V_{EFM} (see Heo et al. (2005) for more details). The dotted line shows the position of the tube. Dark red corresponds to -150 m° whereas dark blue corresponds to 0.0 m° . *Lower panel:* $\Delta\phi$ versus ΔV when the tip is directly above the tube (dotted line in upper panel). Arrows indicate the kinks in the plot. (b) *Upper panel:* color plot of $\Delta\phi$ versus Δx and voltage ΔV when the background is subtracted taken. Dark red corresponds to 20 m° whereas dark blue corresponds to 5 m° . *Lower panel:* $\Delta\phi$ versus ΔV taken along the dotted line exhibiting a series of peaks. The voltages giving the peaks correspond to the voltages yielding the kinks in the data in the lower panel of (a). Adapted from [36]

4.6 KFM Studies of Nanotube Devices

This section describes the use of electrical techniques derived from AFM to probe the electrostatic properties of carbon nanotube devices such as CNT-FETs. To map, for example, voltage drops along nanotubes, KFM-based studies (or surface potential measurements) are here preferred to EFM studies (i.e., charge measurements), which were reviewed in the previous sections. Here we describe the work done using KFM techniques on carbon nanotube devices, starting from the study of charge transfers at the nanotube–contact interfaces, the assessment of diffusive/ballistic character of transport along SWCNTs, and finally, more recent studies performed on CNTFETs under operation.

4.6.1 Charge Transfers at Nanotube–Metal Interfaces

Field effect transistors based on carbon nanotubes with controlled or minimized Schottky barrier heights have been demonstrated by [38] using palladium electrodes. From an electrostatic point of view, the study of the formation of electrical contacts with ohmic or Schottky character between metals and nanotubes can be performed by studying the contact potential

difference (or charge transfer) between metals to nanotubes. This idea corresponds in fact to the first use of the Kelvin probe implementation in an AFM [10], in which the contact potential difference was measured between different metals (gold and palladium), with a sensitivity of a few tens of millivolt.

The idea followed by [39] was based on the fact that CNTFETs operate as p-type devices in air (i.e., Schottky barrier transistors with hole transport) ([2, 3]), even though not intentionally doped, whereas devices probed in ultrahigh vacuum (UHV) behave as n-type devices (i.e., Schottky barrier transistors with electron transport) [40,41]. This change of behavior has been investigated using KFM by dispersing SWCNTs³ on a flame-annealed Au surface (see Fig. 4.23), and monitoring the associated contact potential difference by KFM (contrast in the surface potential image).

The nanotubes on Au in the KFM image of Fig. 4.23b, *taken in air*, appear as bright features (positive surface potential with respect to the surface), indicating a positive charging of the nanotube due to a lower vacuum level. The measured contact potential difference is approximately +60mV. This behavior is, however, opposite (see Fig. 4.23c) after a long stay of the sample in UHV, after which the nanotubes are seen as dark features in the KFM image

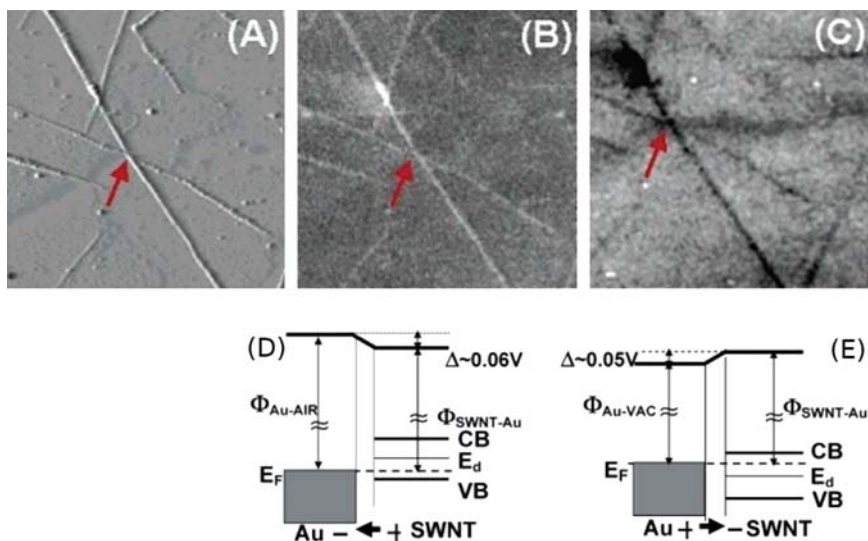


Fig. 4.23. (a) AFM topographic images of nanotubes deposited on a flame-annealed Au (111) surface. (b, c) KFM images obtained in air and after a long period in ultrahigh vacuum (UHV) (Au). The bright contrast indicates a lower vacuum level. (d, e) Energy diagrams showing the energy-level alignment between a SWCNT and Au in air and UHV, respectively. Adapted from [39]

³ The contact potential difference of multiwalled carbon nanotubes have also been investigated by KFM: [18]

(negative nanotube charging) due to a higher vacuum level compared with Au, with here a negative contact potential difference of approximately -50mV . The two diagrams, corresponding to p -type (respectively n -type) carbon nanotubes are shown in Fig. 4.23d,e, and are found consistent with the p -type (respectively n -type) transport behavior of SWCNT FET devices in air or UHV [41].

In order to probe energy level alignments, similar experiments using KFM have been carried out more in the case of ferromagnetic contacts used for spintronic applications. Shiraishi et al. [42] measured the contact potential difference at the Ni-SWCNT and Co-SWCNT interfaces, and found them equal -52 and -27mV , respectively. The corresponding energy-level alignment diagrams (*see* [42]) suggest that holes could be suitable spin carriers for SWCNT spin devices.

4.6.2 Diffusive and Ballistic Transport in Carbon Nanotubes

The transport properties of carbon nanotubes probed by KFM have been discussed in terms of diffusive and ballistic transport by [14] both for MWCNTs and SWCNTs. The authors used an ac-EFM technique, which is equivalent to an open-loop KFM technique (*see* Sect. 4.2 for details), the ac-bias being applied across the nanotube so as to reduce side capacitance effects [43]. Additionally, the ac-EFM signal is here subtracted from the background signal across the electrodes, so as to get rid of the stray capacitive coupling of the tip to the metal electrodes. ac-EFM maps can therefore be viewed as the local measurement of the attenuation of an ac voltage along carbon nanotubes.

ac-EFM measurements are presented in Fig. 4.24 for both an MWCNT and an SWCNT [14]. The ac-EFM signal of the MWCNT is presented in Fig. 4.24a,b, in which the nanotube left electrode is subjected to the ac bias of 150mV . The KFM signal shows here a linear drop, indicating that the nanotube behaves as a diffusive conductor. The ac-EFM signal of a bundle of SWCNTs is then presented in Fig. 4.24c, and shows no measurable intrinsic resistance (plateau in the ac-EFM cross-section of Fig. 4.24d) when the ac-EFM signal is projected across the electrodes, which was – in conjunction with the nanotube transport properties – taken as a signature of the ballistic transport in the bundle of SWCNTs [14].

4.6.3 Kelvin Force Microscopy of CNTFETs

Backgate Operation of CNTFETs

We now turn to the characterization of CNTFETs using KFM techniques. A CNTFET based on a bundle of SWCNTs has been fabricated using Ti electrodes as source and drain contacts, a 300-nm -thick silicon dioxide layer as a gate dielectric and a p -doped silicon wafer as a backgate (*see* Fig. 4.25) [44, 45]. KFM images (here, in vacuum, using an FM-KFM method as described

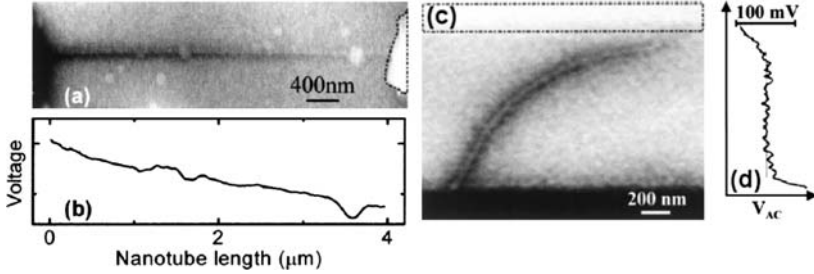


Fig. 4.24. (a) ac-Electrostatic force microscopy (*ac-EFM*) image of a multiwalled carbon nanotube of diameter 9 nm. An ac-bias of 150 mV is applied to the left electrode. (b) ac-EFM signal as a function of the MWCNT length. (c) ac-EFM image of a metallic bundle of SWCNTs with diameter 2.5 nm. An ac bias of 100 mV is applied to the lower electrode. (d) The ac-EFM signal exhibits a plateau along the length of the SWCNT bundle. Adapted from [14]

in Sect. 4.4) have been obtained for various source-drain voltages and various gate potentials. Examples are provided in Figs. 4.25a, b for $V_{DS} = 1$ and 3 V respectively, both at $V_{GS} = 3$ V. KFM profiles are then plotted for various values of gate voltage ranging from $V_{GS} = -2$ to 2 V at $V_{DS} = 1$ V (see Fig. 4.25b) and $V_{DS} = 3$ V (see Fig. 4.25d). In these profiles, the potential drops along the nanotube are strongly influenced by applying a gate potential. The difference of potential at the drain edge is also strongly modulated, in agreement with Schottky barrier FET models for carbon nanotube transistors [46, 47].

KFM Determination of the lever arm of a CNTFET

The crucial issue to understand KFM experiments on transistors is to derive quantitative information from KFM images, even though the role of side capacitances has been pointed out earlier [15, 43]. We here present a KFM study performed on Pd-contacted CNTFETs fabricated on an ~ 320 nm-thick SiO_2 layer, which addresses this issue [48]. It is shown in particular that the lever arm between the backgate bias and nanotube electrostatic potential shifts – of strong interest to characterize the CNTFET operation – can be assessed from such experiments.

The electrostatics of the CNTFET as seen from KFM experiments can be written using the side-capacitance approach introduced by [15]:

$$V_{\text{KFM}} = \alpha_{\text{NT}}V_{\text{NT}} + \alpha_{\text{G}}V_{\text{G}} + \alpha_{\text{D}}V_{\text{D}} + \alpha_{\text{S}}V_{\text{S}} \tag{4.7}$$

Here the KFM dc bias V_{KFM} is written as a weighted average of the nanotube electrostatic potential V_{NT} , of the backgate electrostatic potential V_{G} , and of the source and drain electrodes at voltages V_{D} and V_{S} , respectively. In (4.7), the coefficients α_i refer to the weighted average between capacitance

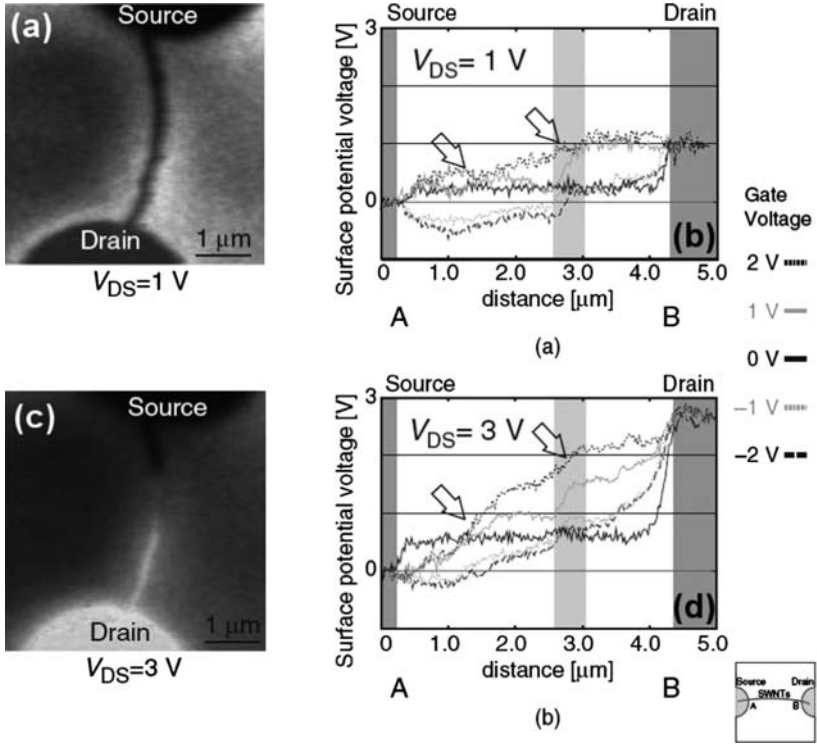


Fig. 4.25. (a) KFM image of a bundle of SWCNTs connected with Ti electrodes at $V_{DS} = 1\text{ V}$ and $V_{GS} = 3\text{ V}$. (b) Surface potential profiles at $V_{DS} = 1\text{ V}$ for V_{GS} from -2 to 2 V . (c) KFM image of the same bundle of SWCNTs as at (a) $V_{DS} = 3\text{ V}$ and $V_{GS} = 3\text{ V}$. (d) Surface potential profiles at $V_{DS} = 3\text{ V}$ for V_{GS} from -2 V to 2 V . In (b) and (d), the dark shaded areas correspond to the source and drain electrodes. Adapted from [44]

derivatives, as described in details in Sect. 4.4, which are here determined phenomenologically. Using $V_{NT} = \beta V_G$ (β being the electrostatic lever arm for the CNTFET), and $\alpha_{NT} + \alpha_G + \alpha_D + \alpha_S = 1$ for the sum of the weights in (4.7), this description requires the experimental determination of four independent coefficients.

This can be done by using the recorded KFM voltage drop across the device when the nanotube is subjected to a source-drain bias $V_{DS} = -3\text{ V}$ (the source is grounded here). From the KFM bias recorded at the source, one obtains $\alpha_D \sim 1\%$ and from the voltage drop across the device (here $600 \pm 50\text{ mV}$), one gets $\alpha_D + \alpha_{NT} = 20 \pm 2\%$, that is $\alpha_{NT} = 19 \pm 2\%$. To determine the two remaining relations and find all the α_i and β , one then uses the KFM maps of the operating device, recorded at different backgate bias V_G (see Fig. 4.26). Such maps give two information: (1) the slope of the variation of V_{KFM} as a function of V_G recorded *off the nanotube*, and which corresponds to $(\alpha_{NT} + \alpha_G)$; and (2) the slope of the variation of V_{KFM} as

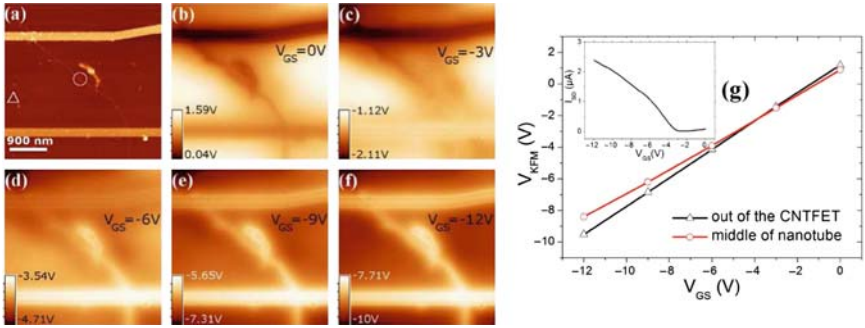


Fig. 4.26. (a) AFM image of a single-walled nanotube field effect transistor. Scale bar: 900 nm. (b–f) Associated KFM images at $V_{DS} = -3$ V, (b) $V_{GS} = 0$ V, (c) $V_{GS} = -3$ V, (d) $V_{GS} = -6$ V, (e) $V_{GS} = -9$ V, and (f) $V_{GS} = -12$ V. (g) Recorded KFM voltage on the nanotube (circles) and off the nanotube (triangles) as a function of the device gate bias V_G , in the range of operation of the CNTFET (see inset for the CNTFET transfer characteristics). Adapted from [48]

a function of V_G recorded with the tip above the nanotube, which provides a direct measurement of $(\alpha_{NT}\beta + \alpha_G)$. Data extracted from Fig. 4.26 give $\alpha_{NT} + \alpha_G = 0.89 \pm 0.02$, and $\alpha_{NT}\beta + \alpha_G = 0.75 \pm 0.02$. These values lead to the following determination of the α_i and β coefficients: $\alpha_{NT} = 19 \pm 2\%$, $\alpha_G = 70 \pm 2\%$, $\alpha_S = 10 \pm 2\%$ and $\beta = 0.26 \pm 0.15$. These figures show that with fairly large tip substrate distances (here $z \sim 80$ nm), the weight α_{NT} associated with the nanotube detection is quite weak. This is in fact strongly enhanced when operating in UHV, because the tip-substrate distance in noncontact AFM measurements is typically less than 20 nm, and thus will enhance the capacitive weight α_{NT} compared with α_G , α_S , and α_D . Also, the electrostatic lever arm of the nanotube has been here determined quantitatively, with a $\beta = 0.26 \pm 0.15$ value close to experimental values already reported in literature for CNTFET devices with similar geometries [49,50], and estimated from, for example, Coulomb blockade effects at low temperature.

Hysteretic Behavior of CNTFETs and Surface Charges

A similar study as in Sect. 4.6.3 has been performed with a 1.1 nm diameter SWCNT field effect transistor [51]. The authors were interested in the origins of the hysteresis of a CNTFET, associated with out-of-equilibrium charge storage in the CNTFET environment, and for which several possibilities have been proposed in literature: charging at the nanotube–ambient air interface [52], at the nanotube–SiO₂ interface [49,53], or at the Si–SiO₂ interface [49,54].

Here, Umesaka et al. first applied a bias stress to the backgate of -5 V (and $+5$ V), and later recorded the KFM images of the device, shown respectively in Fig. 4.27a,b. KFM profiles have then been taken across the nanotube for

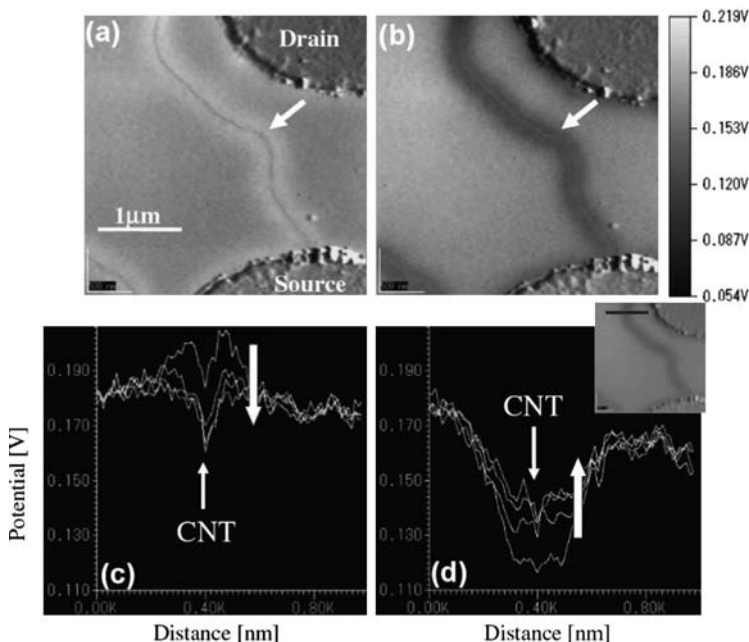


Fig. 4.27. *Top:* surface potential images of a CNTFET at $V_{GS} = 0$ V and $V_{DS} = -0.3$ V (a) just after applying -5 V gate bias stress and (b) just after applying $+5$ V gate bias stress. *Bottom:* Line profiles of the surface potential of the CNTFET at $V_{GS} = 0$ V taken 0, 15, 30, and 45 min after applying the gate bias stress. The graphs in (c) and (d) correspond respectively to a -5 V gate bias stress and a $+5$ V gate bias stress. The bar in the inset indicates the position of the line profiles. Adapted from [51]

different times after the gate stress. The results are shown in Fig. 4.27c and d, respectively, for the -5 V stress and the $+5$ V stress. The nanotube environment shows a positive (respectively negative) surface potential after the -5 V (respectively $+5$ V) gate bias stress, both charging effects decaying with time. This indicates a capacitive response of the nanotube channel to the backgate with opposite charges trapped into the oxide layer surrounded the nanotube, in agreement with the hysteresis observed in the CNTFET operation [51].

4.7 Conclusion

In this chapter, we have reviewed recent experimental studies carried out on carbon nanotubes using electrical techniques derived from AFM. It has been shown, on the one hand, that the dielectric and charging properties of individual carbon nanotubes can be measured experimentally, using EFM, and revealing peculiar electrostatic properties associated with their

one-dimensional, semiconducting, or multishell properties. On the other hand, it has been demonstrated that KFM and related techniques are invaluable so as to bring insight into the operation of devices based on carbon nanotubes, including the formation of electrical contacts and of charge transfer at nanotube–metal interfaces, the assessment of diffusive or ballistic character of transport along SWCNTs, and quantitative measurements on CNTFETs, such as the electrostatic lever arm governing the transistor operation. Such studies, together with the ongoing and future developments of electronic characterization tools using nanoscale force measurements, especially in UHV environments and/or at low temperature, should bring a valuable insight for such fields of research combining nanoscience and nanotechnology.

Acknowledgement

The authors acknowledge fruitful discussions with L. Adamowicz, H. Diesinger, V. Derycke, D. Deresmes, A. Mayer, Z. Wang, and M. Devel.

References

1. S. Iijima, *Nature* **354**(6348) 56 (1991)
2. S.J. Tans, A.R.M. Verschueren, C. Dekker, *Nature* **393** 49 (1998)
3. R. Martel, T. Schmidt, H.R. Shea, T. Hertel, P. Avouris, *Appl. Phys. Lett.* **73** 2447 (1998)
4. G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Phys. Rev. Lett.* **49** 57 (1982)
5. G. Binnig, C.F. Quate, C. Gerber, *Phys. Rev. Lett.* **56** 930 (1986)
6. Y. Martin, D.W. Abraham, H.K. Wickramasinghe, *Appl. Phys. Lett.* **52** 1103 (1988)
7. J.E. Stern, B.D. Terris, H.J. Mamin, D. Rugar, *Appl. Phys. Lett.* **53** 2717 (1988)
8. B.D. Terris, J.E. Stern, D. Rugar, H.J. Mamin, *Phys. Rev. Lett.* **63** 2669 (1989)
9. C. Schönenberger, S.F. Alvarado, *Phys. Rev. Lett.* **65** 3162 (1990)
10. M.N. Nonnenmacher, M.P. O’Boyle, H.K. Wickramasinghe, *Appl. Phys. Lett.* **58** 2921 (1991)
11. T.D. Krauss, L.E. Brus, *Phys. Rev. Lett.* **83** 4840 (1999)
12. M. Zdrojek, T. Mélin, H. Diesinger, D. Stiévenard, W. Gebicki, L. Adamowicz, *J. Appl. Phys.* **100** **114326** (2006)
13. T. Mélin, H. Diesinger, D. Deresmes, D. Stiévenard, *Phys. Rev. Lett.* **92** 166101 (2004)
14. A. Bachtold, M.S. Fuhrer, S. Plyasunov, M. Forero, E.H. Anderson, A. Zettl, L. P. McEuen, *Phys. Rev. Lett.* **84** 26 (2000)
15. H.O. Jacobs, P. Leuchtman, O.J. Homan, A. Stemmer, *J. Appl. Phys.* **84** 1168 (1998)
16. S.C. Chin, Y.C. Chang, C.C. Hsu, W.H. Lin, C.I. Wu, C.S. Chang, T.T. Tsong, W.Y. Woon, L.T. Lin, H.J. Tao, *Nanotechnology* **19** 325703 (2008)
17. S.B. Arnason, A.G. Rinzler, Q. Hudspeth, A.F. Hebard, *Appl. Phys. Lett.* **75** 2842 (1999)

18. C. Maeda, S. Kishimoto, T. Mizutani, T. Sugai, H. Shinohara, *Jpn. J. Appl. Phys.* **42** 2449 (2003)
19. M. Bockrath, N. Markovic, A. Shepard, M. Tinkham, L. Gurevich, L.P. Kouwenhoven, M.W. Wu, L.L. Sohn, *Nano Lett.* **2**(3) 187 (2002)
20. C. Staii, A.T. Johnson, J. Nicholas, *Nano Lett.* **4**(5) 859 (2004)
21. T.S. Jespersen, J. Nygård, *Nano Lett.* **5**(9) 1838 (2005)
22. N. Geblinger, A. Ismach, E. Joselevich, *Nat. Nanotechnol.* **3** 195 (2008)
23. T.S. Jespersen, J. Nygård, *Appl. Phys. Lett.* **90**, 183108 (2007)
24. T. Heim, T. Mélin, D. Deresmes, D. Vuillaume, *Appl. Phys. Lett.* **85** 2637 (2004)
25. M. Zdrojek, T. Mélin, B. Jouault, M. Wozniak, A. Huczko, W. Gebicki, D. Stiévenard L. Adamowicz, *Appl. Phys. Lett.* **86** 213114 (2005)
26. M. Paillet, P. Poncharal, A. Zahab, *Phys. Rev. Lett.* **94** 186801 (2005)
27. M. Zdrojek, T. Mélin, H. Diesinger, W. Gebicki, D. Stiévenard, L. Adamowicz, *Phys. Rev. Lett.* **96** 039703 (2006)
28. M. Paillet, P. Poncharal, A. Zahab, *Phys. Rev. Lett.* **96** 039704 (2006)
29. J. Lambert, G. Loubens, C. Guthmann, M. Saint-Jean, T. Mélin, *Phys. Rev. B.* **71** 155418 (2005)
30. M. Zdrojek, T. Heim, D. Brunel, A. Mayer, T. Mélin, *Phys. Rev. B* **77** 033404 (2008)
31. T. Mélin, H. Diesinger, D. Deresmes, D. Stiévenard, *Phys. Rev. B* **69** 035321 (2004)
32. P. Keblinski, S.K. Nayak, P. Zapol, P.M. Ajayan, *Phys. Rev. Lett.* **89** 255503 (2002)
33. Z. Wang, M. Zdrojek, T. Mélin, M. Devel, *Phys. Rev. B* **78** 085425 (2008)
34. J.W.G. Wildoer, L.C. Venema, A.G. Rinzler, R.E. Smalley, C. Dekker, *Nature* **391** 59 (1998)
35. A.P.M. Barboza, A.P. Gomez, B.S. Archanjo, P.T. Araujo, A. Jorio, A.S. Ferlauto, M.S.C. Mazzoni, H. Chacham, B.R.A. Neves, *Phys. Rev. Lett.* **100** 256804 (2008)
36. J. Heo, M. Bockrath, *Nano Lett.* **5** 853 (2005)
37. W. Lu, Y. Xiong, A. Hassaniien, W. Zhao, M. Zheng, L. Chen, *Nano Lett.* **9**, 1664 (2009)
38. A. Javey, J. Guo, Q. Wang, M. Lundstrom, H. Dai, *Nature* **424** 654 (2003)
39. X. Cui, M. Freitag, R. Martel, L. Brus, P.h. Avouris, *Nano Lett.* **3** 783 (2003)
40. V. Derycke, R. Martel, J. Appenzeller, P. Avouris, *Nano Lett.* **1** 453 (2001)
41. V. Derycke, R. Martel, J. Appenzeller, P. Avouris, *Appl. Phys. Lett.* **80** 2773 (2002)
42. M. Shiraishi, K. Takebe, K. Matsuoaka, K. Saito, N. Toda, *J. Appl. Phys.* **101** 014311 (2007)
43. S. Kalinin, D.B. Bonnell, M. Freitag, A.T. Johnson, *Appl. Phys. Lett.* **81** 754 (2002)
44. Y. Miyato, K. Kobayashi, K. Matsushige, H. Yamada, *Jpn. J. Appl. Phys.* **44** 1633 (2005)
45. Y. Miyato, K. Kobayashi, K. Matsushige, H. Yamada, *Nanotechnology* **18** 084008 (2007)
46. J. Appenzeller, J. Knoch, V. Derycke, R. Martel, S. Wind, P. Avouris, *Phys. Rev. Lett.* **89** 126801 (2002)
47. S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, P.h. Avouris, *Phys. Rev. Lett.* **89** 106801 (2002)

48. D. Brunel, D. Deresmes, T. Mélin, *Appl. Phys. Lett.* **94** 223508 (2009)
49. M. Radosavljevic, M. Freitag, K.V. Thadani, A.T. Johnson, *Nano Lett.* **2** 761 (2002)
50. B. Babic, M. Iqbal, C. Schönenberger, *Nanotechnology* **14** 327 (2003)
51. T. Umesaka, H. Ohnaka, Y. Ohno, S. Kishimoto, K. Maezawa, T. Mizutani, *Jpn. J. Appl. Phys.* **46** 2496 (2007)
52. W. Kim, A. Javey, O. Vermes, Q. Wang, Y. Li, H. Dai, *Nano Lett.* **3** 193 (2003)
53. M.S. Fuhrer, B.M. Kim, T. Durkop, T. Brintlinger, *Nano Lett.* **2** 755 (2002)
54. A. Stirling, A. Pasquarello, J.C. Charlier, R. Car, *Phys. Rev. Lett.* **85** 2773 (2000)

Carbon Nanotube Atomic Force Microscopy with Applications to Biology and Electronics

Edward D. de Asis, Jr., You Li, Alex J. Austin, Joseph Leung,
and Cattien V. Nguyen

Summary. Carbon nanotube (CNT) possesses many desirable properties – electrical conductivity, high mechanical strength combined with ductility, chemical inertness, high aspect ratio, and resistance to wear. These same attributes can extend the capabilities of Atomic Force Microscopy (AFM) beyond the limits imposed by conventional silicon and silicon nitride AFM tips. In this chapter, we will give a brief introduction to the chemical, mechanical, and electrical properties of CNT. Following this will be techniques for fabrication of CNT tips and then a detailed discussion of methods for characterizing CNT AFM tip performance with respect to resolution, tip conductivity, three-dimensional imaging, and long-term reliability, including recently reported performance metrics for CNT tips used in Scanning Probe Microscopy (SPM) imaging studies. A survey of the application of CNT probes to a diverse array of AFM imaging modes such as Chemical Force Microscopy, Electric Force Microscopy, Friction Force Microscopy, Scanning Tunneling Microscopy, Magnetic Force Microscopy, and Scanning Near-Field Optical Microscopy highlights the advantages and performance improvements of using CNT AFM probes over conventional silicon probes. Finally, a presentation of the biological applications of CNT AFM probes illustrates the CNT probe’s present use in aggressive imaging of soft, fragile, feature-rich samples and potential use in targeted nanoscale drug-delivery.

Key words: Carbon Nanotube Scanning Probe Microscopy; Fabrication and Characterization of Carbon Nanotube Probe Tip; Application of Carbon Nanotube Probe to Various Force Microscopy Modes; Biological Imaging; Drug Delivery; Optical Near-Field Microscopy.

5.1 Carbon Nanotube Introduction

Carbon nanotube (CNT) is a cylinder of carbon atoms each making covalent bonds with three nearest neighbors (Fig. 5.1a, b). CNT is currently one of the most researched materials in the world because CNT possesses a myriad of highly desirable electrical, mechanical, and thermal properties that can be exploited in a vast array of diverse applications such as electrical devices, composite materials, thermal dissipation devices and, of course, the subject of this chapter – atomic force microscopy (AFM). To understand the potential role

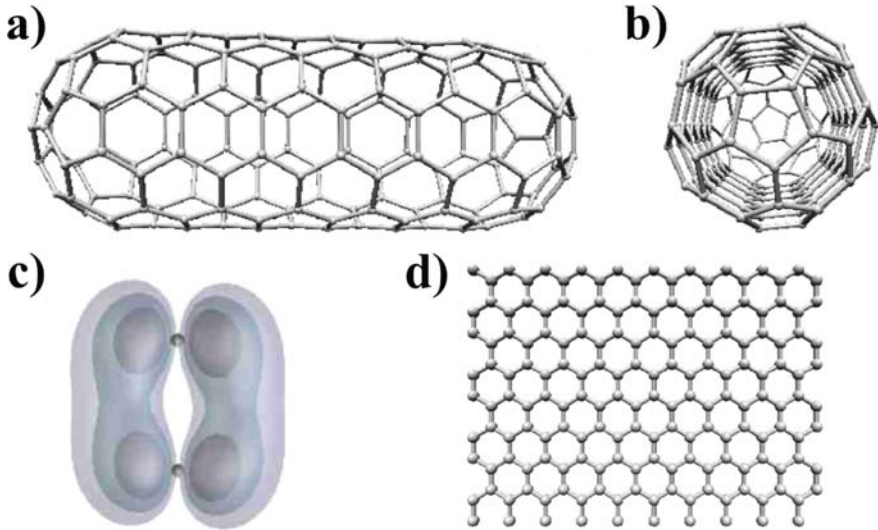


Fig. 5.1. (a) CNT side view, (b) CNT end view, (c) schematic of pi-states in a carbon dimer in grapheme [1], (d) section of graphene sheet

CNT may play in future AFM applications, one must first understand CNT, its properties, and the physics and chemistry underlying those properties. To this end, this section will introduce the reader to CNT and will show that the electrical and mechanical properties can be understood at the basic level by examining the bonds undertaken by the carbon atoms in CNT. As CNT is basically a sheet of graphene rolled up into a cylinder, we will first delve into the properties of an infinite graphene sheet. Then, we will discuss the changes that occur when the graphene sheet is rolled into a CNT. Finally, we will lightly touch upon the effect of defects on CNT properties.

Graphene is a two-dimensional plane of carbon atoms making covalent bonds with three nearest neighbors (Fig. 5.1c). For simplicity in this discussion, we will assume that this plane is free of defects and infinite in extent in order to avoid the thermodynamic and solid-state complications imposed by unterminated dangling carbon bonds in a finite-extent graphene plane. Each pair of carbon atoms is sp^2 hybridized with three σ -bonds radiating out at 120° bond angles from each carbon atom forming covalent bonds with three nearest neighbors (Fig. 5.1d). The mechanical properties of the graphene sheet can be understood within the context of the three in-plane σ -bonds in each carbon atom. With no external forces applied on the sheet, the bond length of the σ -bond between each pair of carbon atoms remains in equilibrium. As tensile or compressive forces are applied perpendicular to the plane normal, the σ -bond length increases or decreases from equilibrium and, assuming that the resulting strain remains in the elastic deformation regime, the bond length or atomic spacing between each pair of nearest neighbor carbon atoms returns to equilibrium when the force is removed.

Rolling up the graphene sheet into a cylinder (i.e., a single-walled carbon nanotube or SWNT) changes the mechanical properties of the CNT slightly changes can drastically modify the electrical conductance properties of the CNT. At the most fundamental level, the unique mechanical as well as electrical properties of CNTs derive from the chemical bonds between the sp^2 hybridized carbon atoms of the graphitic material. The surface curvature and the arrangement of the carbon atoms in CNT may cause the graphitic carbon-carbon bonds to be non-planar. This results in a less efficient overlapping of the pi-orbital between carbon-carbon atoms and hence the carbon atoms may not be perfectly sp^2 hybridized; furthermore, the three sp^2 σ -bonds which radiate out from the carbon atoms are slightly deformed from their preferred trigonal planar orientation. Consequently, these σ -bonds are strained and weakened leading to a slight decrease in CNT Young's modulus compared to the Young's modulus of a graphene sheet [1]. The Young's modulus parallel to the plane of a graphite sheet is about 1.3 Tera pascal and, as expected, the Young's modulus of SWNT is comparable at about 1 Tera pascal, still making CNT one of the strongest materials. Nevertheless, the bonding of the carbon atoms in CNT maintains the same sp^2 -like hybridization characteristics, that is, the carbon-carbon bond as well as the bonding angle can be mechanically distorted in a reversible manner. Figure 5.2 schematically demonstrates the bending of sp^2 hybridized graphitic carbons when an external force is applied. The bonding angle may be twisted out of plane, with the central C atom becoming more sp^3 -like and less sp^2 -like

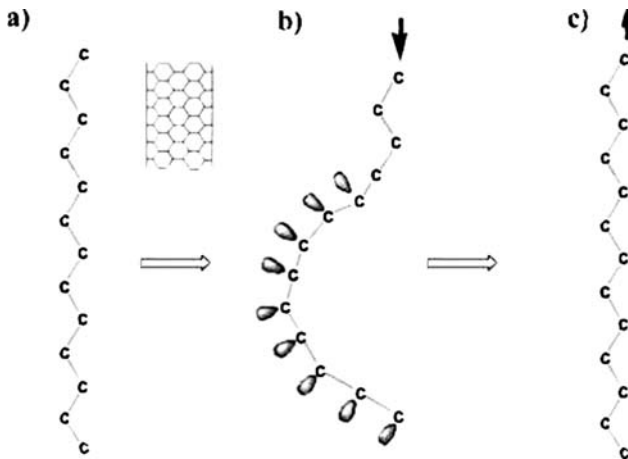


Fig. 5.2. The elastic buckling of the CNT is derived from the chemical bonding energy from the sp^2 carbon atoms. The schematic in (a) is the double bond characteristic of the graphitic carbons along the CNT; (b) external force, as presented by the black arrow, causes bending of the CNT and thus disruption of the pi-bonding of the carbon atoms. The carbon-carbon along the bend are more characteristic of single bond and have more sp^3 -like hybridization; and (c) the carbon atoms revert to their original configuration with graphitic sp^2 bonding when the external force is removed, represented by the arrow

and thus losing most of the bonding contribution by the π -orbital, that is, transforming from double-bond-like to more of a single-bond-like characteristic. When the strain is removed, the bonding of the C–C atoms would revert to its thermodynamically more stable sp^2 configuration and thus retain the original nanotube structure (Fig. 5.2). The bonding energy of a sp^2 carbon–carbon bond is 612 kJ mole^{-1} versus 345 kJ mole^{-1} for sp^3 carbon–carbon bond. The general chemical nature of the sp^2 carbons and the C–C bonds is the basis of the elastic buckling properties of carbon nanotubes. The elastic deformation renders the CNT structurally robust, that is, the CNT will not permanently change its structure when it experiences external forces. This highly desirable mechanical property, coupled with its nanoscale diameter and high aspect ratio, renders CNTs ideal for application as the tip of scanning probes. Therefore, it is worthwhile to explore some of the elastic responses of the CNT in more detail. For the purposes of this AFM applications chapter, the CNT tip is likely to encounter a force that causes it to bend while scanning. The bending force constant of a nanotube is given by

$$k_B = \frac{3Y\pi(r_o^4 - r_i^4)}{4L^3} \quad (5.1)$$

where Y is Young's modulus, r_o and r_i are the outer and inner diameters of the nanotubes, respectively, and L is the length of the CNT [2, 3]. Elastic deformation also invokes a compression response. The compression force constant is given by

$$k_c = \frac{Y\pi(r_o^2 - r_i^2)}{L} \quad (5.2)$$

It can be clearly seen when compared the two equations that the bending response is the larger component of the elastic deformation. Also the length of the CNT, with the length scale typically an order of magnitude larger than the radius, will dramatically influence the bending response of the CNT tip. For multi-walled nanotube (MWNT) which is simply a cylindrical structure of concentric SWNT of increasing radius separated by an interwall spacing of 0.34 nm as shown in Fig. 5.3, the mechanical properties perpendicular to the tube axis depend on van der Waal's forces between adjacent cylinders. Thus, perpendicular to the tube axis, the tensile and compressive strengths are reduced compared to the tensile and compressive strengths parallel to the tube axis.

In the graphene sheet, the remaining p_z orbital in each carbon has two lobes normal to the graphene plane. Overlap of the p_z orbitals of two nearest neighbor carbon atoms leads to the formation of a π -complex which possesses two states – a π state or highest occupied molecular orbital (HOMO) filled with two electrons and a π^* state or lowest unoccupied molecular orbital (LUMO) which is completely empty. When a third carbon atom is considered, the π state and π^* state each split due to the Pauli Exclusion Principle. With the addition of each successive carbon atom in the graphene plane, the π state and π^* state split yet again until, for the infinite graphene plane, there

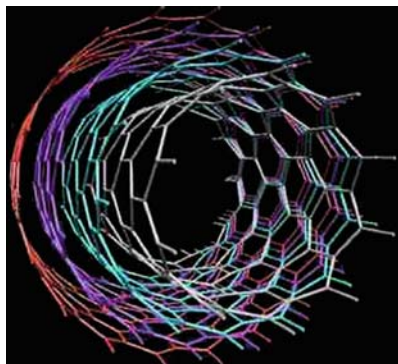


Fig. 5.3. Illustration of MWNT structure with an open end

exists continuous bands of filled π states and unfilled π^* states separated by a Fermi level [1]. Thus, the graphene sheet consists of a HOMO band filled with carriers (electrons) and a LUMO band that is empty. The π band and the π^* band are in fact responsible for the electronic properties of the graphene sheet. Furthermore, when rolling up the graphene sheet into the cylinder, the wrapping angle, which describes the extent of the helicity of a row of hexagons within the CNT, determines whether the CNT is metallic, semimetal, or semiconducting. The theory behind this effect involves solid-state theory which is beyond the scope of this text; thus, we will for simplicity consider only metallic CNTs. For metallic SWNT, once again, rolling up the graphene sheet into a cylinder results in a slight misalignment of the p_z orbitals; however, the mechanism of electrical conduction remains the same as for graphene. Thus, for metallic CNT in which the HOMO band and the LUMO band are not separated by a band gap, the free electrons in the HOMO band act like free electrons in a generic metal.

The above discussion applies to defect-free CNT. Defects which are either pentagons or heptagons of carbon atoms replacing perfect hexagons in the cylinder lead to changes in the CNT structural, mechanical, and electrical properties. Namely, the existence of defects in the CNT leads to indentations in the CNT sidewalls as well as kinks in the CNT; thus, defects prevent CNT from assuming an upright vertically aligned cylindrical shape. Mechanically, defects lead to areas in which the σ -bonds are severely distorted and strained weakening the CNT; thus, defects lead to a decrease in Young's modulus. Electrically, defects in CNT lead to localized changes in the wrapping angle drastically modifying the local electronic structure. Thus, the effect of defects on the electrical properties of CNT can be complicated and provides, in the worst case, an unpredictability in both the electrical and mechanical properties of the CNT AFM tip, and, in the best case, a potential avenue for tailoring CNT electrical conduction properties.

5.2 Carbon Nanotube Synthesis

Four CNT growth processes are currently in wide use: (1) arc discharge, (2) laser ablation, (3) catalytic decomposition, and (4) chemical vapor deposition (Fig. 5.4) [1]. In the arc discharge process, the reactor consists of two closely spaced graphite electrodes inside an airtight chamber. Inert gas (typical pressure is hundreds of mbar) is flowed through the spacing as a DC voltage is applied across the electrodes. The electric field between the plates is sufficient to induce dielectric breakdown causing a current on the order of 100 A to flow raising the temperature to 3,000–6,000°C. Carbon atoms on the graphitic anode are vaporized and deposited on the chamber walls and anode as MWNT and other solid carbon byproducts. Arc discharge SWNT synthesis requires the presence of a metal catalyst (cobalt, iron, nickel) mixed into the graphite anode.

In laser ablation, a tube furnace containing a target consisting of pure graphite for MWNT synthesis or a mixture of graphite and metal catalyst such as Ni–Y or Ni–Co for SWNT synthesis is held at 1,200°C while argon is flowed at sub-atmospheric pressure. A laser ablates the target forming CNT which is collected in soot downstream. The laser may be pulsed or continuous. A pulsed laser requires higher power. The main disadvantages in arc discharge and laser ablation is the production of a large amount of unwanted solid carbon precursors such as amorphous carbons and other graphitic analogues and the high temperatures required leading to the development of catalytic decomposition and chemical vapor deposition processes [1].

Catalytic decomposition and chemical vapor deposition (CVD) refers to CNT synthesis processes for CNT growth in a reaction furnace in which flowing hydrogen gas reduces metal catalyst and a flowing hydrocarbon feedstock gas provides a source of carbon for catalyst-assisted CNT growth. The distinct difference between these two growth processes is in the nature of the metal catalyst employed. The catalytic decomposition refers to the use of an

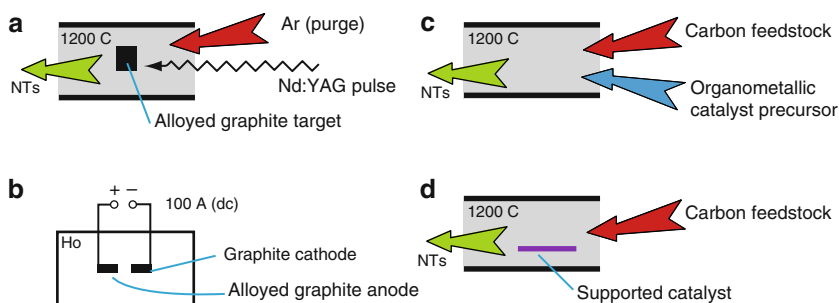


Fig. 5.4. Schematic illustration of various CNT growth methods: (a) laser ablation, (b) arc discharge, (c) catalytic decomposition, and (d) chemical vapor deposition processes for CNT synthesis [1]

organometallic precursor gas flowed into the chamber as the source of the metal catalyst, whereas CVD refers to the synthesis process in which CNT growth occurs on the surface of a supported metal catalyst target. Both of these growth processes use hydrogen for the reduction of the metal catalysts along with hydrocarbon feedstock [1].

5.3 Fabrication of Carbon Nanotube Atomic Force Microscopy Probes

Fabrication of SWNT and MWNT Atomic Force Microscopy (AFM) probes can be grouped into two basic methods: (1) post-growth attachment of a single CNT or a bundle of CNTs on the AFM probe [4–37] and (2) direct growth of the nanotube on the AFM probe [33,38–43]. The current state-of-the-art post-growth attachment involves CNT growth via one of the four primary methods discussed in the preceding CNT synthesis section – arc discharge, laser ablation, catalyzed decomposition, and chemical vapor deposition (CVD) followed by one of three primary methods of attachment: (1) gluing, (2) attachment in an scanning electron microscope, and (3) the pick-up method.

5.3.1 Fabrication of Carbon Nanotube Atomic Force Microscopy Probes by Gluing

Gluing the CNT involves using an adhesive such as acrylic adhesive or conductive carbon tape, to attach CNT to a conventional AFM tip from a CNT source under optical microscopy control [4, 6, 7, 37]. SWNT tips fabricated in this fashion incorporate SWNT ropes and have been used to image DNA [6] while AFM tips with a single MWNT or bundles of MWNT have been used to image gold clusters under water [37], obtain electric force microscopy images of tunnel junctions [7], pattern oxide nanostructures on silicon [4], and compare long-range forces of CNT tips and conventional tips [2].

5.3.2 Mechanical Attachment in Scanning Electron Microscopy

In the second most common method for post-growth attachment, a single CNT is attached to an AFM tip under SEM control (Fig. 5.5) [21]. The fabrication process flow involves three major steps: (1) preparation of the CNT source cartridge via AC electrophoresis (Fig. 5.5a), (2) alignment of and contacting the CNT to the AFM tip inside of an SEM (Fig. 5.5c, d), and (3) welding the CNT to the AFM tip. In step 1, two knife edges are placed on a glass plate separated by a 0.5-mm gap, a drop of CNT–alcohol solution is dispersed in between the gap and an AC voltage is applied between the knife

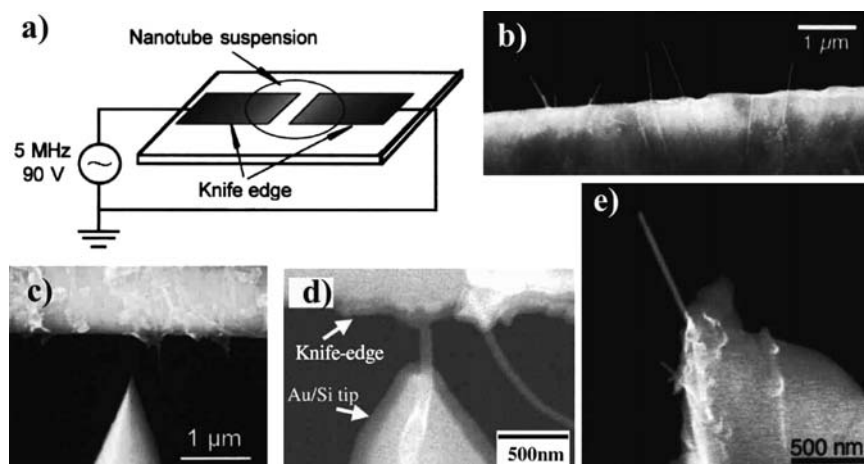


Fig. 5.5. Process for attaching single CNT to AFM tip under SEM control. (a) Setup for preparation of CNT source cartridge using AC electrophoresis [21], SEM micrograph of (b) CNT source cartridge [21] (c) AFM tip and CNT cartridge [21], (d) attachment of CNT to AFM tip [44], (e) completed single MWNT AFM tip [21]

edges to attract CNT to the two knife edges. The applied voltage orients the CNT tube axis parallel to the electric field lines and van der Waal's forces bind the CNT to the knife edge (Fig. 5.5b). In step 2, a CNT cartridge and the AFM probe tip are mounted to two separate and independent translation stages inside a modified SEM. The AFM tip is aligned with or contacted to a CNT. In step 3, the CNT and AFM tip are welded together by either applying a DC voltage, a DC current, or depositing an amorphous carbon film by focusing the electron beam on the CNT/tip junction. Finally, the CNT-tip is withdrawn from the cartridge. Such tips have been used to image DNA [8, 21], for noncontact imaging of DNA helical turns [44], to observe SWNTs on mica [15], to study the effect of MWNT stiffness on AFM imaging [45], magnetic force microscopy [20], materials characterization applications [23], imaging water solvation shells [31], friction force microscopy [46], surface characterization [37], simultaneous observation of sample topography and contact current under contact mode AFM [16], obtaining adhesion force images showing a chemical contrast between patterned hydrophilic and hydrophobic areas on a substrate [47], and comparison of the wear resistance of MWNT tips and polycrystalline silicon tips [37, 48]. Kajeshima et al. used this same method to attach an MWNT to a quartz tuning fork for noncontact AFM in a liquid environment [3]. In a variation of this method, Kuwahara et al. fabricated CNT AFM consisting of double-walled nanotubes [39]. Wolny attached a single CVD Fe-filled CNT to an AFM tip using a micromanipulator mounted inside an SEM to fabricate an MFM tip [14].

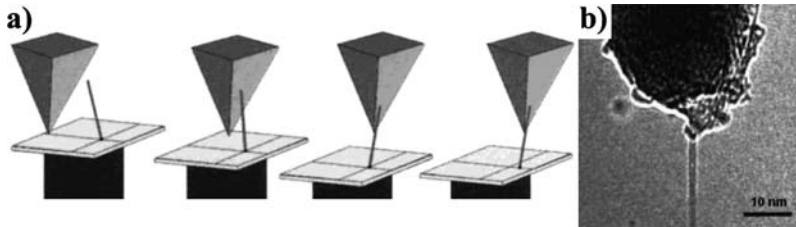


Fig. 5.6. (a) Schematic illustration of the “pick-up method” [11], (b) TEM image of SWNT tip (Scale Bar: 10 nm) [11]

5.3.3 Fabrication of Carbon Nanotube Atomic Force Microscopy Probes by In Situ Pick-Up

The third most common method for post-growth attachment is the “pick-up method” for fabricating single CNT AFM tips (Fig. 5.6) [11]. An AFM tip is scanned in tapping mode along a substrate consisting of vertically freestanding SWNTs. Due to van der Waal’s forces, an individual SWNT adheres to the tip [11–13, 17–19, 25]. This tip has been shown to provide superior AFM and EFM imaging quality compared to conventional silicon tips [12, 13, 17], functionalized with COOH open-end functionality [11], coated with fluorocarbons to increase the mechanical stability of the SWNT [24], and used for characterization of mesoporous materials [34], for imaging hydrated RNA–polymerase–DNA complexes [28] and DNA [22, 36].

5.3.4 Miscellaneous Methods for Post-Growth Attachment of Carbon Nanotube to Atomic Force Microscopy Tips

Other post-growth CNT attachment processes have recently been reported. In a novel technique, a MEMS microgripper was used to fabricate a single-tip CNT AFM probe [32]. The microgripper plucked a CNT from a substrate and placed the CNT onto the AFM tip. A subsequent electron beam deposition of carbon film welded the CNT to the tip. In an alternative technique, Hall et al. used a magnetic field to attach and align arc-discharge synthesized MWNTs dissolved in dichloromethane to gold-coated tips [10].

5.3.5 Metal Catalyst-Assisted Direct-Growth of Carbon Nanotube Atomic Force Microscopy Probes

Mechanical attachment of CNTs is a time-consuming process; therefore, recent studies have focused on developing a batch fabrication process for the mass production of CNT AFM probes (Fig. 5.3) [38, 40–43, 49, 50]. A SWNT AFM probe fabrication process flow is shown in Fig. 5.7a [38] and a MWNT process is shown in Fig. 5.7b [41]. Alternate processes for direct growth of SWNT and MWNT via catalyst-assisted CVD, and direct growth of carbon nanofiber tips

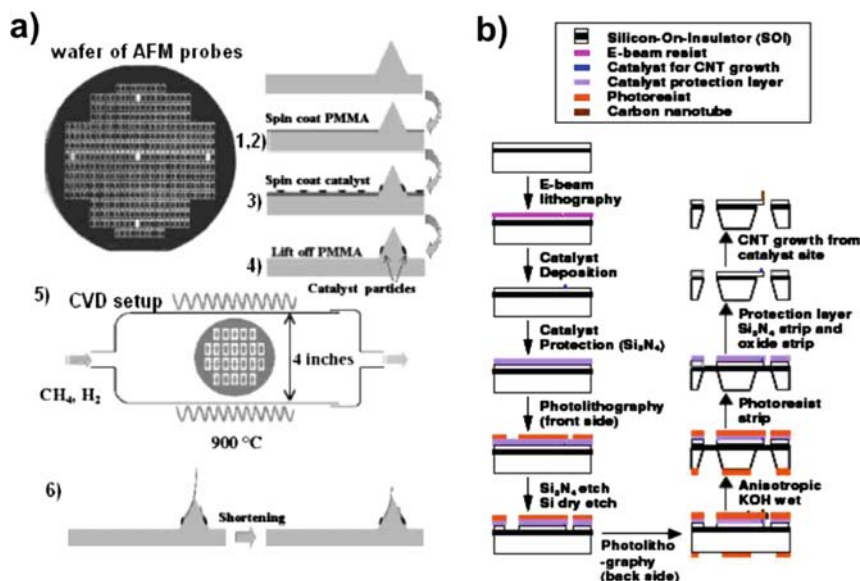


Fig. 5.7. Process flow for batch fabrication of (a) SWNT tips [38] and (b) MWNT tips [41]

on carbon-coated AFM tips via Ar irradiation are described in the literature [40, 43, 49, 50]. The SWNT process flow involves: (1) PMMA spincoat onto a wafer of AFM probes, (2) selective removal of PMMA over the AFM tip area, (3) spin-coat of a metal catalyst layer over the entire wafer, (4) lift-off of the PMMA layer selectively leaving the metal catalyst over the AFM tip only, (5) CVD growth of SWNT onto the AFM tip, and (6) CNT shortening by mounting the probe in an AFM and passing current through the CNT as it contacts a conducting substrate (Fig. 5.7a) [38]. This process yields either tips with single SWNT or bundles of SWNT [38]. Tips fabricated in this fashion have been used to image biomolecules and DNA [51] and Staphylococcus Protein A [38]. The MWNT process flow (Fig. 5.7b) consists of a novel mass production scheme combining fabrication of AFM cantilevers and direct growth of MWNT tips into a single process starting from a bare silicon-on-insulator (SOI) wafer. The fabrication process flow involves seven key steps: (1) wafer-scale nanopatterning and registration via electron beam lithography of a PMMA mask layer to define the catalyst areas, (2) Ti barrier and Ni catalyst deposition via evaporation, (3) lift-off of PMMA leaving Ti/Ni dots over the CNT tip area (4) deposition of a CVD silicon nitride layer to protect the catalyst on the tip area (5) silicon microfabrication of cantilevers via a dry reactive ion etch of the front-side of the wafer to define the cantilever profile and a backside wet KOH etch to define the cantilever thickness, (6) release of the silicon nitride protection to expose the catalyst, and (7) directional growth of CNT on the cantilevers via PECVD. This process

achieved individual MWNT growth from 200-nm catalyst dots [41]. Alternate processes for batch fabrication of MWNT AFM tips include fabrication of the AFM probe from a bare silicon wafer but replace direct CNT growth with dielectrophoresis trapping of SWNT and MWNT onto the patterned cantilever have been reported [26, 52]. Direct growth of CNT as a fabrication process for AFM probes currently still lacks the control necessary to achieve high yield of probes as well as CNT tip structural parameters such as diameter length and orientation. This is mainly due to two following reasons: (1) the inability to controllably pattern nanoscale catalyst in a highly reproducible manner over an entire wafer and (2) CVD growth techniques for CNTs with absolutely uniform diameters and lengths are also not highly reproducible over an entire wafer because CVD reactor technology for such a process is currently not available. It should be noted that for CNT AFM probe applications beside diameter and length control, the control of the orientation of the CNT tip is also extremely important in terms of obtaining CNT probes that would perform well and be easy to use by the AFM end users.

5.3.6 Post-Growth Attachment is Currently the Most Optimal Fabrication Process

The most widely used techniques for fabricating CNT AFM probes involve post-growth attachment of a single CNT or bundles of CNT. To illustrate this process, we present a step-by-step protocol for fabricating a multi-walled carbon nanotube atomic force microscope probe tip (MWNT AFM probe) developed in our laboratory [9, 29, 35, 53–55]. A similar method in which MWNTs dispersed onto a brass plate are mounted to an AFM tip is reported in the literature [27]. Fabrication of the MWNT AFM probe involves three main steps: (1) fabrication of MWNT source, (2) deposition of nickel adhesion layer onto silicon AFM probe, (3) attachment of a single MWNT to metal-coated silicon AFM probe tip.

Step 1: Fabrication of MWNT Source

Vertically aligned MWNTs grown via chemical vapor deposition (CVD) on the sidewalls of iron-coated wires serve as the MWNT source with some of the MWNTs individually separated and vertically aligned with respect to the surface of the iron-coated wire. The wire is cut into several 2-inch sections and placed in a 1-inch diameter quartz tube. The compression port serving as the gas inlet is sealed onto one end of the quartz tube and a compression port serving as the exhaust is sealed onto the opposite end. The entire assembly is placed in a tube furnace and the tube is purged with argon. With argon flowing through the tube, the temperature of the tube furnace is raised to 750°C. Once the temperature reaches 750°C, an additional 20 standard cubic centimeter per minute (sccm) of H₂ gas is flowed through the tube for 10–15 min to purge the tube. Afterwards, the H₂ gas flow is turned off and 80 sccm of Ethylene (C₂H₂) is flowed through the tube for 20 min to grow MWNT. Afterwards,

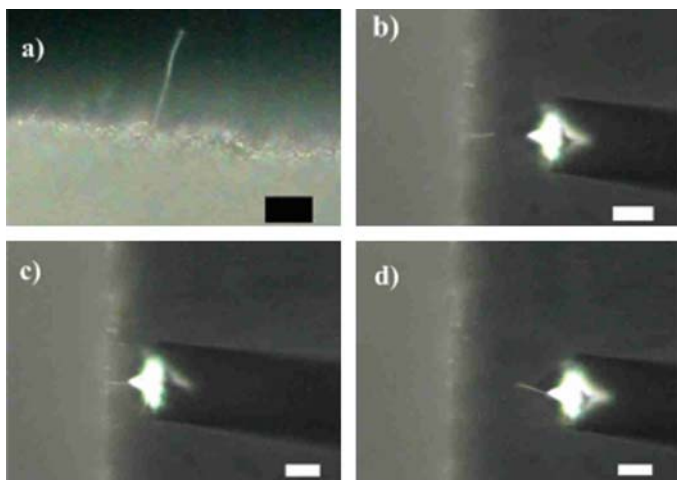


Fig. 5.8. Optical Micrographs at 500X magnification of (a) MWNT grown on a metal wire (scale bar: 40 μm), (b) alignment of MWNT with AFM tip, (c) MWNT in contact with probe tip as DC voltage is applied (d) MWNT tip (b–d: scale bar = 20 μm)

the C_2H_2 flow is turned off and argon is flowed through the tube for 5–10 min to remove any HC gaseous residues from the tube. This step yields MWNT of approximately 10–20 μm in length some of which are vertically aligned perpendicular to the catalyst-coated wire surface (Fig. 5.8a).

Step 2: Deposition of Metal Adhesion Layer onto Silicon AFM Probe

The next step in the fabrication process is to prepare the silicon AFM probe for attachment of the MWNT which requires a metal adhesion layer between the silicon and MWNT. Our own process flow uses sputter deposition of the metal adhesion layer. Experimental observation indicates that nickel provides the most reliable and robust adhesion layer [unpublished results]. The thickness of the nickel adhesion layer influences the MWNT/Ni junction's current–voltage characteristic and the mechanical behavior. Electrically, MWNT/Ni junctions in which the thickness of the nickel adhesion layer is greater than 25 nm are electrically ohmic whereas MWNT/Ni junctions with nickel adhesion layer thicknesses less than 25 nm are rectifying [unpublished results]. However, thinner nickel adhesion layers are more mechanically robust as evidenced by experiments in which the attached MWNT is pushed against a hard object. There is an obvious trade-off in the electrical and mechanical performance of the MWNT/Ni junction and, depending on the application, the thickness of the nickel adhesion layer may be tailored accordingly (i.e., thicker nickel adhesion layer for current-sensing applications and thinner nickel adhesion layer for pure AFM imaging).

Step 3: Attachment of Single MWNT to metal-coated silicon AFM probe tip

Our setup for attaching MWNT to the nickel-coated AFM probe tips consists of two XYZ micropositioners mounted to a sample stage on an inverted microscope providing $\times 50$ and $\times 500$ magnification. The nanotube source and the nickel-coated AFM probe are mounted to the micropositioners. The two mounts provide electrical connections to the nanotube source and to the AFM probe. The output and ground leads of a 0–50 V DC power supply are connected to each mount. The DC power supply is turned on and initially outputs 0 V. Under low magnification and using the micropositioners, the nanotube source and the AFM probe are advanced toward each other into close proximity. Under high magnification, the tip of the AFM probe is brought into contact with an MWNT on the nanotube source and the output voltage of the DC power supply is increased until the MWNT breaks at a defect location (Fig. 5.8b, c). The MWNT is attached to the tip of the AFM probe as shown in Fig. 5.8d.

Step 4: Application-specific post-fabrication processing

Typically, CNT AFM probe requires post-attachment modification such as sharpening and shortening in order to enhance the attainable resolution and adjust the mechanical response of the CNT. Sharpening the CNT tip improves the lateral resolution of the CNT AFM probe. Shortening the CNT influences the CNT mechanical response particularly when imaging samples in solution. Sharpening and shortening on the CNT are accomplished using the same equipment and setup described in step 3 except that a sharp gold STM tip is used in place of the nanotube source. Under optical control, the CNT is contacted with the STM. The voltage of the DC power supply is increased until Joule's heating between the STM tip and the CNT burns off portions of the CNT resulting in a sharpened or shortened CNT. Alternative methods for sharpening and shortening are given in the literature in which the sharpening process is performed in situ in an AFM employing the force distance curve [11, 30].

5.4 Characteristics and Characterization of Carbon Nanotube Atomic Force Microscopy Tips

Following CNT tip fabrication, scanning electron microscopy (SEM) characterization of the completed CNT tip's length and attachment angle defined as the angle between the CNT axis and the surface normal of a flat substrate is a necessary step prior to imaging. Tip resolution is limited by noise from thermal vibration of the CNT and by lateral force components which act on the CNT perpendicular to the tube axis and which increase in magnitude as the attachment angle increases. To limit thermal noise, the vibration of the tip must be limited to less than 1 nm. Therefore, the length of an SWNT tip must

be less than 100 nm and the length of an MWNT tip with diameter greater than 10 nm must be less than approximately 1 μm [56, 57]. Thus, an obvious trade-off between optimal resolution and three-dimensional depth profiling of samples with nonplanar high aspect ratio features exists. To reduce the lateral force components acting on the CNT, the attachment angle must be as close to zero as possible [25, 58]. SEM characterization of the CNT tip structure will provide a visual first clue of the attainable image quality. To more accurately characterize the imaging capabilities of the CNT tip, the user may then employ more refined techniques for quantifying the maximum attainable resolution [11, 25, 54] and for gauging CNT tip lifetime [54, 56, 59, 60].

To motivate the need for CNT tip characterization and highlight the advantages of CNT tips over conventional microfabricated silicon AFM tips, a survey of reported investigations is conducted describing (1) the mechanical behavior of CNT substrate interaction, (2) the effect of CNT tip length, high aspect ratio, attachment angle, and tip radius on two-dimensional image resolution and three-dimensional depth profiling, (3) CNT tip electrical characteristics, and (4) tip lifetime. Embedded in the discussion are procedures that the reader can use in his or her own laboratory for characterizing the CNT tip structure, resolution, electrical conductivity, and lifetime. We refer the reader to the cited literature for more detailed characterization procedures. For an introduction to SEM, we refer the reader to an excellent reference [60].

Investigations conducted to elucidate the mechanics underlying the CNT tip and substrate interaction utilize Force Spectroscopy [11, 53], test imaging of nonplanar structures [45, 58], and Frequency Modulation AFM (FM-AFM) [33, 55, 61]. In Force Spectroscopy, the CNT AFM probe is retracted from and extended to a hard surface in contact mode while monitoring the cantilever deflection versus the piezoelectric displacement generating a force versus distance curve (Fig. 5.9). From this force distance curve, information regarding and SWNT [11] and MWNT [53] tip mechanical behavior in response to applied forces can be extracted. Force-distance plots indicate that the CNT

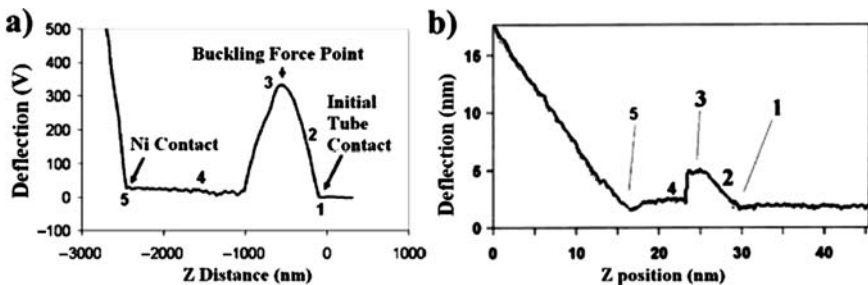


Fig. 5.9. (a) MWNT tip and substrate interaction [53], (b) SWNT tip and substrate interaction [11] showing five interaction regimes: (1) end contact of the CNT with the substrate, (2) cantilever deflection, (3) CNT bending, (4) CNT sliding, (5) contact of the support layer with the substrate

substrate contact involves a five-step mechanism as shown in Fig. 5.9: (1) initial end-contact of the CNT with the substrate, (2) cantilever deflection which is similar to that of a conventional silicon tip and which is mitigated by CNT stiffness withstanding the axial compressive force, (3) initial CNT bending starting at the buckling force point, (4) CNT sliding along the surface leading to a near constant deflection, and (5) contact of the underlying nickel-coated or bare silicon tip apex with the substrate [11, 53].

Test imaging of three-dimensional structures correlates CNT stiffness with imaging quality. The imaging capabilities of AFM tips consisting of a single unsupported MWNT and stiffer MWNT AFM tips consisting of a single MWNT supported with MWNT bundles at the MWNT tip's base were compared in imaging experiments performed on DVD pits [45]. Due to van der Waal's forces causing the sidewalls of MWNTs to adhere to the walls of the pits, unsupported single MWNT tips buckle due to the adhesive force between the MWNT sidewall and the walls of the DVD pits leading to instability along the pit sidewalls (Fig. 5.10a); however, the stiffer supported MWNT tips were able to resist adhesion force-induced MWNT buckling improving image stability (Fig. 5.10b) [45]. A similar experiment imaging silicon dioxide trenches

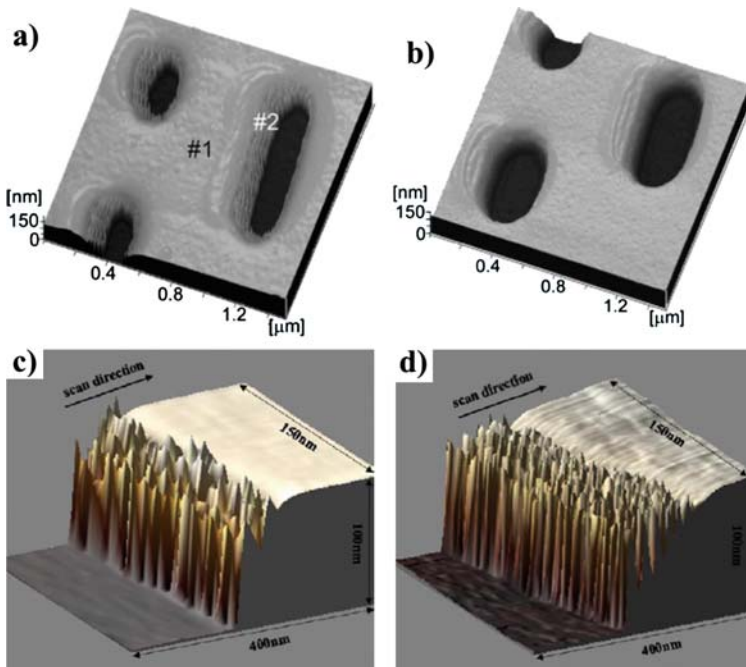


Fig. 5.10. AFM images of DVD pits showing (a) instability along pit sidewalls when imaging with unsupported MWNT tips and (b) improved stability with supported MWNT pits [45]. Ringing artifacts appearing when imaging trench sidewalls with (c) 0.4- μm length MWNT probe and (d) 1.2- μm length MWNT probe [58]

with MWNT tips found that adhesion and stiction of single unsupported MWNT tips to the trench sidewall led to an imaging artifact composed of a ringing pattern (Fig. 5.10c) [58]. These results suggest that stiffer CNT tips which are better able to resist adhesion force-induced CNT bending produce better images.

Force measurements obtained through FM-AFM of CNT tips on noncompliant substrates allows for the estimation of the magnitude of the adhesion force and the elastic bending constant which measures CNT stiffness. FM-AFM is a special technique in which positive feedback is used to oscillate an AFM tip at a resonance frequency. Tip substrate interaction forces result in a frequency shift which can be used to generate images or estimate the magnitude of the forces acting on the tip. From this resonant frequency shift, the CNT elastic bending constant can be extracted. The adhesion force mitigating the interaction between the substrate and the CNT free-end can be determined from the FM-AFM damping signal [61]. FM-AFM studies of SWNT tips on silicon and graphite substrates were performed to extract the SWNT elastic bending constant and the magnitude of the SWNT substrate adhesion force [61]. From the resonant frequency shifts, the effective SWNT bending spring constant was estimated to be approximately 4.8 mN m^{-1} for the case in which the free end of the SWNT is unable to slide freely along the substrate. From the damping signal, the magnitude of the adhesive force was estimated to be approximately 0.2 nN [61].

Bernard et al. proposed a model describing CNT – both MWNT and SWNT – AFM tip substrate interaction in FM-AFM. The interaction involves three regimes: (1) a no interaction regime in which the CNT does not interact with the sample and no frequency shift is observed, (2) an intermittent contact regime consisting first of a main attractive force from adhesion of the sample to the CNT manifesting in a negative frequency shift and followed by an increasing repulsive interaction in which the CNT restoring force counteracts the adhesion force resulting in a positive frequency shift, and (3) a permanent contact regime resulting in a rapid increase in frequency. Thus, in imaging applications, the CNT tip experiences two counteracting forces – an adhesion force attracting the CNT to the substrate causing the CNT to bend elastically counteracted by a repulsive restoring force in which the CNT reverts back to its original vertically aligned orientation [33]. A similar analysis for a coiled CNT tip is presented in Dietzel et al. [55].

From the studies presented above, it is evident that the major forces acting on the CNT tip are axial forces applied parallel to the CNT axis, lateral forces applied perpendicular to the CNT axis, and adhesive forces between the substrate and CNT sidewalls. Clearly, these forces affect the imaging capability of the CNT tip. The magnitudes of these forces and their relative contributions to imaging resolution and stability are determined by the CNT stiffness, the CNT length, attachment angle, and tip radius.

Studies examining the effect of CNT length on AFM imaging indicate that long CNTs suffer from reduced resolution due to larger amplitude thermal

vibration of the CNT tip and increased susceptibility to CNT bending. Thermal vibration of the CNT contributes noise degrading the signal-to-noise ratio (SNR) and the image quality. The thermal vibration is given by (5.3) which was proposed for a DWNT tip [39]:

$$L^3 = \frac{3\pi r^4 Y}{4k_B T} u^2 \quad (5.3)$$

where u is the upper limit amplitude vibration, r is the radius of the DWNT measured by high-resolution TEM, Y is Young's modulus, T is temperature in Kelvins, and k is Boltzmann's constant. Using this equation, the optimal DWNT length was found to be approximately 100 nm [39]. Long SWNT tips are also prone to higher thermal vibration and higher noise [11]. For an SWNT tip, the estimated thermal vibration amplitude is given by (5.4 and 5.5):

$$X_{tip} = \sqrt{\frac{k_B T}{k}} \quad (5.4)$$

$$k = \frac{3\pi r^4 Y}{4l^3} \quad (5.5)$$

where k is the SWNT bending constant. These equations show that, as CNT length increases, the thermal vibration increases. This same trend has been observed for MWNT tips.

In addition to increased thermal vibration, long CNT tips are also more prone to bending induced by lateral forces acting on the CNT tips. Both long and short SWNT tips are able to produce stable images of flat surfaces; however, in imaging a nonplanar surface consisting of circular GaSb quantum dots, only shorter SWNTs produce stable images. Longer SWNTs bend leading to feedback instability and blurring (Fig. 5.11) [56]. Also, for SWNT tips with an aspect ratio greater than 10, spatial resolution is degraded due principally to CNT bending from lateral forces exerted by the sample on the CNT [25]. When imaging patterned pits, using MWNT tips of lengths below 200 nm eliminates the imaging artifacts observed when imaging DVD pits (Fig. 5.12) [23, 45].

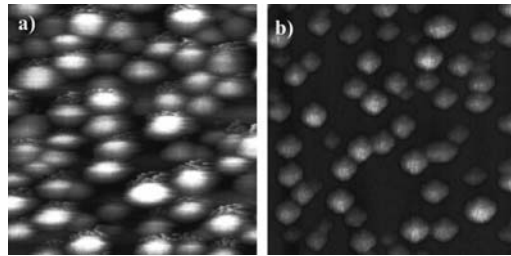


Fig. 5.11. Noncontact image of GaSb quantum dots taken with (a) long and (b) short SWNT tips. Blurring is evident in the left image [56]

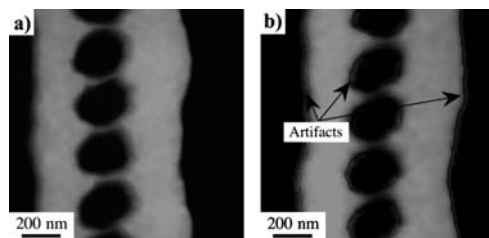


Fig. 5.12. Pit patterns taken with MWNT tips of length (a) 150 nm and (b) 300 nm. The image on the right is free of image artifacts [23]

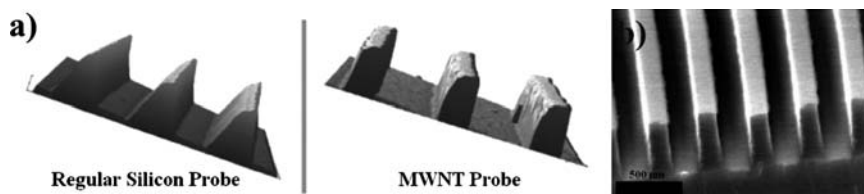


Fig. 5.13. (a) Illustration of high aspect ratio MWNT more accurately profiling 280-nm/line photoresist pattern compared to silicon AFM tips, (b) SEM micrograph of resist profile (scale bar = 500 nm) [54]

The limits on CNT tip length imposed by tip thermal vibration lead to an obvious trade-off between CNT tip aspect ratio and resolution. For imaging applications allowing for less aggressive resolution requirements, the high aspect ratio of the CNT can be exploited when performing three-dimensional profiling of deep trenches and line patterns. High aspect ratio MWNT AFM tip (diameter = 60 nm and length = 5 μm) generated an accurate three-dimensional profile of a 90-nm line and space photoresist pattern [41]. Furthermore, compared to a silicon AFM probe, a CNT tip was able to map a near vertical sidewall when profiling the height of deep trenches in photoresist (Fig. 5.13) [26, 54].

Having fully accounted for the effect of CNT length on image resolution, the attainable image resolution of the CNT tip is limited by the tip radius of curvature which is the contact point between the sample and the CNT. The radius of curvature determines the image sharpness which is improved as the radius of curvature decreases. The tip radius is estimated by using the CNT tip to image a sample of gold colloid particles on mica and measuring the full width at half-maximum height (fwhm) of the gold particles [11]. With this technique, the radius of curvature were estimated to be 6 and 3.4 nm for mechanically assembled SWNT and MWNT tips and 1.5–4 nm for CVD direct grown SWNT tips [11]. An alternative method for measuring the radius of curvature involves measuring the fwhm of SWNT lying on a silicon substrate [25]. SWNT AFM tips are able to resolve 3-nm silicon nitride grains [54]. In imaging DC sputtered amorphous silicon films, Hudspeth et al. showed

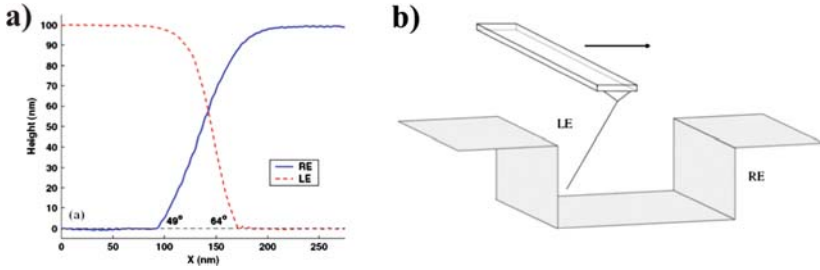


Fig. 5.14. (a) Illustration of distortion in image of trench sidewalls due to nonzero attachment angle, (b) schematic showing effect of attachment angle [58]

that the obtainable lateral resolution of a CNT tip was many times better compared to a silicon tip [35].

Another factor influencing the CNT tip's attainable image resolution and fidelity is the attachment angle. For SWNT tips, an attachment greater than 30° degrades resolution and produces shadowing artifacts due to an increase in the force perpendicular to the CNT axis causing the SWNT to buckle slightly near the tip [25]. In imaging trench sidewalls, non-zero attachment angle results in ringing artifacts. Furthermore, when imaging identical opposing vertical sidewalls, a non-zero attachment angle resulted in a distorted and asymmetrical two-dimensional profile of the trench (Fig. 5.14) [58].

Electrical characterization of an MWNT tip can be performed using the setup described in the Fabrication section for CNT shortening or contacting the MWNT AFM tip to a Platinum (Pt) surface via the AFM approach function. Provided a metal adhesion layer exists between the base silicon cantilever and the MWNT, the equivalent resistance consisting of three resistance in series: (1) the bulk resistivity of a single MWNT, (2) the contact resistance the MWNT's free end with the Pt substrate or Pt STM tip, and (3) the contact resistance of the MWNT with the adhesion layer. Using the latter method for a 60-nm diameter MWNT tip, the equivalent resistance was found to be less than 44 k Ω [13]. These electrical data may be valuable for CNT tip EFM and STM applications.

Finally, tip wear influences the long-term resolution capabilities of the CNT AFM tip. Continuous scanning for long periods of time typically results in a gradual blunting of the silicon tip radius of curvature leading to a reduction in image resolution. Determining the tip lifetime involves continuous scanning of a noncompliant surface containing nanoscale feature sizes for long periods of time while periodically measuring the change in resolution or surface roughness statistics. A 10-nm diameter MWNT probe scanning a hard silicon nitride surface shows no detectable lateral resolution degradation after 15 h of continuous scanning and was able to resolve 8–10 nm silicon nitride grain sizes. In contrast, a conventional silicon AFM tip shows lateral resolution degradation after 12 h of continuous scanning resolving grain sizes of 18–20 nm (Fig. 5.15) [27]. Using dynamic force AFM, the tip-wear resistance

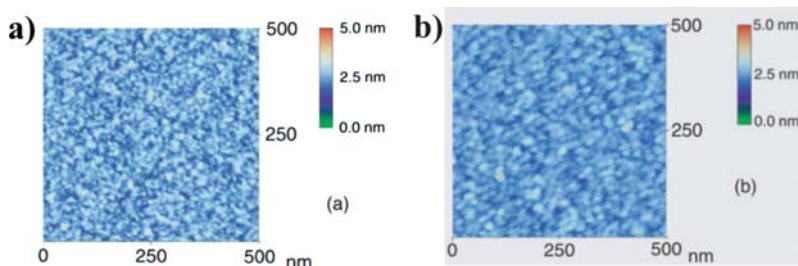


Fig. 5.15. Reduced tip-wear of (a) MWNT tip compared to (b) Si tip in imaging silicon nitride grain sizes [54]

of a super sharp silicon AFM tip and a MWNT AFM tip were compared by periodically measuring the RMS surface roughness of a P-type (100) silicon wafer over 150 scan frames and found that the MWNT tip was less prone to tip wear and that the lifetime of the MWNT tip was five times that of the silicon tip [59].

5.5 Applications of Carbon Nanotube Scanning Probe Microscopy

5.5.1 Functionalization of Carbon Nanotube Tips for Chemical Force Microscopy

Chemical Force Microscopy (CFM) is a variation of Friction Force Microscopy (FFM) in which the interaction between a chemically modified CNT tip and the substrate map into an image. For example, generating image contrast of a flat surface consisting of regions of varying hydrophilicity requires an AFM tip functionalized with a hydrophobic or hydrophilic group. Compared to conventional silicon and silicon nitride tips, CNT tips offer some advantages for CFM, namely: (1) CNT ductility is suited for CFM characterization of delicate biological and organic materials which would be damaged by hard noncompliant tips and (2) in imaging surfaces with varying hydrophobicity, the hydrophobic nature of the CNT sidewall would reduce adsorption of water along the tip enhancing contrast [47].

Two major methods for CNT covalent modification involve end functionalization via a carboxylic acid functional group and sidewall fluorination which allows for a succeeding substitution reaction (Fig. 5.16) [1]. Application of these methods to CNT AFM probes produces tips for adhesion studies [47, 62, 63] and stiff sidewall fluorinated high-resolution SWNT tips [24].

MWNT tips with end carboxylic acid functional groups are fabricated under SEM control resulting in an MWNT tip in which the free-end is open and the individual sidewalls are terminated by carboxylic acid groups [47]. Alternatively, closed ended MWNT may be purified and glued via an acrylic

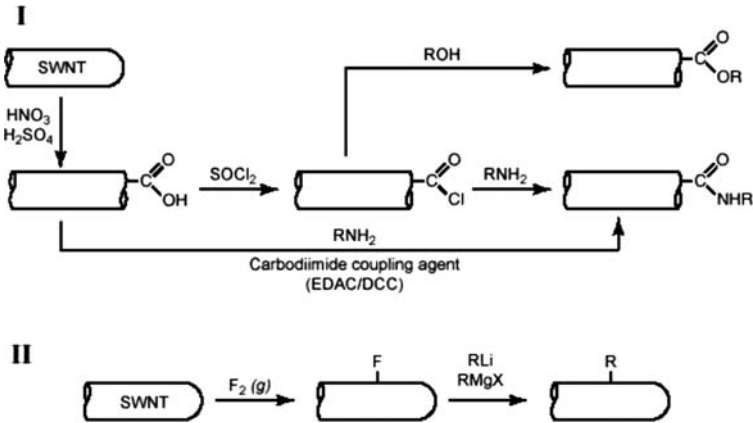


Fig. 5.16. Covalent modification of CNT via: (I) end carboxylation followed by conversion to an ester or amide functional group, (II) sidewall fluorination preceding organometallic substitution [1]

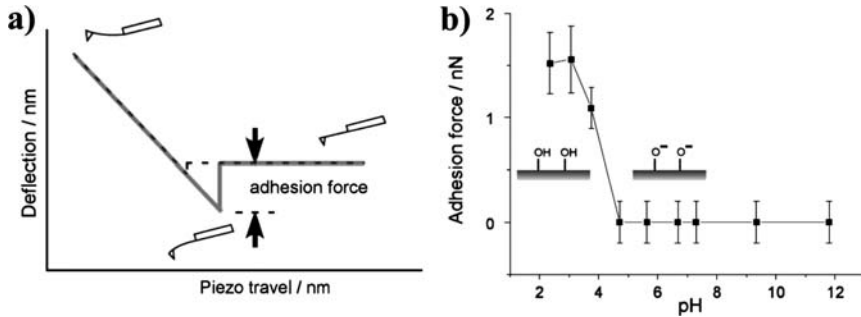


Fig. 5.17. (a) Force versus distance curve showing extraction of adhesion force, (b) force–titration curve on mica surface [47]

adhesive to a supporting tip and shortened thereby opening the MWNT end by end contacting the MWNT to a metal surface and applying a DC voltage (0.5–20 V) to pass a DC current [63]. For both processes, the mechanism leading to carboxylic acid termination is unknown although several models have been proposed [45, 63]. Carboxylic acid group end termination is verified by performing a force titration curve on a mica surface or a hydroxyl-terminated self-assembled monolayer (SAM) immersed in an adjustable pH phosphate buffer solution [47, 63]. At different pH levels, a force spectroscopy is performed generating a force versus distance curve from which the adhesion force which measures the stickiness of the MWNT tip to the surface is extracted (Fig. 5.17a). As shown in Fig. 5.17b, at pH above 4.2, the adhesion force is the weakest; furthermore, the end carboxylic acid group is deprotonated and negatively charged and the mica surface is negatively charged at this pH

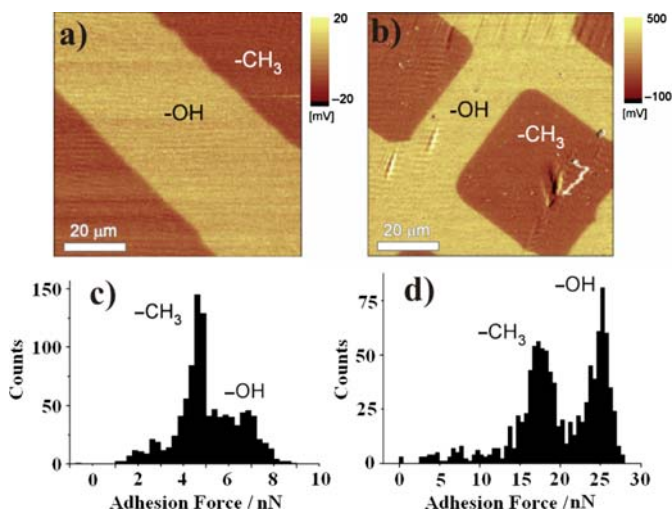


Fig. 5.18. FFM images of patterned hydrophobic and hydrophilic surfaces taken with (a) silicon nitride tip and (b) an open-ended MWNT tip. Histogram of adhesion force measurements taken on surface using (c) silicon nitride and (d) open-ended MWNT tip [47]

level verifying the existence of a carboxylic group at the tip end [47]. Force-titration experiments performed on a SAM on Au surface, APTES-treated, amine-terminated ($-\text{NH}_3$) mica surface, and carboxylic acid-terminated surface further attest to this interpretation [47,63]. FFM imaging of a patterned hydrophobic ($-\text{CH}_3$) and hydrophilic ($-\text{OH}$) surface using silicon nitride and end-functionalized MWNT tips clearly indicate that both tips show well-defined contrast between hydrophilic and hydrophobic surfaces (Fig. 5.18) [47]. Both tips are hydrophilic; however, adhesion force measurements taken at different locations in the sample show that, because the hydrophobic MWNT sidewalls resist water condensation around the tip unlike the silicon nitride tip which is easily wetted, the MWNT tip is better able to differentiate between hydrophilic and hydrophobic areas as evidenced by the adhesion force histogram's nearly unimodal distribution obtained using the silicon nitride tip versus the bimodal distribution obtained with the MWNT tip (Fig. 5.18) [47]. Alternative terminating groups are achievable. Shortening MWNT in an activated H_2/N_2 environment results in H-terminated MWNT ends [63]. Exposing a hydroxyl-terminated MWNT to a reaction solution consisting of AET and DEC converts the termination from a carboxylic acid to an amide [47]. Analogous results have been demonstrated with shortened SWNT tips [62].

A recently reported technique involves fluorocarbon Teflon-like polymer coating of SWNT which stiffens the CFM probes while producing potentially high-resolution probes with unique electrical properties [24]. Using the

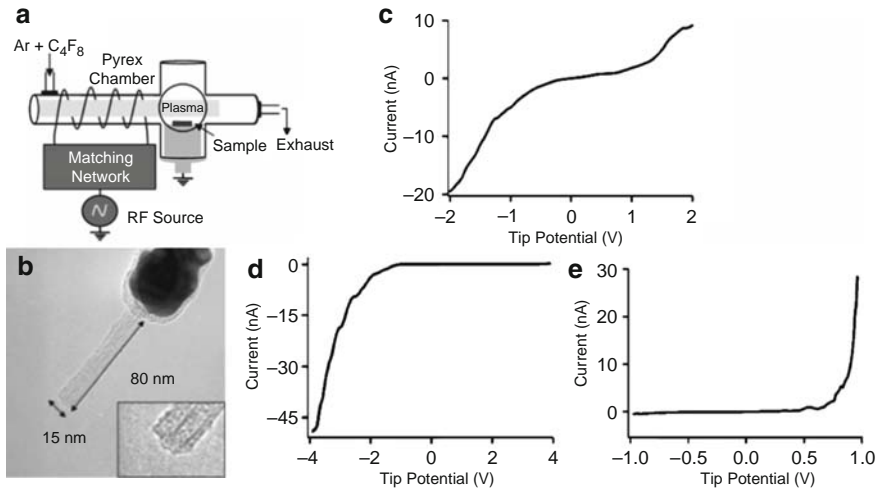


Fig. 5.19. (a) Schematic of ICP reactor, (b) TEM image of Teflon-polymer-coated SWNT electrode after tip exposure (Inset: close-up of tip), I–V curves of (c) semiconducting SWNT attached to gold tip, and I–V curves of SWNT showing (d) negative bias rectification and (e) positive bias rectification [24]

pick-up technique, CVD grown 5-nm diameter SWNT are attached to silicon or gold AFM supporting tips producing protrusion lengths of 40–350 nm. Thin layer polymer deposition was performed using an inductively coupled plasma (ICP) reactor (Fig. 5.19a) [24]. Power was delivered by a 600-W variable RF amplifier with a 21-MHz sinusoidal signal serving as the source through a pi-impedance matching network to a four turn copper antenna wrapped around the upstream extension tube of the chamber. The reactor conditions for polymer deposition were 7 sccm argon, 1.2 sccm octafluorocyclobutane (C_4F_8), 160 mTorr, 50–75 W for 60–90 s. Etching with a 5-V electrical pulse exposed and opened the SWNT tip resulting in carboxylic acid terminated tips (Fig. 5.19b).

Electrical testing was performed by contacting the tip with liquid mercury (Hg). For low bias voltages, the etched tip yielded a resistance of 177 k Ω which agrees with published estimates of SWNT resistivity suggesting that the deposition process did not affect the SWNT. However, these nanoelectrodes displayed unique electrical properties. Esplandiu et al. performed current–voltage measurements on several such nanoelectrodes and found that semiconducting SWNT attached to an Au tip showed no rectification (Fig. 5.19c), and some nanoelectrodes in which SWNT was attached to *n*-type silicon showed rectification at negative bias (Fig. 5.19d) whereas other SWNT attached to Au tips showed rectification at positive voltages (Fig. 5.19e) [24].

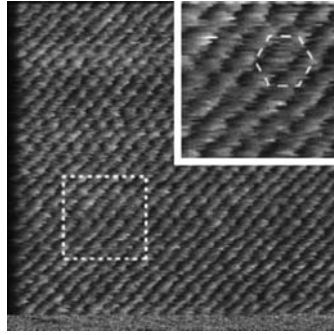


Fig. 5.20. FFM $12\text{ nm} \times 12\text{ nm}$ micrograph of mica surface taken with CNT probe 50 nm in length. Hexagonal symmetry and $\approx 5.5\text{ \AA}$ periodic spacing of mica is shown [46]

5.5.2 Carbon Nanotube Friction Force Microscopy

Friction Force Microscopy (FFM) is a contact mode technique in which the sample exerts a lateral force on the tip producing torsional bending of the cantilever. As the tip is scanned over the surface, the variation in friction is mapped into a two-dimensional image. Image contrast can result from steps or holes on the sample and from differences in tip-sample forces on a heterogeneous material. With sharp tips, FFM can attain atomic-scale resolution.

CNT tip's high-resolution stemming from small achievable radius of curvature has been used in FFM imaging of mica. CNT attached FFM tips are produced under SEM control [45, 46]. Successful CNT-FFM imaging is length dependent because long CNT buckle and bend preventing atomic resolution [64]. Moreover, estimates from the CNT bending spring constant as estimated from the equation $k_{\text{CNT}} = 3EI/L^3$ indicates that, for atomic-scale FFM imaging, the bending spring constant must match the torsional spring constant of the adjoining cantilever [46]. Otherwise, the lateral force on the CNT will result in CNT elastic deformation only without the necessary corresponding translation into cantilever bending. Using 50-nm length CNT tips, atomic-scale resolution of the surface of mica can be achieved (Fig. 5.20) [46].

5.5.3 Carbon Nanotube Electric Force Microscopy

CNT Electric Force Microscopy (EFM) is a noncontact AFM technique which is derived from tapping mode and relies on the capacitive coupling between a conducting tip and the substrate. In response to an applied voltage between the tip and substrate, opposite charges develop on the tip and surface leading to an electrostatic Coulombic force. Variations in the electrostatic force between the tip and substrate depends on the local electronic properties of the substrate and are transduced into variations in resonant frequency and phase

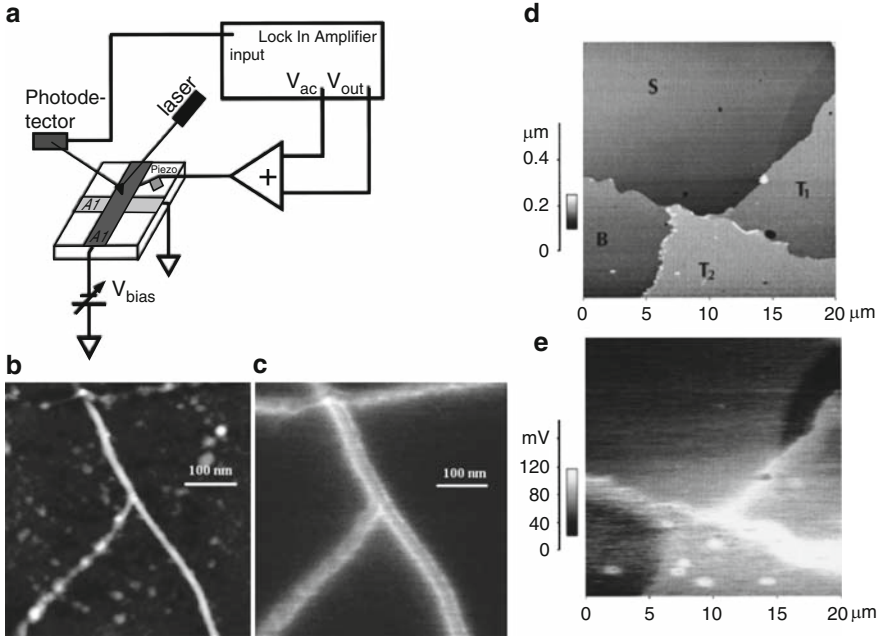


Fig. 5.21. (a) Schematic diagram of typical CNT-EFM setup [7], (b) topography and (c) EFM images of Cr electrodes at ends of SWCNT lying on silicon dioxide taken with SWCNT tip [13], (d) topography and (e) EFM images of Al–Al₂O₃–Al tunnel junction [7]

shift from which an EFM image is generated (Fig. 5.21a). Unlike STM which can only be performed on conducting substrates, EFM can image insulating, conducting, or semiconducting surfaces and is ideal for characterization of and defect detection in nanoscale semiconducting devices and determining the distribution of different materials in a composite. EFM suffers from the following drawbacks: (1) limited spatial resolution, (2) tip effects which complicate interpretation of image data, and (3) increased sensitivity to topographical differences on the sample.

Because of CNT high aspect ratio and small achievable radius of curvature, CNT EFM tips can potentially alleviate EFM sensitivity to topographical differences while simultaneously enhancing the spatial resolution. The high aspect ratio also minimizes stray capacitive coupling between the AFM cantilever and the substrate further enhancing image quality. SWCNT and MWNT EFM tips can be fabricated using the pick-up technique or using an adhesive to paste CNT to a supporting AFM tip [11, 65]. SWCNT and MWNT must be metallic and are attached to highly doped silicon AFM tips. MWNT attached using adhesive must be checked for ohmic conductance between the MWNT and silicon junction and can potentially be used for several months [7]. A bundle of SWCNT attached through the pick-up technique results in a single

SWNT protruding from bundle with a length of 150 nm and a diameter of 5 nm [13]. Topographical and EFM images of CVD grown SWNT on silicon dioxide with both ends covered with Cr electrodes were simultaneously collected with an SWNT EFM tip and are shown in Fig. 5.21b, c [13]. Topographical and EFM images of Al–Al₂O₃–Al tunnel junction simultaneously obtained with an MWNT tip are shown in Fig. 5.21d, e [7]. SWNT and MWNT EFM tips are capable of resolving 15 nm and 10 nm minimum feature sizes [7, 13]. An added advantage of CNT–EFM tips is simplified interpretation of EFM image data. As demonstrated by Wilson et al., the electrostatic force between a CNT tip and a metallic substrate can be modeled with an inverse power law fit which simplifies tip effect deconvolution [13].

5.5.4 Carbon Nanotube Scanning Tunneling Microscopy

Scanning tunneling microscopy (STM) is an AFM technique that probes the density of electron states in a material via a tunneling current. In STM, the tip is brought into close proximity with the surface and a bias voltage is applied between the tip and surface inducing an electron current to tunnel through the thin vacuum separating the tip and the sample surface. An image is formed by mapping the differences in tunneling current amplitude over the sample surface. The necessary prerequisite for successful STM imaging is a sharp tip and a clean environment; consequently, STM is often performed in ultra high vacuum (UHV).

The CNT AFM tip satisfies many requirements for successful STM imaging in that the CNT tip is resistant to crashes due to its mechanical flexibility. Furthermore, sharpening allows for the fabrication of CNT tips with a small radius of curvature that enhances the electron tunneling current. Moreover, CNT high aspect ratio and ductility enable STM studies of sample trenches and soft samples while reducing the risk of sample damage. Several recent investigations have examined the performance of CNT STM tips [65–68]. STM tips fabricated by attaching MWNT to gold tips under electron microscope control have successfully imaged Au [111] surfaces; however, STM imaging with MWNT tips is prone to sample contamination by adsorbates originating from the CNT tip itself [67, 68]. Adsorbate contamination can be eliminated by thermal annealing in UHV [67]. Furthermore, tip resolution is dependent on CNT length with lengths less than 200 nm giving the best resolution and imaging stability.

A second obstacle to CNT STM is the contact resistance which originates from the junction between the CNT and the supporting tip and which impedes electron tunneling [66]. This problem can be alleviated by blanket deposition of a thin layer metal over the entire CNT AFM probe. A CNT STM tip fabricated by MWNT attachment to a tungsten tip followed by pulsed laser deposition (PLD) of a thin W layer over the MWNT and supporting W tip was used to image *n*-type Si [111] surface in UHV [66]. The thin W coating preserves MWNT flexibility. Using this tip, an STM micrograph was

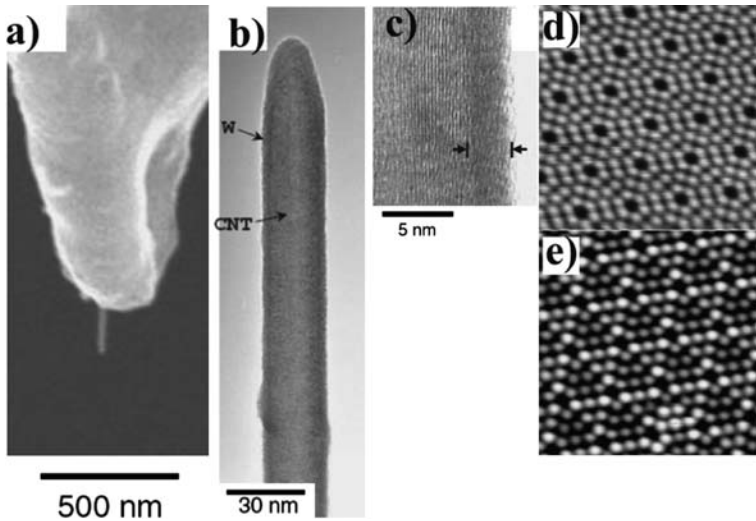


Fig. 5.22. (a) SEM micrograph of W-coated CNT STM tip. TEM micrographs of (b) tip and (c) W thin layer. (d) STM micrograph of filled (*bottom*) and empty (*top*) electronic states in Si [111] [66]

produced showing the empty and filled electron states in the silicon sample (Fig. 5.22) [66].

5.5.5 Carbon Nanotube Magnetic Force Microscopy

Magnetic Force Microscopy (MFM) is a microscopy technique that maps the interaction between the tip and sample mediated by magnetic fields. MFM has been used to study the surface of magnetic storage systems such as hard drives and contrast mechanisms. The main prerequisite for MFM imaging is an AFM tip with a magnetic moment that renders the tip sensitive to magnetic fields originating from the sample surface. Conventional microfabricated MFM tips consist of silicon or silicon nitride tips covered with a thin layer of hard magnetic material; however, these tips suffer from the following drawbacks: [1] the magnetic coating is prone to oxidation necessitating a protective coating or a UHV imaging environment and (2) the complex tip geometry leads to stray magnetic fields along the tip sides making the MFM measurements dependent upon tip geometry, thus complicating the extraction of quantitative data. Because of stray magnetic moments, the resolution of MFM tips is typically limited to approximately 30 nm.

Because of CNT high aspect ratio and relative inertness of the sidewalls, CNT MFM probes may potentially solve the problems associated with tip oxidation and stray magnetic fields in that CNT filled with magnetic material resist oxidation and the magnetic moment of vertically aligned CNT is perpendicular to the sample surface reducing the stray magnetic field. Several

candidate CNT MFM probes are based on iron-filled MWNT [14], catalyst-assisted CVD grown MWNT with magnetic metal catalyst at the tip [20], and MWNT blanket coated with a thin magnetic metal layer [69].

Iron-filled MWNT exploit CNT sidewall stability and high aspect ratio for durable long-lifetime high-resolution MFM probes; furthermore, the Fe-filling stiffens the MWNT mechanically stabilizing long ($\approx 20\ \mu\text{m}$) MWNT tips [14]. The process for fabricating such tips involves: (1) synthesis of Fe-MWNT from a ferrocene precursor, (2) magnetization of the Fe-core through an applied external step-wise increasing magnetic field (3) attachment of a single Fe-MWNT to an AFM probe, and (4) tailoring of attached Fe-MWNT by electron beam-induced CNT oxidation to improve the attainable resolution. Fe-MWNT are grown using thermal CVD in a two zone furnace on silicon wafers coated with a 10-nm aluminum buffer layer to improve MWNT alignment and a 2-nm iron top layer [14]. The two zone reaction chamber consists of a sublimation zone, where ferrocene is sublimated under a constant argon flow rate and the substrate is held at a constant 300°C temperature to prevent ferrocene condensation. In the reaction zone, the temperature is raised at a rate of $0.6\ \text{K s}^{-1}$ to support CNT alignment from 300°C to 600°C where ferrocene decomposes on the Fe-layer to produce Fe-filled MWNT $10\ \mu\text{m}$ to $25\ \mu\text{m}$ long [14]. SEM characterization of Fe-filled MWNT MFM probes shows that most MWNT contain unfilled parts along the MWNT length (Fig. 5.23)

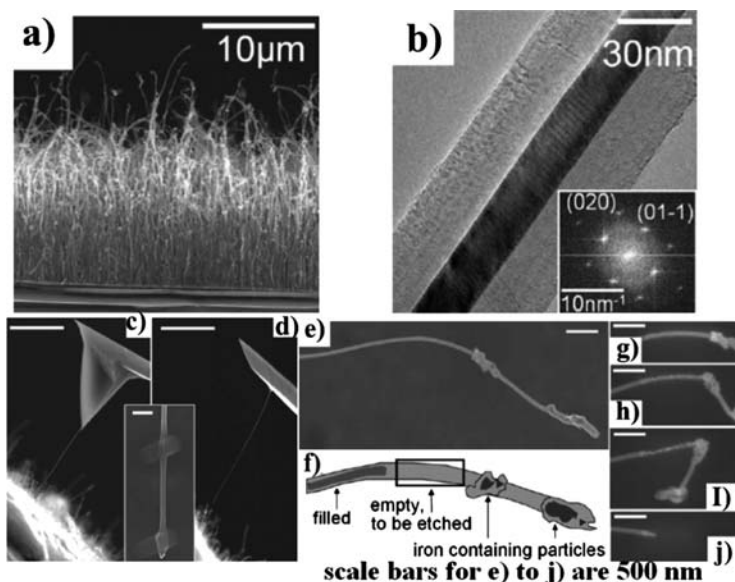


Fig. 5.23. (a) SEM micrograph of as-grown CVD Fe-CNTs on a silicon substrate, (b) TEM micrograph of Fe-filled MWNT (inset: high-resolution electron diffraction measurement of Iron-filling). SEM micrographs of attachment to (c) AFM tip and (d) tipless cantilever (inset: close-up of tip contamination) (e) partially filled Fe-MWNT with (f) schematic, and (g–j) process of oxidizing MWNT shell to filled end [14]

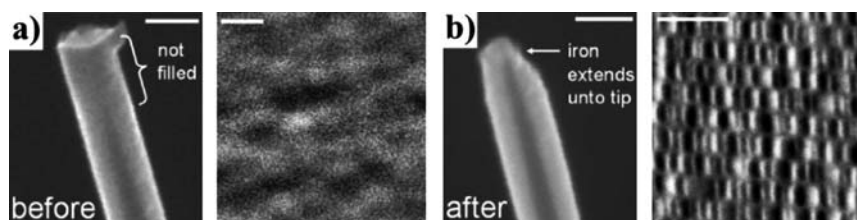


Fig. 5.24. Improvement in MFM lateral resolution evident in comparison of MFM image (right) of magnetic hard disk taken by (a) unfilled and (b) filled Fe-MWNT tip [14]

[14]. For high-resolution MFM imaging, the MWNT tip must be Fe-filled with large diameter iron particles; thus, unfilled portions at the MWNT tip are removed by SEM-controlled oxidation wherein the specimen chamber is filled with a 0.2-mbar water vapor pressure and a focused 10 kV, 2 nA electron beam burns off the unfilled portion of MWNT [14]. Test imaging of a magnetic hard drive shows that Fe-filled MWNT lengths up to 20 μm produce high-quality topography images attesting to the stiffness imparted by the Fe-filling on the MWNT (Fig. 5.24) [14]. Furthermore, MFM test imaging indicates that Fe-filled MWNT are able to resolve 60-nm magnetic domains.

MWNT MFM probes with magnetic metal localized at the CNT tip may minimize stray magnetic moments and enhance MFM resolution. Such tips are prepared from magnetic-catalyst assisted CVD growth of MWNT via a free-end-growth mechanism resulting in MWNT with magnetic metal particles localized at the tip [20]. To fabricate such MFM probes, a quartz substrate coated with 10-nm nickel is placed on a susceptor in a quartz tube and heated for 1 h at 800°C, 60 mTorr to form round 20-nm diameter nickel particles on the substrate. Afterwards, using helium as a carrier gas, benzene is flowed into the chamber producing 45-nm diameter MWNT several microns in length with nickel particles located at the free-end. Under SEM control, the MWNT is attached to a silicon AFM tip and magnetized axially by applying a pulsed 10 ms 12.5 T magnetic field resulting in SEM artifact swelling of the MWNT and nickel particle. Test imaging of a hard disk drive resolved 750-nm magnetic domains (Fig. 5.25) [20].

An alternative method for fabricating CNT MFM tips involves blanket deposition of magnetic alloy CoFe on a commercially available MWNT AFM probe by radio frequency RF sputtering (Fig. 5.26) [69]. Both MWNT shortening due to plasma-induced burning and tip widening from CoFe deposition were observed via SEM. Experimentally, it was found that probes coated with CoFe sputtering rates of 2.4 nm min^{-1} at 0.3 Pa and 100 W, 1.5 nm min^{-1} at 1 Pa and 100 W, and 1.2 nm min^{-1} at 1 Pa and 50 W measured minimal surface roughness [1 nm] of sputtered films indicating that MWNT MFM probes fabricated under these growth conditions obtain optimal MFM image resolution. Test imaging of an ultra-high-density magnetic storage media indicates that the CoFe covered MWNT resolve 10-nm features [69].

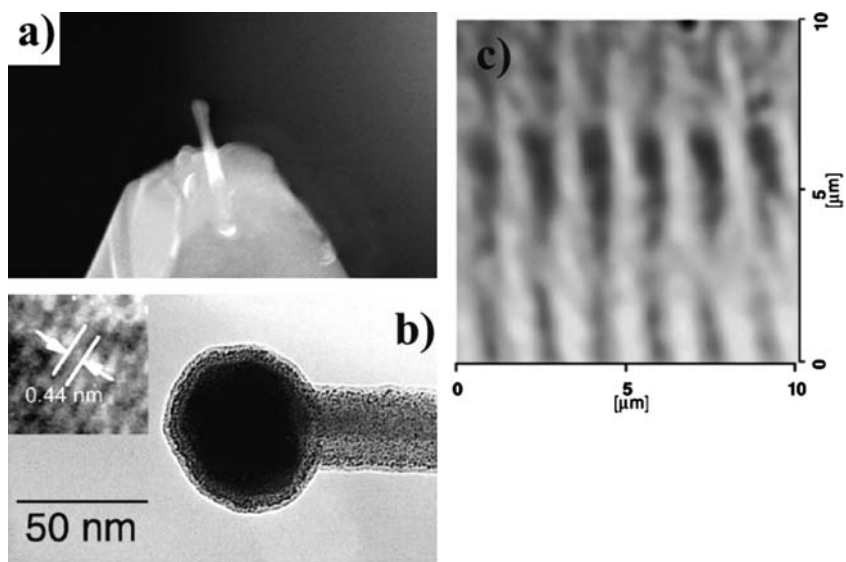


Fig. 5.25. (a) SEM micrograph of magnetized MWNT with nickel particle at tip. Swelling of CNT and tip is due to magnetization. (b) TEM micrograph of magnetized Ni₃C particle at MWNT tip (inset: magnified micrograph of observed fringes). (c) MFM micrograph of hard disk drive obtained with magnetized tip [20]

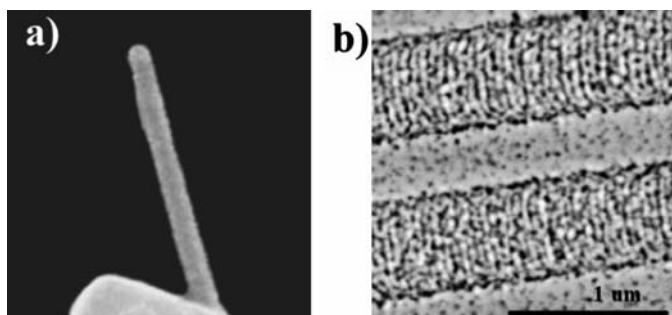


Fig. 5.26. (a) High-energy SEM micrograph of CoFe covered CNT probe, (b) MFM micrograph of ultra-high-density magnetic recording media (scale bar = 1 μm) [69]

5.5.6 Carbon Nanotube Scanning Near-Field Optical Microscopy

Another application of carbon nanotube AFM probes is in the scattering-type of scanning near-field optical microscopy (s-SNOM) [70–74]. In s-SNOM, a laser beam is focused on the apex of a sharp tip positioned nanometers above the sample. The electromagnetic field is enhanced at the tip generating a near-field interaction between the tip and sample [75, 76]. Light is scattered from this near-field interaction and is collected, amplified, and filtered at the far-field forming a single pixel of the sample. Concurrently at the same position,

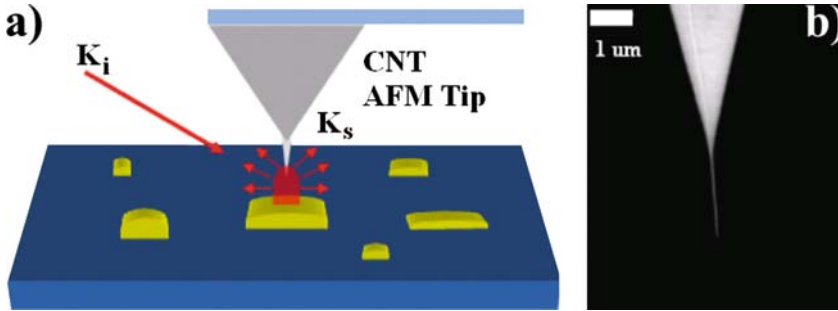


Fig. 5.27. Principle of the scattering-type of near-field CNT scanning optical microscopy. The incident laser beam, K_i , is focused on the apex of the carbon nanotube and the sample. The enhanced near-field at the tip interacts with the sample and K_s is scattered from this region. At the far-field, K_s is extracted from the background scattering forming a single pixel

the AFM image pixel is also collected. To form the complete optical and AFM image, the sample is raster scanned, see Fig. 5.27. The near-field is localized and enhanced at the tip and decays rapidly as r^{-6} , where r is the radius of the tip [77]. Simulation showed that the lateral resolution of metallic tips is of the order of the radius of the tip in air and in water [78]. Wavelengths in the radio waves microwave [79], infrared [80–86], visible [87–92], and X-ray have been used. Because the image is formed from the scattered light, this version of SNOM is known as the scattering-type SNOM (s-SNOM) [74]. A resolution of ~ 10 nm has been attained in s-SNOM [93].

Hillenbrand et al. [94] were the first to use a carbon nanotube AFM probe to image plasmon fields of gold nanostructures using s-SNOM. The microscope is shown in Fig. 5.28 [95]. A 633-nm laser polarized vertically is focused with an aspheric lens onto the tip of the 4- μm long nanotube bundle at an angle 30° from the sample surface. Back-scattered light is collected with the same lens. The AFM tip is operated in tapping mode at the resonance frequency $\Omega = 330\text{kHz}$. The reference frequency to the phase detector is $\omega + \Delta$, where ω is the laser frequency and $\Delta = 80\text{MHz}$. The phase detector is operated at $\Delta + n\Omega$ to extract the amplitude and phase of the scattered light at the harmonics of the tapping-mode frequency. Simultaneous AFM and optical images taken with the CNT optical probe is shown in Fig 5.29.

5.5.7 Biological Applications of Carbon Nanotube Atomic Force Microscopy

The same CNT structural and mechanical properties – ductility, high aspect ratio, small radius of curvature, and vertical alignment – that offer myriad advantages toward the specialized CNT AFM applications discussed above also bestow advantages toward biological imaging and injection. Namely, CNT AFM tips are robust and durable probes that provide high-resolution

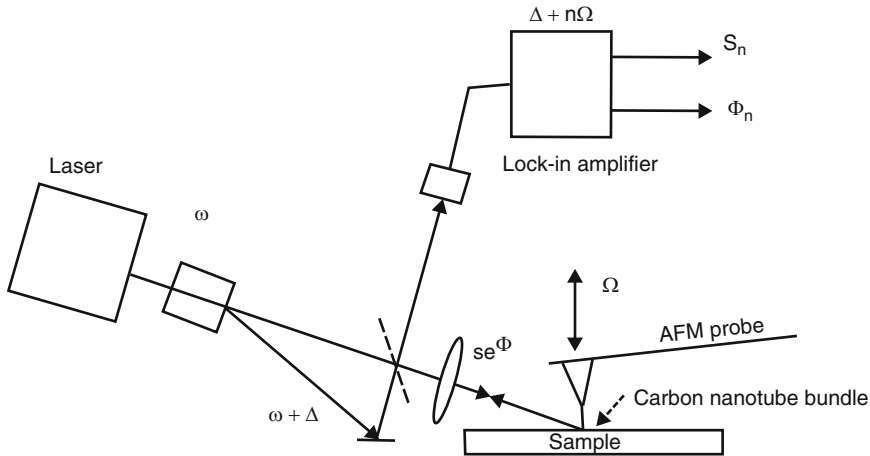


Fig. 5.28. Schematic representation of an s-SNOM microscope coupled with an AFM with a CNT tip (from [95])

images of delicate three-dimensional, contour, crevice, and cavity rich biological samples without damage; furthermore, CNT-based needles combined with the AFM's nanoscale positioning system allow for precise targeting of the probe [96–98]. Current biological applications involve aggressive imaging applications of nanoscale biological macromolecules and structures, drug delivery via CNT tip injection, and force-spectroscopy on cell membrane. (Fig. 5.29)

Based on results from imaging experiments performed on DNA and protein, CNT AFM bioprobes offer the following advantages: (1) high resolution, (2) imaging free from tip artifacts, and (3) reduced tip wear. A 300-nm long, 4–8-nm diameter CNT tip fabricated under SEM control reliably and repeatedly resolved the 3.4-nm interval spacing between helical turns in right-handed B-DNA [44]. Compared to a conventional silicon tip, MWNT of 140-nm length, 11-nm diameter, 3.5-nm radius of curvature prepared using the same technique was shown to provide better resolution of colon bacteria *E. Coli* pUC plasmid imaged in air (Fig. 5.30a,b) [8]. Imaging experiments on Lambda-DNA performed with SWNT tips attached using acrylic adhesive also showed improved resolution compared to silicon tips [6]. SWNT tips with 2.4-nm and 5.8-nm tip diameters have also been shown to reliably image DNA in liquid environments [51]. SWNT tips mounted using the pick-up technique are able to resolve hydrated RNA polymerase DNA (RNAP–DNA) complexes [28]. Direct CVD grown SWNT tips have also successfully imaged IgM protein molecules (Fig. 5.30c) [49].

Imaging experiments comparing silicon tips with attached MWNT tips demonstrate that images generated using MWNT tips are free from tip artifacts because of the MWNT high aspect ratio. Tip artifacts are present in

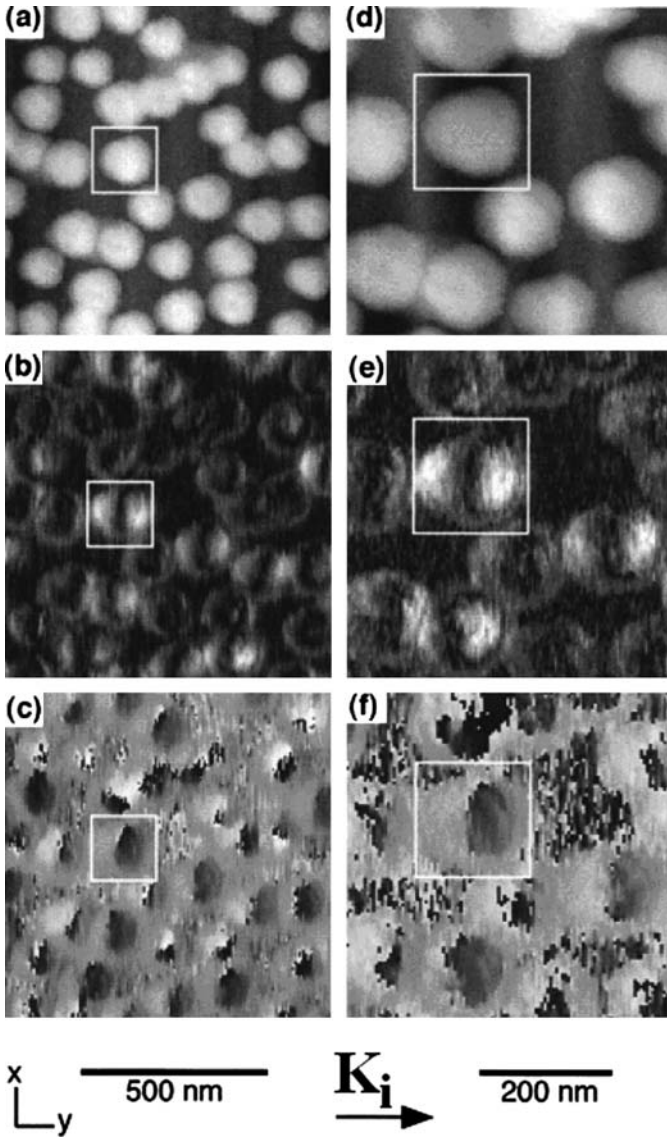


Fig. 5.29. Images of Au islands on glass taken with the CNT probe. (a) Topography (25 nm full scale) with (b) simultaneously recorded optical amplitude E_2 and (c) optical phase φ_2 (360° full scale); (d–f) repeat with reduced scan area showing (d) topography with (e) simultaneously recorded optical amplitude E_3 and (f) optical phase φ_3 . k_i marks the direction of the incident light (from [94])

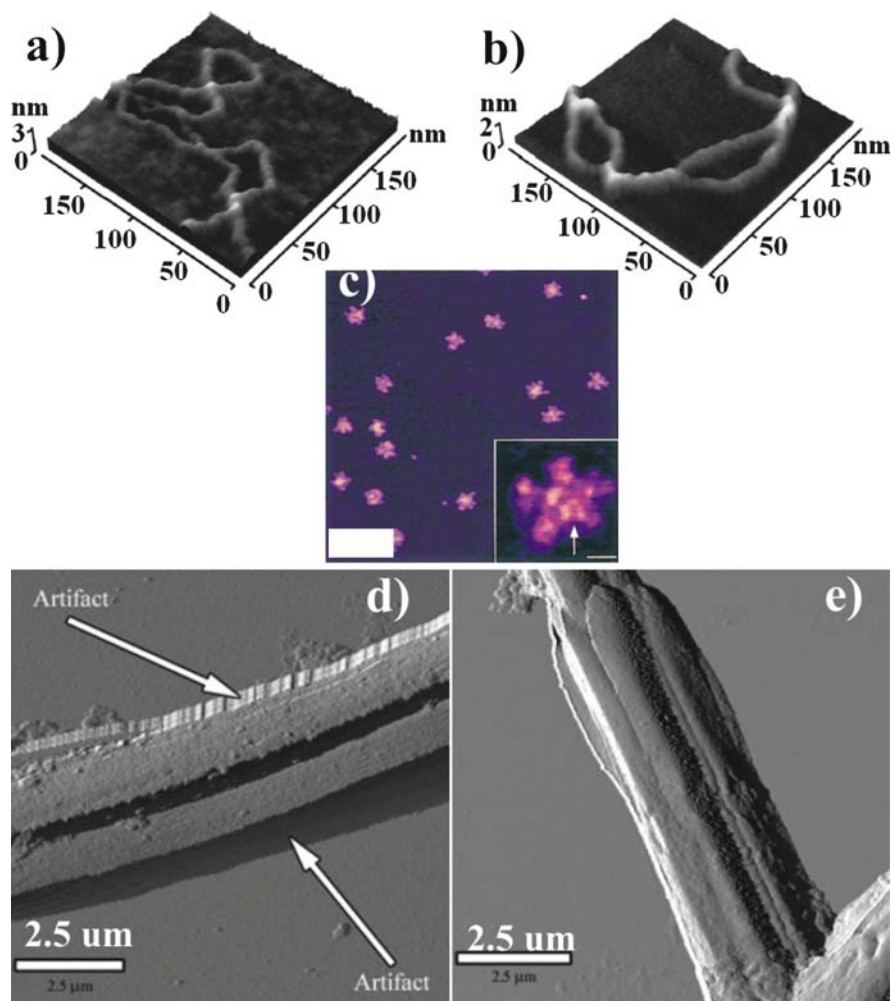


Fig. 5.30. *E. Coli* colon bacteria, pUC AFM image taken with (a) MWNT tip and (b) silicon tip [8]. (c) IgM protein image taken with SWNT tips (scale bars 100, inset: 10 nm) [49]. Image of *T. Aarantia* protein structure taken with (d) silicon tip and (e) MWNT tip [27]

silicon tips due to the pyramidal shape causing the tip sides to contact the sample before the tip apex (Fig. 5.30d, e) [27]. CNT probes are also prone to imaging artifacts but this problem can be alleviated by using shorter tips. An arc-discharge grown MWNT tip was also used to image DNA. Following attachment under SEM control, the MWNT was shortened by scanning a focused ion beam perpendicular to the MWNT axis and sharpened by applying a 2 V DC voltage across the MWNT producing a 1.8-nm tip radius [30]. Long tips (>100 nm) produce sharp images of DNA accompanied by a ghost

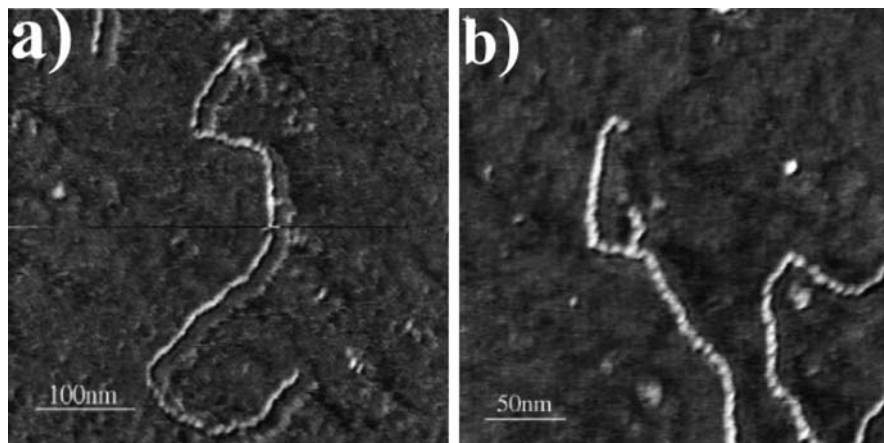


Fig. 5.31. Noncontact mode AFM image of DNA taken with (a) 300-nm MWNT tip and (b) 50-nm MWNT tip [30]

image artifact due to MWNT bending; however, images taken with short tips (≈ 50 nm) are not corrupted by a ghost image and resolve the interval spacing of 3.4 nm between the major grooves in DNA (Fig. 5.31) [30]. SWNT tips fabricated using a direct growth fabrication process have reliably imaged Staphylococcus Protein A adsorbed to SWNT side walls and show no appreciable degradation of resolving power due to tip wear [38].

A CNT-based nanoinjection system based on CNT AFM tips enables high-resolution drug delivery in that the nanoscale diameter CNT can probe spatially confined regions of cells and the AFM nanoscale manipulation system can precisely position the probe. Starting with an MWNT AFM tip fabricated under SEM control, compound 1 is absorbed on the MWNT sidewalls during coincubation with methanol and is then loaded with streptavidin-coated quantum dots in phosphate buffer (Fig. 5.32) [97]. Compound 1 consists of a CNT-binding pyrene moiety conjugated to a biotin molecule via a disulfide bond that can be cleaved in the cell's cytosol. Nanoscale injection proceeds as follows: (1) using the AFM positioning system, the nanoinjector is placed over a target cell and allowed to approach the cell surface, (2) upon contact with the cell membrane, the MWNT injector is lowered further to pierce the cell membrane, (3) once inside the cell, the reducing environment in the cytosol severs the disulfide bond releasing the quantum dot-loaded streptavidin into the cell, and (4) the MWNT injection system is withdrawn (Fig. 5.33) [97]. Quantum dots fluoresce and this intracellular signal can be captured by fluorescence microscopy. Testing this mechanism on HeLa cell cultures indicates that the probe must be allowed to remain inside the cell for 15–30 min in order to allow time for disulfide bond cleavage. After QD delivery, the injected cell fluoresces indicating successful delivery; furthermore, viability and membrane integrity assays performed after injection indicate that, for up to 10 h, the

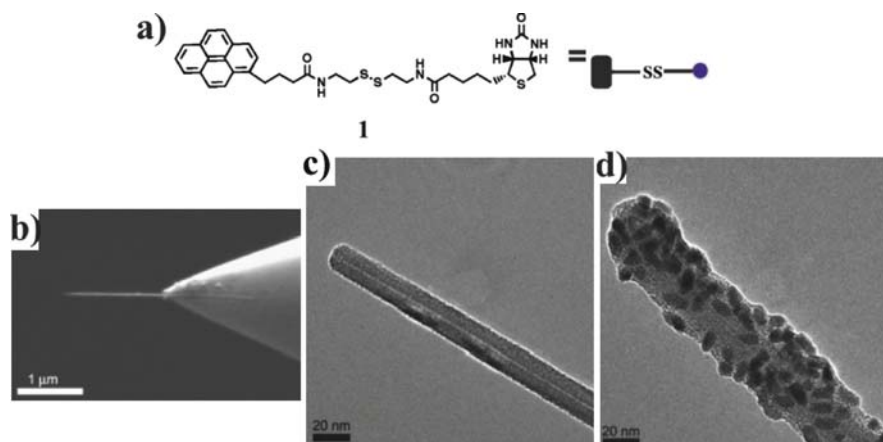


Fig. 5.32. (a) Compound 1 which provides a disulfide bond, (b) SEM image of MWNT attached to supporting tip. TEM images of (c) MWNT tip, (d) and MWNT after coating of linker 1 and conjugation with streptavidin [71]

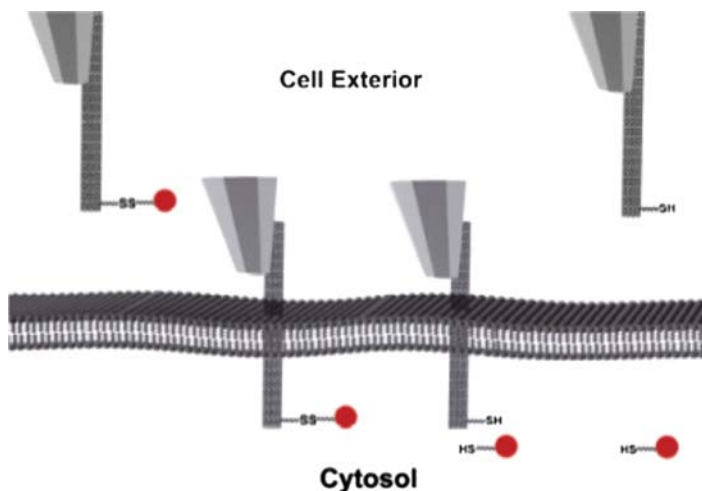


Fig. 5.33. Mechanism for drug injection utilizing a chemically functionalized CNT AFM tip [71]

cells remain viable and the membrane remains intact; thus, the CNT nanoinjection system can deliver drugs into the cytosol without damaging the cell (Fig. 5.34) [97].

A proposed alternative arrangement for CNT-based injection involves a gold-coated MWNT AFM tip conjugated via gold–thiol bonds to a self-assembled monolayer bound to a chemical agent via pH-dependent covalent bonds or intramolecular forces [98]. The gold coating is needed to stiffen the CNT allowing the rigid nanoneedle to penetrate the cell membrane.

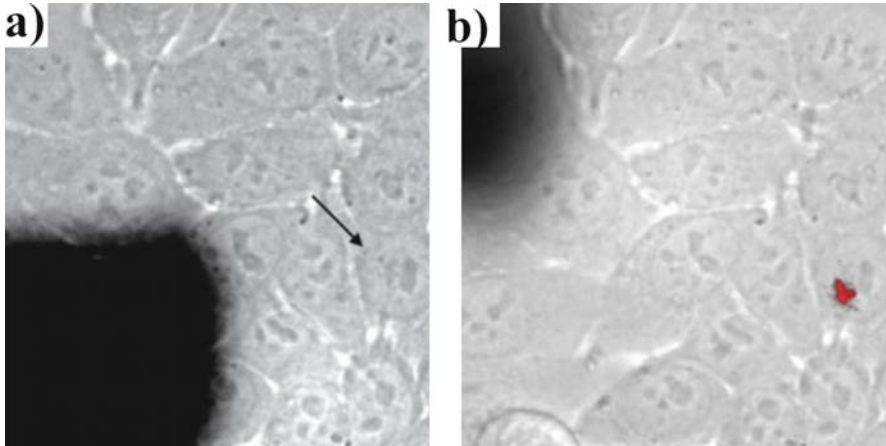


Fig. 5.34. Fluorescence microscopy image of HeLa cells (a) before injection and (b) after injection [71]

Force spectroscopy on human pleural mesothelial cells indicate that 10–20-nm diameter, 100–200-nm long MWNT tips coated with carbon via ion beam deposition are able to penetrate the cell with forces as low as 0.1–0.2nN which is an order of magnitude lower than forces required by etched silicon nanoneedles [98].

References

1. M. Di Ventra, S. Evoy, J.R. Hefflin, Jr., *Introduction to Nanoscale Science and Technology* (Kluwer, Norwell, MA, 2004)
2. V. Barwich, M. Bammerlin, A. Baratoff, R. Bennewitz, M. Guggisberg, C. Loppacher, O. Pfeiffer, E. Meyer, H.J. Guntherodt, J.P. Salvetat, J.M. Bonard, L. Forro, *Appl. Surf. Sci.* **157**, 269–273 (2000)
3. M. Kageshima, H. Jenseusius, M. Dienwiebel, Y. Nakayama, H. Tokumoto, S.P. Jarvis, T.H. Oosterkamp, *Appl. Surf. Sci.* **188**, 440–444 (2002)
4. H. Dai, N. Franklin, J. Han, *APL* **73**, 1508–1510 (1998)
5. G. Nagy, M. Levy, R. Scarmozzino, R.M. Osgood, Jr., H. Dai, R.E. Smalley, C.A. Michaels, G.W. Flynn, G.F. McLane, *APL* **73**, 529–531 (1998)
6. S.S. Wong, A.T. Woolley, T.W. Odom, J. Huang, P. Kim, D.V. Vezenov, C.M. Lieber, *APL* **73**, 3465–3467 (1998)
7. S.B. Arnason, A.G. Rinzler, Q. Hudspeth, A.F. Hebard, *APL* **75**, 2842–2844 (1999)
8. H. Nishijima, S. Kamo, S. Akita, Y. Nakayama, K.I. Hohmura, S.H. Yoshimura, K. Takeyasu, *APL* **74**, 4061–4063 (1999)
9. R. Stevens, C.V. Nguyen, A.M. Cassell, L. Delzeit, M. Meyyappan, J. Han, *APL* **77**, 3453–3455 (2000)
10. A. Hall, W.G. Matthews, R. Superfine, M.R. Falvo, S. Washburn, *APL* **82**, 2506–2508 (2003)

11. J. Hafner, C.L. Cheung, T.H. Oosterkamp, C.M. Lieber, *J. Phys. Chem. B* **105**, 743–746 (2001)
12. L.A. Wade, I.R. Shapiro, Z. Ma, S.R. Quake, C.P. Collier, *Nano Lett.* **4**, 725–731 (2004)
13. N.R. Wilson, J.V. Macpherson, *J. Appl. Phys.* **96**, 3565–3567 (2004)
14. F. Wolny, U. Weissker, T. Muhl, A. Leonhardt, S. Menzel, A. Winkler, B. Buchner, *J. Appl. Phys.* **99**, 4304–4309 (2006)
15. N. Choi, T. Uchihashi, H. Nishijima, T. Ishida, W. Mizutani, S. Akita, Y. Nakayama, M. Ishikawa, H. Tokumoto, *Japan. J. Appl. Phys.* **39**, 3707–3710 (2000)
16. M. Ishikawa, M. Yoshimura, K. Ueda, *Japan. J. Appl. Phys.* **41**, 4908–4910 (2002)
17. N.R. Wilson, D.H. Cobden, J.V. Macpherson, *J. Phys. Chem. B* **106**, 13102–13105 (2002)
18. I.R. Shapiro, S.D. Solares, M.J. Esplandiu, L.A. Wade, W.A. Goddard, C.P. Collier, *J. Phys. Chem. B* **108**, 13613–13618 (2004)
19. S.D. Solares, M.J. Esplandiu, W.A. Goddard, C.P. Collier, *J. Phys. Chem. B* **109**, 11493–11500 (2005)
20. T. Arie, H. Nishijima, S. Akita, Y. Nakayama, *J. Vac. Sci. Technol. B*, **18**, 104–106 (2000)
21. Y. Nakayama, H. Nishijima, S. Akita, K.I. Hohmura, S.H. Yoshimura, K. Takeyasu, *J. Vac. Sci. Technol. B*, **18**, 661–664 (2000)
22. S. Carnally, K. Barrow, M.R. Alexander, C.J. Hayes, S. Stolnik, S.J.B. Tendler, P.M. Williams, C.J. Roberts, *Langmuir* **23**, 3906–3911 (2007)
23. M. Kuwahara, H. Abe, H. Tokumoto, T. Shima, J. Tominaga, H. Fukada, *Mat. Char.* **52**, 43–48 (2004)
24. M.J. Esplandiu, V.G. Bittner, K.P. Giapis, C.P. Collier, *Nano Lett.* **4**, 1873–1879 (2004)
25. L.A. Wade, I.R. Shapiro, Z. Ma, S.R. Quake, C.P. Collier, *Nano Lett.* **4**, 725–731 (2004)
26. J. Tang, G. Yang, Q. Zhang, A. Parhat, B. Maynor, J. Liu, L.C. Qin, O. Zhou, *Nano Lett.* **5**, 11–14 (2005)
27. R.M.D. Stevens, N.A. Frederick, B.L. Smith, D.E. Morse, G.D. Stucky, P.K. Hansma, *Nanotech.* **11**, 1–5 (2000)
28. J.S. Bunch, T.N. Rhodin, P.L. McEuen, *Nanotech.* **15**, S76–S78 (2004)
29. S.I. Lee, S.W. Howell, A. Raman, R. Reifengerger, C.V. Nguyen, M. Meyyappan, *Nanotech.* **15**, 416–421 (2004)
30. J. Martinez, T.D. Yuzvinsky, A.M. Fennimore, A. Zettl, R. Garcia, C. Bustamante, *Nanotech.* **16**, 2493–2496 (2005)
31. T. Uchihashi, M. Higgins, Y. Nakayama, J.E. Sader, S.P. Jarvis, *Nanotech.* **16**, S49–S53 (2005)
32. K. Carlson, K.N. Anderson, V. Eichhorn, D.H. Peterson, K. Molhave, I.Y.Y. Bu, K.B.K. Teo, W.I. Milne, S. Fatikow, P. Boggild, *Nanotech.* **18**, 345501–345507 (2007)
33. C. Bernard, S. Marsaudon, R. Boisgard, J.P. Aime, *Nanotech.* **18**, 345501–345507 (2007)
34. I.S. Nandhakumar, T.J. Gordon-Smith, G.S. Attard, D.C. Smith, *Small* **1**, 406–408 (2005)
35. Q.M. Hudspeth, K.P. Nagle, Y.P. Zhao, T. Karabacak, C.V. Nguyen, M. Meyyappan, G.C. Wang, T.M. Lu, *Surf. Sci.* **515**, 453–461 (2002)

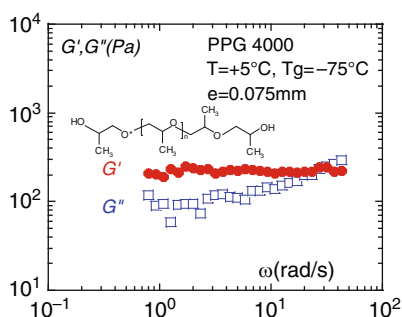
36. C.T. Gibson, S. Carnally, C.J. Clive, *Ultramicroscopy* **107**, 1118–1122 (2007)
37. K. Moloni, M.R. Buss, R.P. Andres, *Ultramicroscopy* **80**, 237–246 (1999)
38. E. Yenilmez, Q. Wang, R.J. Chen, D. Wang, H. Dai, *APL* **80**, 2225–2227 (2002)
39. S. Kuwahara, S. Akita, M. Shirakahara, T. Sugai, Y. Nakayama, H. Shinohara, *Chem. Phys. Lett.* **429**, 581–585 (2006)
40. M. Tanemura, M. Kitazawa, J. Tanaka, T. Okita, R. Ohta, L. Miao, S. Tanemura, *Japan. J. Appl. Phys.* **45**, 2004–2008 (2006)
41. C.V. Nguyen, Q. Yi, M. Meyyappan, *Meas. Sci. Tech.* **16**, 2138–2146 (2005)
42. Q. Yi, A.M. Cassell, H. Liu, K.K. Chao, J. Han, M. Meyyappan, *Nano Lett.* **4**, 1301–1308 (2004)
43. C.L. Cheung, J.H. Hafner, T.W. Odom, K. Kim, C.M. Lieber, *APL* **76**, 3136–3138 (2000)
44. T. Uchihashi, N. Choi, M. Tanigawa, M. Ashino, Y. Sugawara, H. Nishijima, S. Akita, Y. Nakayama, M. Ishikawa, H. Tokumoto, K. Yokoyama, S. Morita, M. Ishikawa, *Japan. J. Appl. Phys.* **39**, L887–L889 (2000)
45. S. Akita, H. Nishijima, Y. Nakayama, *J. Phys. D.* **33**, 2673–2677 (2000)
46. M. Ishikawa, M. Yoshimura, K. Ueda, *Phys. B.* **323**, 184–186 (2002)
47. H. Azehara, Y. Kasanuma, K. Ide, K. Hidaka, H. Tokumoto, *Japan. J. Appl. Phys.* **47**, 3594–3599 (2008)
48. T. Larsen, K. Moloni, F. Flack, M.A. Eriksson, M.G. Lagally, C.T. Black, *APL* **80**, 1996–1998 (2002)
49. J.H. Hafner, C.L. Cheung, C.M. Lieber, *Nature* **398**, 761–762 (1999)
50. S. Kleckley, G.Y. Chai, D. Zhou, R. Vanfleet, L. Chow, *Carbon* **41**, C01–836 (2003)
51. L. Chen, C.L. Cheung, P.D. Ashby, C.M. Lieber, *Nano Lett.* **4**, 1725–1731 (2004)
52. J. Tong, Y. Sun, *IEEE Trans. Nanotech.* **6**, 519–523 (2007)
53. A.J. Austin, Q. Ngo, C.V. Nguyen, *J. Appl. Phys.* **99**, 4304–4309 (2006)
54. C.V. Nguyen, K.J. Chao, R.M.D. Stevens, L. Delzeit, A.M. Cassell, J. Han, M. Meyyappan, *Nanotech.* **12**, 363–367 (2001)
55. D. Dietzel, S. Marsaudon, J.P. Aime, C.V. Nguyen, G. Couturier, *Phys. Rev. B.* **72**, 1–16 (2005)
56. E.S. Snow, P.M. Campbell, J.P. Novak, *APL* **80**, 2002–2004 (2002)
57. C.V. Nguyen, R.M.D. Stevens, J. Barber, J. Han, M. Meyyappan, M.I. Sanchez, C. Larsen, W.D. Hinsberg, *APL* **80**, 801–803 (2002)
58. M.C. Strus, A. Raman, C.S. Han, C.V. Nguyen, *Nanotech* **16**, 2482–2492 (2005)
59. M. Yasutake, Y. Shirakawa, T. Okawa, S. Mizooka, Y. Nakayama, *Ultramicroscopy* **91**, 57–62 (2002)
60. J. Goldstein, D.E. Newbury, D.C. Joy, C.E. Lyman, P. Echlin, E. Lifshin, L.C. Sawyer, L.C. Michael, *Scanning Electron Microscopy and X-Ray Microanalysis* 3rd ed. (Springer, New York, NY, 2003)
61. D. Dietzel, M. Faucher, A. Iaia, J.P. Aime, S. Marsaudon, A.M. Bonnot, V. Bouchiat, G. Couturier, *Nanotech* **16**, S73–S78 (2005)
62. S.S. Wong, A.T. Woolley, E. Joselevich, C.L. Cheung, C.M. Lieber, *J. Amer. Chem. Soc.* **120**, 8557–8558 (1998)
63. S.S. Wong, A.T. Woolley, E. Joselevich, C.L. Cheung, C.M. Lieber, *Chem. Phys. Lett.* **306**, 219–225 (1999)
64. M. Ishikawa, M. Yoshimura, K. Ueda, *Appl. Surf. Sci.* **188**, 456–459 (2002)
65. H. Dai, J.H. Hafner, A.G. Rinzler, D.T. Colbert, R.E. Smalley, *Nature* **384**, 147–150 (1996)

66. T. Ikuno, M. Katayama, M. Kishida, K. Kamada, Y. Murata, T. Yasuda, S. Honda, J.G. Lee, H. Mori, K. Oura, *JJAP* **43**, L644–L646 (2004)
67. T. Shimizu, H. Tokumoto, S. Akita, Y. Nakayama, *Surf. Sci.* **486**, L455–L460 (2001)
68. W. Mizutani, N. Choi, t. Uchihashi, H. Tokumoto, *JJAP* **43**, L644–L646 (2004)
69. H. Kuramochi, H. Akinaga, Y. Semba, M. Kijima, T. Uzumaki, M. Yasutake, A. Tanaka, H. Yokoyama, *JJAP* **44**, 2077–2080 (2005)
70. F. Zenhausern, M.P. O’Boyle, H.K. Wickramasinghe, *Applied Physics Letters* **65**(13), 1623–1625 (1994)
71. Y. Inouye, S. Kawata, *Optics Letters* **19**(3), 159–161 (1994)
72. D. Courjon, C. Bainier, *Rep. Prog. Phys.* **57**, 989 (1994)
73. L. Novotny, R.X. Bian, X.S. Xie, *Phys. Rev. Lett.* **79**, 645 (1997)
74. F. Keilmann, R. Hillenbrand, *Physical and Engineering Sciences* **362**(1817), 787–805 (2004)
75. L. Novotny, S. Stranick, *Annu. Rev. Phys. Chem.* **57**, 303 (2006)
76. A. Downes, D. Salter, A. Elfick, *J. Phys. Chem. B* **110**, 6692–6698 (2006)
77. Z. Ma, J.M. Gerton, L.A. Wade, S.R. Quake, *Phys. Rev. Lett.* **97**, 260801, (2006)
78. A. Downes, D. Salter, A. Elfick, *J. Microscopy*, **229**(2), 184–188 (2008)
79. B. Knoll, F. Keilmann, A. Kramer, R. Guckenberger, *Appl. Phys. Lett.* **70**, 2667–9 (1997)
80. B. Knoll, F. Keilmann, *Appl. Phys. Lett.* **77**, 3980–3982 (2000)
81. A.J. Huber, D. Kazantsev, F. Keilmann, J. Wittborn, R. Hillenbrand, *Adv. Mater.* **19**, 2209–2213 (2007)
82. R. Hillenbrand, T. Taubner, F. Keilmann, *Nature* **418**, 159–162 (2002)
83. T. Taubner, F. Keilmann, R. Hillenbrand, *Nano Lett.* **4**, 1669–1672 (2004)
84. S. Schneider, J. Seidel, S. Grafström, L.M. Eng, S. Winnerl, D. Stehr, H. Helm, *Appl. Phys. Lett.* **90**, 143101–143101 (2007)
85. T. Taubner, R. Hillenbrand, F. Keilmann, *Appl. Phys. Lett.* **85**, 5064–5066 (2004)
86. M. Brehm, T. Taubner, R. Hillenbrand, F. Keilmann, *Nano Lett.* **6**, 1307–1310 (2006)
87. R. Hillenbrand, F. Keilmann, *Phys. Rev. Lett.* **85**, 3029–3032 (2000)
88. R. Hillenbrand, F. Keilmann, *Appl. Phys. Lett.* **80**, 25–27 (2002)
89. R. Bachelot, G. Lerondel, S. Blaize, S. Aubert, A. Bruyant, P. Royer, *Microsc. Res. Tech.* **64**, 441–452 (2004)
90. N. Anderson, A. Bouhelier, L. Novotny, *J. Opt. A* **8**, S227–S233 (2006)
91. Z.H. Kim, S.R. Leone, *J. Phys. Chem. B* **110**, 19804–19809 (2006)
92. Y.C. Martin, H.K. Wickramasinghe, *J. Appl. Phys.* **91**, 3363–8 (2002)
93. R. Hillenbrand, F. Keilmann, *Appl. Phys. Lett.* **80**, 25 (2002)
94. R. Hillenbrand, F. Keilmann, P. Hanarp, D.S. Sutherland, *Appl. Phys. Lett.* **83**, 368–370 (2003)
95. R. Hillenbrand, F. Keilmann, *Phys. Rev. Lett.* **85**(14), 3029–3032 (2000)
96. N. Sinha, J.T.W. Yeow, *IEEE Trans. Nanobiosci.* **4**, 180–195 (2005)
97. X. Chen, A. Kis, A. Zettl, C.R. Bertozzi, *PNAS* **104**, 8218–8222 (2007)
98. I.U. Vakarelski, S.C. Brown, K. Higashitani, B.M. Moudgil, *Langmuir* **23**, 10893–10896 (2007)

Novel Strategies to Probe the Fluid Properties and Revealing its Hidden Elasticity

Laurence Noirez

Summary.



Solid-like response of the propyleneglycol at 80°C above T_g and $75\ \mu\text{m}$ gap thickness.

Key words: Dimensionality, Elasticity, Liquid state, Solid-like correlations.

This chapter reports on the measurement of a non-vanishing shear elasticity in the liquid state of various fluids, away from phase transition temperature and at a macroscopic length scale. A series of experimental observations carried by different groups, using different techniques, on various samples (simple liquids, unentangled polymers, entangled polymers, glass formers, etc.) has revealed very large solid-like length scale correlations in the liquid state. This elastic property is typically observed at the micron scale in simple liquids and at several tens of microns up to the millimeter scale in liquid polymers and glass formers. It has been measured by rheometry using drastic boundary conditions, and by grazing incidence synchrotron studies. All rheometry measurements point out the influence of the gap thickness on the elasticity value. An increase of the elastic character is observed by decreasing the gap thickness. Correlatively, the elastic character vanishes at high thickness.

The concept of finite shear elasticity in liquids is new and absent from the theoretical conventional considerations in both simple liquids and viscoelastic liquids. It contrasts with the classical viscoelastic results that foresees a continuously decrease of the shear elasticity towards low frequencies. The consideration of a finite shear elasticity sheds new light on various unpredictable flow instabilities and allows a better analysis and interpretation of nanoscopic, microscopic studies with respect to the true nature of the bulk properties.

6.1 Introduction

‘From the three states of matter, the liquid state is the last studied and the least understood’. Hildebrandt & Scott 1950.

More than a half century later, predicting the properties of a liquid from its constitutive molecular characteristics remains a challenge [1–4]. The liquid state is a condensed matter state that flows. But how is it flowing? van der Waals, polar, Keesom forces, and hydrogen bonding provide the necessary intermolecular attraction to ensure the cohesive nature of this condensed state. Local interactions in liquids are not different from the solid state. But unlike solids, the strong thermal agitation in liquids weakens the intermolecular interactions and provides their puzzling flowing ability. A realistic interpretation of the macroscopic rheological properties of fluids in terms of a microstructural molecular description is thus one of the major challenges of the last 50 years. The need of molecular-based models concerns as well ordinary liquids, as complex fluids, polymers, colloids, emulsions, foams and involves multidisciplinary approaches including wetting, friction, adhesion, micro to macrofluidics. It is the key that allows any prediction in the huge field covered by the analysis of the fluid properties, its behaviour and its applications.

To probe and describe the complexity of the fluid properties, the discipline of rheology has emerged since several decades as the technique of characterisation of the fluidic state. The elaboration of a protocol of measurement, in particular for the oscillatory experiments of dynamic relaxation became allowed by the adoption of the concept of no-slip boundary condition between the liquid and the wall. The question of the slip at the wall is as old as the first attempts of understanding of the flow mechanisms. The works of Maxwell (1866) played a major role in the story by claiming firmly the no-slip assumption. This assumption was accredited by the reproducibility of the viscosimetry values deduced on the basis of the no-slip boundary conditions. Nowadays, no-slip is usually accepted as the wall boundary condition for flows at scales above several molecular lengths. Finally, it is interesting to note that small strain oscillatory experiments (dynamic relaxation method) have never been suspected of violating the no-slip boundary condition [5, 6]. Dynamic relaxation became particularly famous in the frame of the phenomenological description of the linear viscoelastic behaviour of polymers remarkably described in analogy with mechanical models in terms

of springs and dashpots [5, 6]. However, this remarkable phenomenological approach also fails at explaining the continuous emergence of new and unpredictable phenomena. The conventional viscoelasticity is unable to predict basic observations such as why do some fluids climb stirrer rods, why do they swell at the exit of an extruder, why do they produce spectacular instabilities [e.g. the ‘shark-skin’ instability, gross melt fracture, Weissenberg effects, shear induced phase transitions (Fig. 6.1)] [7–10], why a slippage length (Navier length) is systematically introduced to model even the hydrodynamics of simple fluids [11, 12], why macrofluidic properties are incompatible with the micro- and nanoscopic scale [13–19]. This chapter draws attention to these unresolved

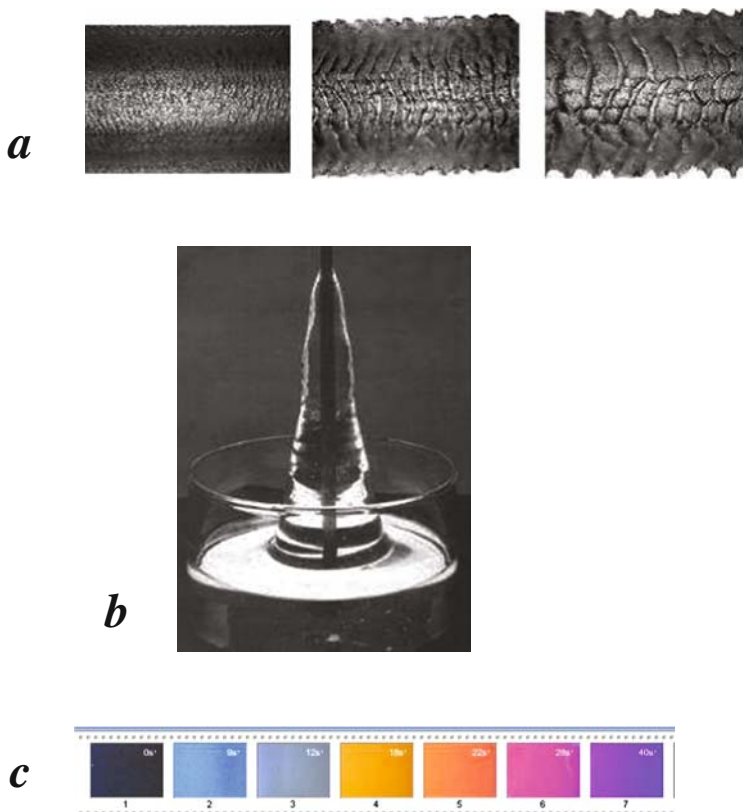


Fig. 6.1. Spectacular non-linear rheological behaviours: (a) Extruded polyethylene exhibiting the ‘stigmata’ due to the shark-skin instability. Controlling this interface instability is crucial for extraction processes. (b) Weissenberg effect. The liquid climbs on the rotating rod as a response to the shear stress. (c) Shear induced phase transition in the isotropic phase of liquid-crystal polymers. Above a critical shear rate, the isotropic liquid (*left photo*) transforms in a strongly birefringent oriented phase (*right photos*). The change of colors indicates the increase of the birefringence by increasing shear rate (photos between crossed polarisers)

topics, and to adopt a critical point of view on some fundamentals of the conventional descriptions. The pertinence of the usual protocol of determination of the viscoelastic parameters is questioned with respect to the boundary conditions, slippage transitions, and autophobicity effects. These parameters are known to have a significant impact on the measurement and the interpretation [20–27]. Emphasising the interest at complementing the conventional protocols of rheology by microscopic, wettability, and surfacic studies, this chapter highlights new strategies to probe the fluid dynamics. The conventional relaxation dynamic method that gives access to the so-called viscous and elastic moduli characterising the material is revisited considering in particular the impact of the no-slip boundary conditions on the measurement and on the interpretation of the results. In this frame, a series of experimental observations carried by different groups, using different techniques, has evidenced the identification of so far neglected huge solid-like length scale correlations in the liquid state far away from any phase transition. This solid-like character measures the cohesive character of the liquid, that is, its condensed state nature. This fundamental component is absent in conventional theories for viscoelastic liquids. The concept of finite shear elasticity in liquids contrasts with the theory of the viscoelasticity that foresees a continuously decrease of the shear elasticity towards low frequencies. The consideration of finite shear elasticity should open new routes for a better approach of the fluid dynamic and of its substrate interaction.

6.2 Basic Theoretical Considerations: Conciliating Simple Liquid Approach to the Viscoelasticity Theory?

6.2.1 Simple Liquid Description

The theoretical analysis of the fluid dynamic properties differs following the size of the molecules. Simple liquids are considered as governed by short-range interparticle interactions [1]. The ad-equation of the modelling of the structure factor displayed by liquids at the molecular scale, to pair correlation functions, has provided significant progresses in the description of the microstructure of liquids [4, 28, 29]. However, scattering methods (neutron scattering, X-ray scattering) give access to pair interaction only whereas a multi-body interaction description would certainly be preferable. At a macroscopic scale, the concept of cohesion illustrates the high degree of connectivity in liquids that forms a cohesive network; the cohesion is calculated on the basis of the energy of the local interactions (van der Waals, H-bonding, etc.), estimated from the sublimation energy, from the tensile strength of the fluid to negative pressure [30].

6.2.2 The Viscoelastic Approach

The analysis of liquids of high molecular weight systems (typically polymers in solution or in the bulk) is radically different. Polymers are usually described in light of molecular theory of the linear viscoelasticity initiated by Rouse to describe first diluted solutions of polymer (Fig. 6.2), concentrated polymer solutions, and then unentangled polymers, and by Doi-Edwards for the first descriptions of high molecular weight entangled polymers [31–35]. In these pioneering approaches, the interparticle interaction is neglected; the liquid state is supposed to be entirely dominated by the intramolecular forces due the intrachain connectivity in the case of unentangled polymers [33, 34, 36], or extended to first neighbour interactions in the case of a description in terms of the tube model [28, 29 and references therein]. Correlatively, the notion of macroscopic cohesion is absent. This concept of cohesion due to intermolecular interactions is absent even for low molecular (unentangled) polymers that do not fundamentally differ from the case of viscous liquids. There is thus a contradiction between the description of low molecular weight molecules proposed by the viscoelasticity and by the theory of liquids.

The viscoelastic theory is a phenomenological approach. Its success is explained by the formalisms of Rouse or of reptation, able to model the rheological responses displayed by a wide class of polymers in terms of individual molecular dynamic. Promoting the single molecular dynamic, it neglects intermolecular interactions [37, 38]. How this strong assumption has been justified? The origin of the viscoelasticity theory is very instructive. This approach was

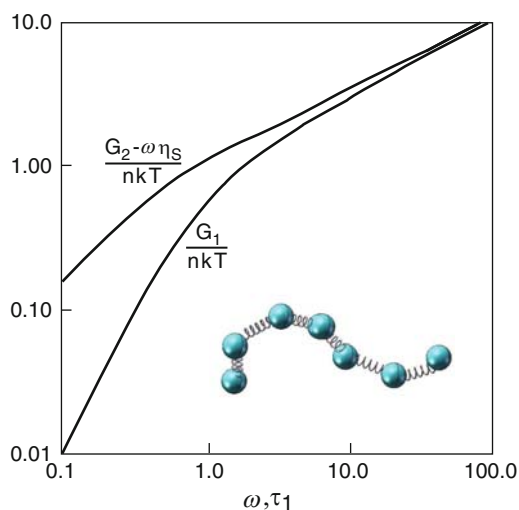


Fig. 6.2. The graph displays the viscoelastic curve from the original paper of Rouse [31]. This is the bead-spring model. This free-joined chain without interaction with its neighbours has been elaborated to model the viscoelastic properties of diluted polymer solution

initially elaborated to depict the viscoelastic behaviour of polymer chains in a diluted solution [39, 40]. In this frame, the properties of the diluted solution were interpreted as governed essentially by the chain properties; it is a single chain picture approach where all the relaxation modes (generating viscous and elastic contributions) originate from the intrachain properties. The chain is seen as a succession of independent elastic springs ('bead-spring' model, Fig. 6.2) [39–41]. Each spring representing a group of monomers (subchain) is sufficiently long to be decoupled from others (gaussian statistics).

This single molecular approach has been then extended to concentrated solutions and to melts. The justification of this assimilation is based on the experimental similarity of the experimental viscoelastic signatures exhibited by all of them [40]. Therefore, because of the single molecular dynamic assumption made in analogy with diluted solutions, the theory of viscoelasticity neglects, even for the melt state, intermolecular interactions and thus ignores the cohesive nature of the fluidic state. There is no fundamental difference between a low molecular weight (unentangled) polymer and a simple liquid from the point of view of the nature of the interactions. There is thus a contradiction between the theory of simple liquids and the viscoelasticity theory which neglects the possibility of intermolecular interactions.

After a short presentation of the usual procedure employed to determine the viscoelastic parameters of a material, we list a series of experimental observations that drive to identify extra long collective effects in the liquid state, and to conclude that this conventional procedure does not actually probe the dynamic signature of the individual molecules but that of a (huge) assembly of molecules. It comes out that the single molecule dynamic may not be the most adapted to describe the viscoelastic curves.

6.3 Conventional Procedure to Determine the Dynamic Properties of Fluids

Rheological testing involves most of the time, either a frequency scanning small-amplitude oscillatory shear, referred to as relaxation dynamic measurements or a steady-state shear measurements referred to as non-linear rheology. In both experiments, the material is placed in contact with and between two fixtures [disk-like or cylinders in the case of conventional rheometry (Fig. 6.3) or plates in the case of a piezorheometer], one oscillating or rotating, the other fixed one measuring the motion transmitted by the sample via a sensor.

These two principal types of measurements are referred to linear and non-linear rheology.

6.3.1 Linear Rheology

- Linear rheology applies a small amplitude oscillatory shear stress (dynamic relaxation measurements) to probe the equilibrium state of the material

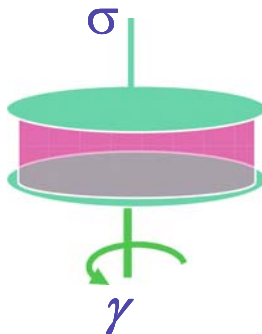


Fig. 6.3. Scheme of the measuring cell for conventional rheology. The sample is placed between two-disks and the interaction is ensured by simple contact to the substrates. The upper surface is fixed and the bottom surface is rotating (applying small oscillations or a steady-state rotation for the linear rheology and non-linear rheology, respectively)

[33–35]. The strain amplitude has to be sufficiently small to be considered as non-perturbative with respect to the equilibrium state of the material. This principle is referred to as the causality-linearity principle. These conditions are necessary to access to relevant physical parameters.

The causality means that the measured parameters (elastic and viscous moduli) are the consequence of the application of an external field (oscillatory strain stress).

The linearity indicates that the external field is so weak that the response function (torque) is only proportional to the applied field (the strain amplitude). In such conditions, the modulus which is the ratio of the torque to the strain does not depend on the strain amplitude within the regime of small strain amplitudes. When the response is no more proportional to the applied strain (entrance in the non-linear regime of viscoelasticity), the sample is no more in its equilibrium state (deformation, alignment, slipping, fracturing, etc.). However, a material is able to exhibit several linear regimes related to different physical mechanisms, as shown in the paragraph 8 of this chapter. Therefore, the identification of the first linear regime is not always trivial.

In the ideal case, the material (liquid or solid) should be probed in a non-dissipative way, that is, should not flow during a dynamic relaxation measurement and should respond linearly with the shear strain amplitude or measured in a strain range corresponding to a linear response (without being necessarily Newtonian, nor a fluid). From the difference between the input and the output signals, two parameters are extracted; the component in phase with the strain determines elastic (or storage) modulus (G'), whereas the out of phase component defines the viscous (or loss) modulus (G''). In controlled strain mode, a sinusoidally varying strain $\gamma = \gamma_0 \cdot \sin(\omega \cdot t)$ is imposed on the

sample. Here, ω is the frequency, t the time, δ the phase, and γ_0 is the strain amplitude. The shear stress in the sample follows:

$$\sigma = \sigma_0 \cdot \sin(\omega \cdot t + \delta) = |G^*| \cdot \gamma_0 \cdot \sin(\omega \cdot t + \delta),$$

where σ_0 is the stress amplitude and G^* is the magnitude of the complex modulus. Conventionally, $\sigma = G' \cdot \gamma_0 \cdot \sin(\omega \cdot t) + G'' \cdot \gamma_0 \cdot \cos(\omega \cdot t)$, where $G' = |G^*| \cdot \cos(\delta)$ and $G'' = |G^*| \cdot \sin(\delta)$.

Finally, it should be noted that a relationship (established by Kramers and Kronig for the treatment of optical signals [42–44]), links the elastic modulus (G') to the viscous modulus (G''). It allows the determination of one when the other one is known. However, this relationship requires the integration of the relaxation functions over the full frequency spectrum ($\omega \rightarrow \infty$) to be experimentally applicable.

6.3.2 Non-Linear Rheology

- Non-linear rheology consists typically in the observation of the response of a fluid submitted to a continuous shear stress versus shear rate; the material is forced to flow and becomes fluid, or enters into a slip regime if the sample presents strong cohesive forces. This method is clearly not indicated to probe the characteristics of gels, pastes, foams, solid-like assemblies, etc. Non-linear rheology allows the construction of the constitutive relationship between the shear stress, σ and the shear rate, $\dot{\gamma} = \frac{\delta\gamma}{\delta t}$: $\sigma = \eta(\dot{\gamma}) \cdot \dot{\gamma}$ where η is the shear viscosity. A Newtonian fluid is defined by a linear shear stress response versus shear: $\eta(\dot{\gamma}) = Cst$. The discontinuities in the constitutive curve indicate a coupling with characteristic relaxation times, τ of the material (life time of fluctuations, viscoelastic relaxation times). The flow coupling is expressed following the Deborah relationship: $\tau \cdot \dot{\gamma}_c \approx 1$, where $\dot{\gamma}_c$ is the critical shear rate at the discontinuity. Shear induced phase transitions and banded flow profiles are typically analysed on the basis of the shear stress versus shear rate constitutive relationship [45–49]. In these models, the liquid is supposed to separate within a shear rate interval, in an inhomogeneous flow submitted to different shear rates, but keeping the shear stress constant. Dissipative phenomena as partial slippage, fracturing, shear induced thermal effects (increase of temperature induced by friction) are irrelevant in the context of the above models. Geometry considerations are taken into account (gradient banding and vorticity banding), but gap thickness effects are irrelevant since the shear stress is supposed to be constant. Gap thickness effects are also irrelevant in the frame of conventional viscoelasticity.

Figure 6.4 displays the expected parallelism between the linear viscoelastic curve of polymer melts and their non-linear rheological behaviour.

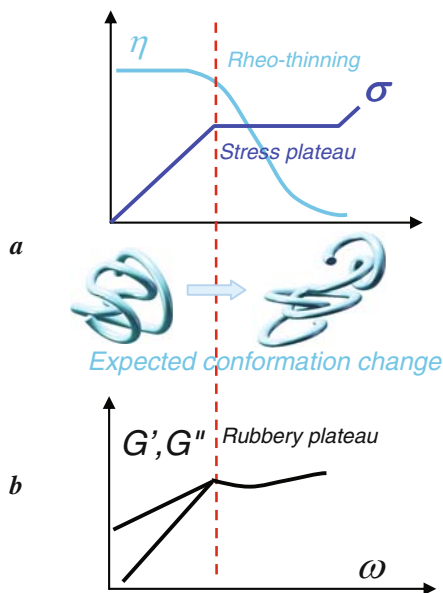


Fig. 6.4. Theoretical equivalence between flow curves and the viscoelastic spectrum. (a) Typical polymer flow curve. The shear viscosity (η) and the steady-state shear stress (σ) are represented as a function of the shear rate ($\dot{\gamma}$). (b) Typical evolution of the elastic (G') and viscous (G'') moduli as a function of the frequency (ω)

Now, how can an experimentalist establish conditions under which the experimental measurement reflects effectively the stress in the material? These methods are entirely dependent on the efficiency of the transmission of the shear stress from the surface to the sample (and from the sample to the sensor). This transmission is deeply related to the forces of the interaction between the material and the substrate. The following paragraph shows on a series of examples that the situation may not be so clear.

6.4 Unpredicted Phenomena and Unsolved Questions: Flow Instabilities, Non-Linearities, Shear Induced Transitions, Extra-Long Relaxation Times, Elasticity in the Liquid State

The success of the linear viscoelasticity to model the experimental curves of a broad category of complex fluids cannot hide a conceptual contradiction with the simple fluid description which is based on interparticle interactions (paragraph 2) [1–4]. Another body of viscoelastic materials, polymers [50, 51] but also miscible blends [52–54], breaks down the theoretical predictions disobeying the time–temperature superposition principle [33–35]. The situation

is worse considering the difficulty to conciliate non-linear characteristic relaxation times to the viscoelastic time scales. One example displayed by polymer melts showing this attempt of conciliation of viscoelastic relaxation times to the non-linear time scales issued from the shear stress versus shear rate curve is given in [55]. The non-linear rheology contains a long list of behaviours unpredictable on the basis of the analysis of the viscoelastic study. Still unsolved non-linear behaviours in polymers date from 1942, when birefringence experiments suggested an affineness loss in stretched entangled melts [56,57]. Similarly, over large time scale relaxations observed on entangled melts by neutron scattering and particle tracking [58, 59]. Later, despite considerable efforts, the problem of the excessive strain/shear softening in polymers still resists to the best predictions [60]. Recently, experimental approaches so different as NMR [61], ‘nanofishing’ [62–65], particle tracking velocimetry [66–68], still reveal the inadequacy with a description in terms of an individual molecular dynamic model and invite to explore the possibility of collective effects.

But, probably the most convincing observations in favour of the primacy of collective effects is a set of recent [69–72] and older converging experimental observations, indicating by dynamic relaxation at narrow geometry that solid-like length scale correlations govern the liquid dynamics. The first claim of identification of finite shear elasticity as a bulk property of liquids seems to be due to B. Derjaguin (1902–1994) [69,70].

In 1989, B. Derjaguin and collaborators published ‘liquids exhibit a low-frequency shear elasticity – a property unknown before – which is associated with collective interactions of liquid molecules’ (Fig. 6.5). This result was obtained at room temperature, on a series of ordinary liquids (water, hexadecane, cyclohexane, etc.) and glass formers (diethylene glycol) using a resonance method based on a piezoquartz crystal performing small shear oscillations with respect to the sample placed between this surface and a quartz prism. This elastic property has been identified at about a few microns scale and its

The complex shear modulus of polymeric and small-molecule liquids

B. V. Derjaguin, U. B. Bazarov, K. T. Zandanova and O. R. Budaev
Institute of Physical Chemistry of the USSR Academy of Sciences, Department of Surface Phenomena, 31 Lenin Prospect, Moscow 117915, USSR
 (Received 11 May 1988; accepted 7 June 1988)

The complex shear moduli of some ordinary liquids and their dependence on the shear deformation angle have been measured by the resonance method at a frequency of 73.5 kHz. From the results obtained, it is concluded that liquids exhibit a low-frequency shear elasticity—a property unknown before—which is associated with collective interactions of liquid molecules.

Fig. 6.5. First observation of finite shear elasticity reported in 1989 by B.V. Derjaguin and coworkers [Polymer **30** (1989) 97]

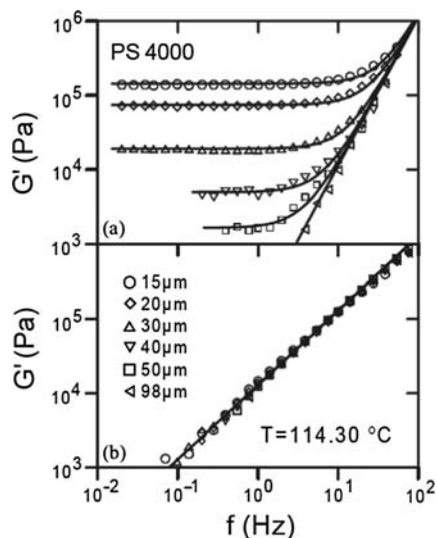


Fig. 6.6. Terminal elastic (gel) behaviour and its gap thickness dependence reported by D. Collin and coworker in 2003 on a low molecular weight polystyrene ($M_w = 4,000$) at 114.3°C , using a piezorheometer and treated glass substrates. The upper figure displays the evolution of the elastic modulus (G') versus frequency down to $15\ \mu\text{m}$. The lower figure displays the viscous modulus (G''). After Collin et al. [reprinted with permission of *Physica A* **320** (2003) 235]

modulus increases as the thickness decreases [69]. The authors attribute this low frequency elasticity to collective interactions of molecules.

Four years later (2003), Collin and Martinoty published that low molecular weight (unentangled) polystyrene melts display an anomalous behaviour when the gap of their piezorheometer became less than about $50\ \mu\text{m}$ [71]. They observed an elastic response at low thicknesses and low frequency (Fig. 6.6) at about 40°C above the glass transition. In contrast, the viscous modulus did not exhibit an anomalous behaviour and scales as the frequency; this gel behaviour was interpreted in terms of pretransitional glass transition effects.

Mendil et al. published in 2006 [72–74], the identification of a solid-like response (both elastic and viscous moduli are constant with a dominant elastic part) in unentangled polybutylacrylates, at about 100°C above the glass transition in low gap geometry (from 20 up to $100\ \mu\text{m}$) using a commercial rheometer (Fig. 6.7). This macroscopic measurement was reproduced by S.Q. Wang in 2007 [75] using the protocol of the former authors. This protocol has been also applied to reveal the solid-like component in glass formers (H-bond liquids) and entangled polymers [76].

Concerning glass formers, Orrit et al. [77] declare that they were able to access by rheological measurements at very weak stresses, to the solid-like network nature of glycerol and ortho-phenyl, up to $30\ \text{K}$ above the glass transition (Fig. 6.8). They also explain that this result would not be observable

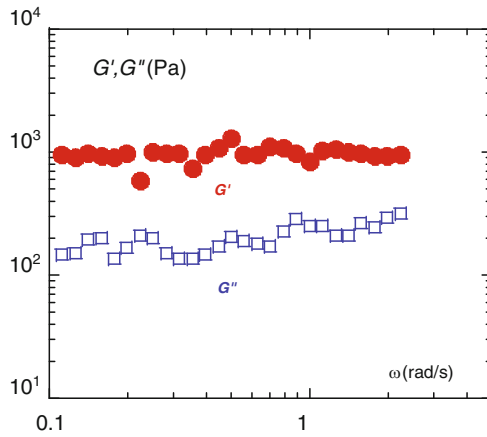


Fig. 6.7. Observation of a terminal solid-like response displayed by an unentangled polymer at 90°C above T_g with a conventional rheometer equipped with alumina substrates (full wetting boundary condition, plate-plate geometry, gap thickness 0.020 mm). After Mendil et al. [reprinted with permission of EPJ E **19** (2006) 84]

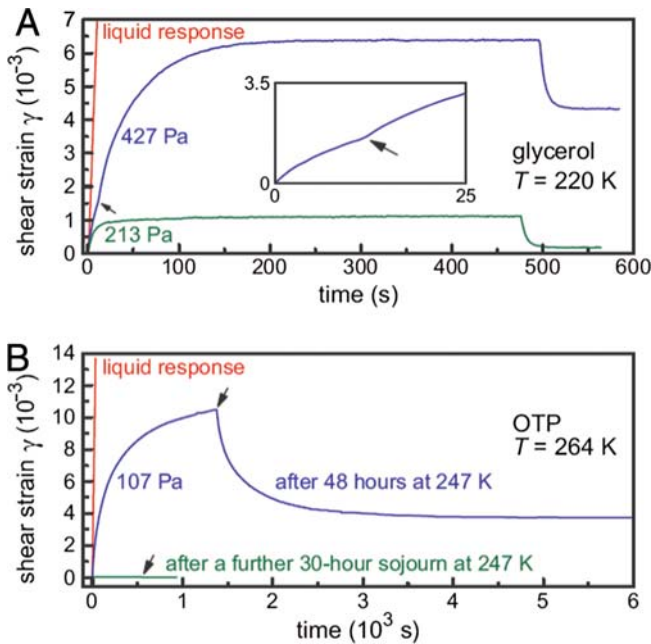


Fig. 6.8. Observation of a finite shear strain response (solid-like response) in the relaxation spectrum at a constant shear stress in glass formers (*upper figure*: glycerol and *lower figure*: *ortho*-phenyl) above 7 and 3°C above the glass transition. After Orrit et al. [reprinted with permission of PNAS **105** (2008) 4993]

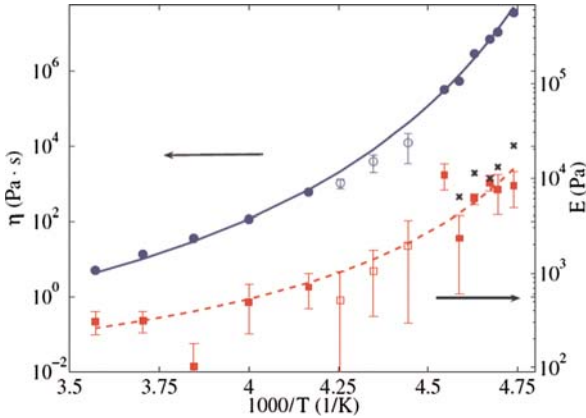


Fig. 6.9. Determination and variation versus temperature of an elastic component (*filled square*) extracted from grazing incidence synchrotron radiation. The circles show the viscosity. After Y. Chushkin et al. [reprinted with permission of EPL **83** (2008) 36001]

by conventional rheology, because the large stresses employed produce the melt of the solid-like structure.

Finally, besides rheology experiments, an interesting grazing incidence synchrotron radiation study [78] has also revealed the need to include a low frequency elasticity (Fig. 6.9), to describe adequately the surface dynamic of capillary waves displayed by a polypropyleneglycol far above the glass transition.

6.5 From Macro to Micro and Nanofluidics

The preceding results show that a finite low frequency shear elasticity is observable at the sub-millimetre scale in the liquid state. It might be expected that the same property should dominate at even lower length scales. Experiments relevant to the present discussion are measurements in the surface force apparatus (SFA), and microrheometry experiments. McKindley and collaborators revealed by micro-rheology, at a 3–4 μm thickness, an elastic behaviour on polysiloxanes, but the authors attributed it to the friction of trapped dust particles [79]. But the much larger body of measurements in very narrow gap geometry is made by SFA, dominated by the developments of Israelachvili [16–18] and Granick [13–15] and co-workers. More recently, high-sensitivity AFM measurements by Aimé and collaborators showed the quasi-solid nanoscale behaviour of simple liquids on graphite [19] (Fig. 6.10).

Granick and co-workers were pioneering in identifying a linear elastic response on a low molecular weight polymer (polyphenylmethylsiloxane) at several molecular thicknesses (Fig. 6.11). They observe an elastic response

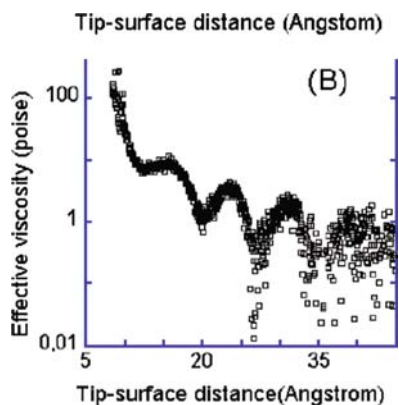


Fig. 6.10. High-resolution Atomic Force Microscopy results revealing the damping and the increase of the viscosity when the gap between the tip and the surface is reduced. After Maali et al. [reprinted with permission of Phys. Rev. Lett. **96** (2006) 86105]

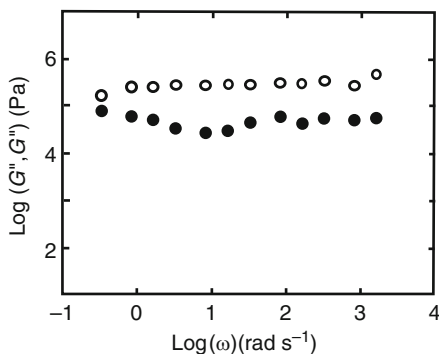


Fig. 6.11. Surface Force Apparatus measurements revealing a solid-like response displayed by a low molecular weight polyphenylmethylsiloxane ($M_w = 4,620$) at 42 \AA and 44°C above the glass transition temperature. After Hu and Granick [reproduced with the permission of the American Association for the Advancement of Science, Science 258 (1992) 1339]

increasing as the thickness decreases. The disappearance at large thicknesses was interpreted by a transition towards the conventional bulk properties. However, increasing the sample thickness, increases dramatically the experimental difficulty to keep the linear conditions. The intrinsic mobility and the various and numerous defects (voids, impurities) lower the chances to let the system unperturbed under strain. The mica surface used in SFA measurements is hydrophilic. It does not offer the best interactivity with respect to organic materials as polymers. As noticed by the author himself ([13–15] Science 1992), the force deformation can produce non-linear responses and the first experiments were not designed to measure linear elastic effects. The

transition from low to high molecular thicknesses might be connected to a non-linear regime, which produces an underestimation of the elastic response. At high strain and/or high slippage conditions (paragraph 9), the non-linear behaviour results in a perfectly linear and reproducible conventional viscoelastic behaviour. Accessing to the real first linear regime is a true challenge and the boundary conditions play an important role for the determination of the optimised parameters. Nanoscale and microscale results, as those above cited, are not measured with the same boundary conditions (different substrates, different roughness, etc.). They are hardly quantitatively comparable, but they describe very likely the same elastic properties.

6.6 Analysis of the Viscoelasticity Scanning Method

A large fraction of rheological testing involves relaxation dynamic measurements. With the emergence of the rheometers [80, 81] and suppliers [Rheometrics (TA-Instruments), Physica (Anton-Paar), Bohlin], this oscillatory shear frequency scanning method rapidly became established as the technique for characterising the dynamic properties of fluids, in particular for materials displaying behaviours intermediate between solids and liquids, that is, viscoelastic behaviours.

In this method, the material is placed in contact with and between two disk-like fixtures, one oscillating, the other fixed one measuring the motion transmitted by the sample via a sensor. From the difference between the input and the output signals, two parameters are extracted; the component in phase with the strain determines elastic (or storage) modulus (G'), whereas the out of phase component defines the viscous (or loss) modulus (G'').

The discipline usually ignores the nature of the boundary conditions between the material and the substrate. Rheology supposes that these conditions have no influence on the measured parameters at the macroscopic scale. The validity of the measurements is ensured by Newtonian calibration of standards of known viscosity and by the reproducibility of the results. However, the determination of the wettability on metallic substrates, as the Aluminium, shows that the boundary conditions are not optimum since the wetting is only partial. We will see in the next paragraph that the consequences are multiple and important.

In the frame of polymer dynamics, the conventional study is carried out on Aluminium or stainless steel fixtures (partial wetting). The dynamic relaxation measurements generate, versus frequency (ω), the typical viscoelastic curve (Fig. 6.12) [33–35]. This curve is linear, that is, displayed a response independent on the strain amplitude. It displays at low frequencies, a terminal behaviour of the flow type. This flow behaviour is depicted by a ω -scaling decrease of the viscous modulus, G'' and a ω^2 -scaling of the elastic modulus G' , G' having a negligible value compared with the viscous modulus. The flow behaviour describes thus the frequency part where the viscous component

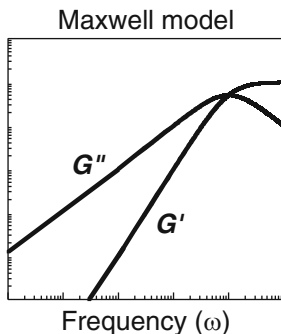


Fig. 6.12. Logarithmic representation of the frequency dependence of the viscous and the elastic moduli fitting with the Maxwell model

becomes major. It also indicates that the lower the frequency, the more rapidly both moduli tighten to zero.

This viscoelastic curve plays an extremely important role: most of the well-known elegant theoretical approaches (in particular Rouse and reptation model) have been developed on the basis of this experimental viscoelastic spectrum which is observed as well in polymers solutions, as in concentrated solutions, polymer melts, and some gel-like behaviours [31, 33–35].

The viscoelastic curve including the terminal flow behaviour has been first modelled by Rouse [14]. It has been constructed to describe the viscoelastic behaviour of diluted polymer solutions; i.e. of non-interacting molecules. Despite this strong limitation, the Rouse model has been also adopted for describing the viscoelastic behaviour of entanglement-free polymer melts, concentrated polymer solutions and entangled melts (modified Rouse approaches). It is the foundation for modern theories of polymer dynamic and viscoelasticity. Basically, in its simplest version, the viscoelastic response (of polymer solution or melts) is described as a single chain response where all the relaxation modes (generating viscous and elastic contributions) originate from the intrachain properties (intermolecular interactions are neglected). The chain is decomposed in a finite number of independent coarse-grains linked to elastic springs (“bead-spring” model – Fig. 6.2) [18–20]. Each grain represents a group of monomers (subchain) sufficiently long to be decoupled from others (gaussian statistics). The bulk viscoelastic quantity is the sum of the statistically-averaged values from individual molecules.

The elastic modulus takes the form of the Maxwell expression (Fig. 6.12):

$$G'(\omega) = nkT \sum_{p=1}^Z \frac{\omega^2 \tau_p^2}{1 + \omega^2 \cdot \tau_p^2},$$

where τ_p is the relaxation time of the group p , Z the number of springs, n the number of chain per volume unit and T the temperature.

The low frequency regime should exhibit a ω^2 -scale behaviour and the zero-frequency limit predicts a collapse of the elastic component.

The viscous part is expressed as

$$G''(\omega) = nkT \sum_{p=1}^Z \frac{\omega\tau_p}{1 + \omega^2 \cdot \tau_p^2}$$

The low frequency regime should exhibit a ω -scale behaviour and the zero-frequency limit predicts a collapse of the viscous component. The slope of the ω -scale line corresponds to the expression of the dynamic viscosity.

Following Ferry [36], to distinguish viscoelastic liquids from viscoelastic solids, a term G_e should be added to the viscoelastic function:

$$G'(\omega) = nkT \sum_{p=1}^Z \frac{\omega^2\tau_p^2}{1 + \omega^2 \cdot \tau_p^2} + G_e(T)$$

G_e is the equilibrium modulus. It expresses infinite relaxation times (i.e. of infinite correlation ranges) that define the elastic character of solid-like materials (typically cross-linked polymers). G_e equals zero for the viscoelastic liquids [36].

This distinction is, however, arbitrary. It supposes that the state (liquid or solid) of the material is known before doing the rheology, whereas precisely, the experiment is supposed to determine the nature of the material. It seems that the criterium for a non-zero G_e is the inability of the material to flow (cross-linked polymers do not flow). The non-flowing property is certainly the signature of a solid-like property but the reciprocal proposition is misleading; a material under flow does imply that G_e is initially equal to zero. Various transitions from an initial solid-like state to a flow regime are possible if the external parameters (gravity, strain, or stress field) are strong enough. This is typically the case of delicate foams, gels, self-assemblies, etc. The flowing capacity is disconnected from the value of G_e .

Now, if it is believed that the nature of the material is defined by the measurement, then the definition should specify within which device resolution limits (weak values of G_e may be not accessible in the past and now accessible with more performing devices). It should be also proven that the measurement is non-perturbative (*first* linear regime). Therefore, it is not trivial to prove experimentally that the liquid state corresponds to G_e equalling zero. This situation may reflect more an ideal and simplified case than the result of the rheological measurement.

Another conceptual difficulty lays in the zero value limit of G' when the frequency lowers to zero. The liquid state is a cohesive state resulting from the interacting nature of its constituting molecules. van der Waals, hydrogen, polar interactions govern the forces of the interactions. The description of liquids [1–4, 82, 83] takes into account these weak force interactions defining an energy of cohesion. The molecule motion (flow) becomes allowed when the friction forces transfer an impulsion that overcomes the mobility energy of the molecules. In other words, because of the interaction forces, any fluid is

supposed to resist before flowing. The energy threshold should be finite. It is incompatible with an absence of elastic quiescent properties at zero frequency limit ($G_e \equiv \theta$) as it is postulated in rheology for a condensed state as the liquid state.

6.7 The Question of the Boundary Conditions: Surface Effects, Wetting, and Slippage

Curiously, whereas the surface is at the central place in a rheological experiment, this discipline generally ignores the possible influence of the surface roughness, the wettability properties on the dynamic measurement, the presence of gas and (micro)bubbles in the fluid, treating in an equal way, rough and smooth surfaces, gas polluted liquids and homogeneous liquids, hydrophilic and hydrophobic surfaces. Rheology supposes that none of these parameters has an influence on the parameters of the measurement at a macroscopic scale. However, surface roughness, presence of gas, and wettability are known to favour the violation of the no-slip boundary condition [84]. The following relationship (Tolstoi formula) proposes to relate the wetting property (determined via the drop method providing the contact angle θ) to the slippage ability (via the slippage length b defined as the distance from the surface corresponding to the origin of the velocity gradient as shown in Fig. 6.13): $b \sim \exp(\sigma^2 \cdot \gamma(1 - \cos \theta)/kT) - 1$, $\theta = 0$ in the case of total wetting (σ is related to the molecule size and γ to the surface tension).

The surface plays a crucial role in a rheological experiment; it is at the origin of the motion transfer. If this dynamic transfer is incomplete (partial slippage), the measurement underestimates the material response. To improve the dynamic transfer and avoid the liquid to move/slip on the substrate, a total wetting is preferable. The total wetting ‘catches’ the molecules on the surface. The interaction of the molecules with the surface may be seen as an activated process: $t \propto e^{-(U/kT)}$ where t is the dynamic (e.g. the life time) of the contact and U the energy of interaction of the molecules with the surface. In the case of total wetting, U is large and t small. Short life time contacts are less disturbed under shear strain than long life time contacts. This mechanism favoured in total wetting may preserve the no-slip boundary condition.

An easy determination of the degree of wetting can be made using the contact angle method [85, 86]. The evolution of the profile of a polymer drop

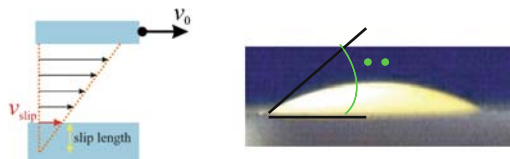


Fig. 6.13. Definition of the slippage length b in a shear field and of the contact angle θ (photograph of a polymer droplet deposited on a flat substrate)

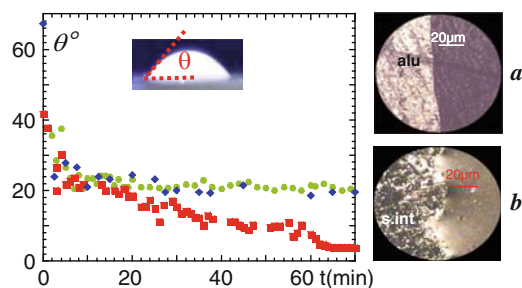


Fig. 6.14. Evolution of the contact angle versus time for an organic material (liquid polybutylacrylate) deposited on different substrates (*sea green circle*: aluminium, *blue circle*: glass, *blue filled circle*: alumina). The contact angle reaches a finite stationary value (20°) in the case of Aluminium and glass substrates (partial wetting), whereas it vanishes in the case of alumina (total wetting). After Noirez et al. (Polymer International in press)

deposited on the surface is measured versus time. The contact angle θ is the angle made by the drop tangent to the surface. In the case of a total wetting, the drop extends itself entirely on the surface. The liquid is progressing via a thin layer (called molecular part) in front of the drop (which is typical of a complete wetting). This molecular part avoids the formation of an air layer between the liquid and the substrate.

Figure 6.14 shows that the metallic substrates conventionally used for the rheological measurements (Aluminium, stainless steel, glass) do not provide a total wetting interaction; the molecular part is absent and a stationary contact angle θ of 20° is observed. In contrast, other substrates as alumina show a rapid propagation of the liquid via a molecular part and a total wetting final state ($\theta \equiv 0$). The boundary conditions are known to have significant quantitative impact on the interpretation of experimental results [22, 23, 84]. We will see in the next paragraph that the differences of wetting produce under the same stress conditions, different viscoelastic behaviours under the same stress conditions, that are perfectly reproducible. The boundary conditions have thus important and multiple consequences.

6.8 Novel Description of Conventional Fluids: from Viscous Liquids, Glass Formers to Entangled Polymers. Experiments in Narrow Gap Geometry: Extracting the Shear Elasticity in Viscous Fluids

A non-vanishing elasticity has been observed far above any phase transition by piezorheometry measurements at the micron scale, on ordinary liquids by Derjaguin since 1989 [69, 70], on liquid-crystal polymers since 1994 by Gallani et al. [87, 88], since 2005 by Noirez and coworkers [89–91], and since 2001 by Colby by dynamic relaxation but interpreted as defects [92, 93]. It was

later identified on a polystyrene melt [71], by micro-rheology on polysiloxanes but the authors McKindley et al. reported on a 3–4 μm thickness elasticity but attributed it to trapped dust particles [94, 95], by viscoelastic measurements carried out at the millimeter scale, using a conventional rheometer for liquid-crystal polymers using standard Aluminium substrates [89–91], and at the sub-millimeter scale, using total wetting surfaces, for the glass formers and the ordinary polymers [72, 73, 75], and finally by Wang [75] who reproduced the results of [61] using the total wetting protocol. The piezorheometer measurements [69–71, 87, 88] were carried on (eventually treated) glasses.

What is the common point between all these experiments? The applied strains originating from the molecular vibration of the piezomembrane are extremely low. Liquid crystal moieties possess intrinsically a high power of anchoring that might help to avoid wall slippage. Finally, small gap thicknesses (reduced volumes) are probed reducing the dissipation by nanobubbles. The common point of these experiments is the absence of dissipation at the surface/liquid boundary and in the volume. Another important point is that the elasticity is accessible using an ordinary rheometer but a strong fluid/substrate interaction (total wetting) is needed to allow it. Even if it is not a necessary condition, the use of wetting substrates facilitates the observation of the elasticity.

6.9 Tribology Meets Rheology. Novel Methods for the Determination of Bulk Dynamic Properties of a Soft Solid or a Fluidic Material

Conventional viscoelastic measurements are carried out under partial wetting conditions. A full contact of the material to the substrate is thus preferable to access to the complete dynamic signature of the fluid. First results, obtained using the full wetting contact method, have been published in [72–74]. As illustrated in Fig. 6.12, organic liquids as polymer melts wet totally Alumina substrates. The following results are obtained using this protocol.

To illustrate the influence of the boundary conditions on the rheological measurement, the elastic modulus of an ordinary melt (Polybutylacrylate) measured at 90°C over T_g ($T_g = -64^\circ\text{C}$) is displayed using partial (conventional aluminium fixtures) and total wetting conditions (Fig. 6.15) [72–74]. The optimisation of the boundary conditions produces a spectacular increase of the elastic modulus. The elastic response is higher and independent of the frequency: a terminal solid-like behaviour replaces the conventional flow behaviour. Except the nature of the substrates, these measurements have been carried out with the same experimental conditions (1% strain amplitude, 20 mm plate–plate geometry, 0.100 mm gap thickness).

Do these results reflect a spectacular surface effect? Surface effects are known to change the thermodynamic at the molecular scale or a multiple of that only [96–98] but not at these macroscopic scales. The authors did

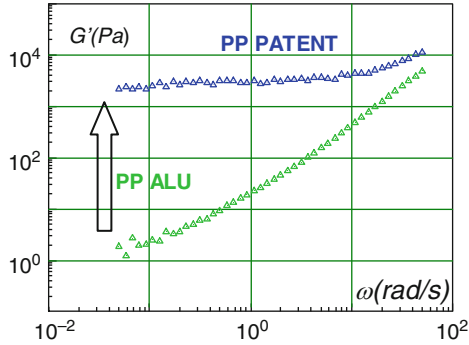


Fig. 6.15. Elastic modulus G' exhibited at room temperature by a molten polymer [Polybutylacrylate ($T_g = -64^\circ\text{C}$) : $M_w = 47,500$, $I = 1.1$] on Aluminium substrate (*green curve*) and with improved boundary conditions (*blue curve*), on similar experimental conditions: 0.100 mm gap thickness, 1% strain amplitude (Polymer International in press)

not notice any calorimetry change on measuring the heat flow on aluminium or on alumina (at 50–100 μm length scales) [76]. Another interesting feature noted by these authors is that the conventional viscoelastic behaviour is also observable on wetting substrates but at high strain amplitudes.

The following scheme (Fig. 6.16) illustrates the evolution of the shear elasticity versus strain amplitude. The conventional linear viscoelastic measurement corresponds to the *Regime 1*. The elasticity is measured in a *regime prior* to *Regime 1*, called '*Regime 0*'. '*Regime 0*' corresponds typically to strain rates of 0.05% up to 1%.

At higher strain amplitudes, the low frequency elasticity lowers (entrance in the non-linear regime ' θ_{NL} ') and vanishes below the resolution device. The strain induced lowering affects all frequencies except a high frequency zone (typically above 10 rad/s). This high frequency zone corresponds precisely to the zone where the conventional viscoelastic curve emerges with a shear modulus scaling as ω^2 . The viscoelastic behaviour is the product of the non-linear regime of the elastic component.

The increase of the strain amplitude is similar as a transition from total to partial wetting conditions. This transition is dynamic: Fast relaxation time contacts are easily restored and give rise to the conventional flow behaviour whereas long time relaxation (solid-like) contacts bear only small strains and exhibit the elastic response.

Since the lowest strain amplitudes ensure the weakest perturbations of the sample, that is, to fulfil the causality–linearity conditions, the elastic response of the fluid corresponds to the real first linear response of the material. Figure 6.17 shows the relaxation behaviour of this polymer liquid submitted to a constant shear strain of 1%. As it can be seen, the polymer does not flow under strain; the shear stress does not collapse after 10^4 s. This is typically a solid-like response, coherent with the dynamic relaxation observations.

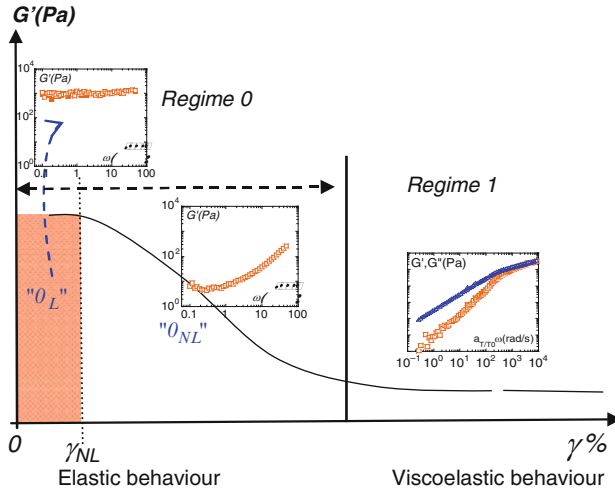


Fig. 6.16. Scheme of the evolution of the elastic modulus (at a given frequency) versus strain amplitude. θ_L linear *elastic* regime, θ_{NL} non-linear *elastic* regime, i_L conventional linear *viscoelastic* regime. At higher strains, the elasticity vanishes and is replaced by the conventional linear viscoelastic response. The spectra display the response of PBuA ($M_w = 40\text{ k}$) at 0.5, 100, 500% strain amplitude and 0.100 mm gap thickness

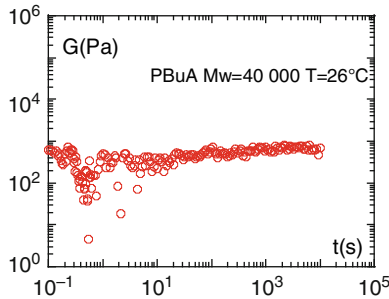


Fig. 6.17. Relaxation modulus (G) versus time (strain amplitude: 1%, gap thickness: 0.2 mm). Measurements carried out on a polybutylacrylate at $T = 25^\circ\text{C}$ ($M_w = 40,000$, $T_g = -64^\circ\text{C}$) using total wetting conditions (Polymer International in press)

The conventional linear viscoelastic behaviour that displays a lowering of both elastic and viscous moduli at low frequency is a second linear regime, obtained once the solid-like response enters into a strong non-linear regime by increasing the strain amplitude.

The remarkable reproducibility of the conventional viscoelastic curve (obtained after a preshear recommended by the rheology protocol) may indicate that the ‘solid-like’ contacts are probably lost by the formation of

an autophobic layer [24–27]. Its superposition to non-linear (steady-state) rheometry [99] may indicate that similar slippage states have been reached.

The elasticity in fluids is hardly observable under conventional (partial) wetting conditions. The wetting parameter, or more generally the nature of the interaction between the fluid and the substrate, is usually not taken into account in conventional rheology, probably for historical reasons. Indeed, wetting, adhesion, friction are new disciplines whose importance were ignored 30–40 years ago, at the time of emergence of the first rheometers [78]. This discipline has been rapidly imposed as the technique of characterisation of fluids and viscoelastic materials with schools, societies (industrial and scientific institutions) to transmit the knowhow to use and to extract the ‘right’ signal from these machines. Because of the easy reproducibility of the viscoelastic curve and because of its modelling in terms of molecular dynamics, a formal protocol has been instituted. Nowadays, the boundary conditions are at the centre of all research areas related to filmic, surfacic, tribology, slippage processes [100–102]. Likely because its established protocol produces reproducible robust results, the dynamic relaxation method did not upgrade the protocol considering the advances of other fields, particularly concerning the boundary conditions. But actually, the converging results showing elasticity at narrow gaps under weak stresses indicate the importance of the boundary conditions. The surface acts as a revealer. When the slippage is avoided, it is then possible to fix and reveal the elastic properties of the material. Recent results show that the elastic response is also observable on glass former liquids [76–78]. The low frequency elasticity of the Polypropyleneglycol PPG-4000 (Fig. 6.18) is here observed at 0.075 mm and 80°C away from the glass transition ($T_g = -75^\circ\text{C}$). An interpretation in terms of entanglement effects rules out and these observations do not establish a straightforward link to the glass transition. It rather suggests a generic property that may originate from condensed state intermolecular forces.

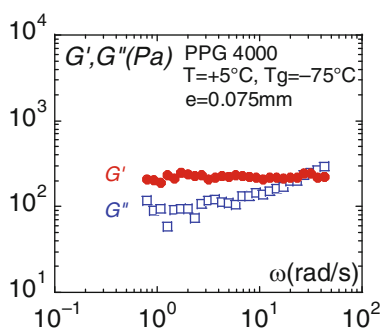


Fig. 6.18. Frequency dependence of the elastic $G'(\omega)$: *red filled circle* and viscous moduli $G''(\omega)$: *open square*, measured for a Glass Former liquid (Polypropyleneglycol – PPG4000, $M_w = 4,000$, $T_g = -75^\circ\text{C}$) at $T = +5^\circ\text{C}$ (total wetting conditions, 0.075 mm gap thickness)

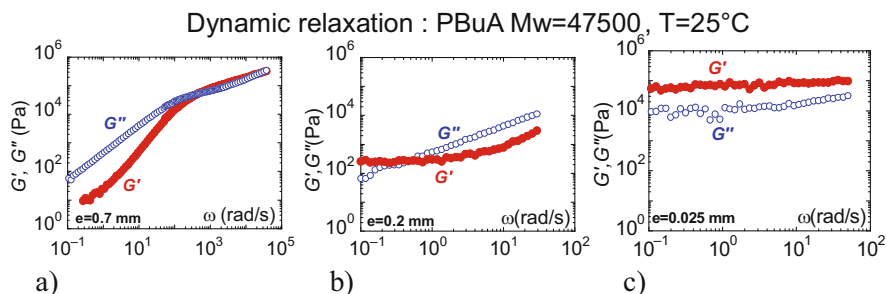


Fig. 6.19. Elastic and the viscous moduli, $G'(\omega)$: filled circle and $G''(\omega)$: open square measured at different gap thicknesses (a: 0.700 mm, b: 0.200 mm, c: 0.025 mm), strain amplitude: 0.5%, measurements carried out on a polybutylacrylate at $T = 25^\circ\text{C}$ ($M_w = 47,500$, $T_g = -64^\circ\text{C}$) (Polymer International in press)

6.10 Elasticity and Dimensionality in Fluids

Interestingly, it seems that the low frequency elasticity is restricted to small gaps, typically below $50\ \mu\text{m}$ using a piezorheometer on treated glass substrates (Fig. 6.6), accessible below the millimeter with a conventional rheometer under full wetting boundary conditions (Fig. 6.19) [72–74, 76]. Figure 6.19a–c displays the dynamic relaxation of a polybutylacrylate liquid measured at room temperature (i.e. 90°C above T_g) at different thicknesses. The terminal elastic mode is visible on typically 0.025–0.5 mm sample thicknesses. Above this thickness, the conventional viscoelastic spectrum dominates and the low frequency elasticity vanishes under the sensibility limit of the rheometer (Fig. 6.19c). The dependence with the thickness is not trivial. It may reveal a dimensional character unknown in fluid dynamics. The dimensional character is likely related to dynamic effects (self-diffusion, voids, mobility). It also reveals the limits of the mechanical way to probe materials (competition between self-diffusive motions, internal mobility and the extremely small mechanical to ensure non-perturbative linear measurements). An analogy with size effects reported for the rheology of cements seems to be pertinent [103]. It is now indeed admitted that cements, concrete beams, obey to scaling mechanical properties. This concept has been established after having for years observed and modelled a size effect in shear failures.

6.11 General Summary and Perspectives

This chapter points out the utmost importance of the role of the characterisation techniques to improve our understanding. It also points out the emergency to upgrade the long-dated knowhow, in particular, in rheology by emphasising the interest of using complementary methods such as the control of the substrate–fluid boundary conditions and the viscoelastic measurements.

Dynamic relaxation probes the macroscopic dynamic of a sample by imposing an oscillatory shear frequency scanning. This mechanical stress method, which is often considered as a 'no-smart' engineering method, is actually of the rare ones able to access the true dynamical properties of materials at a macroscopic scale, especially in the low frequency domain which is hardly accessible by most of the dynamic measurements (NMR, inelastic scattering, dynamic diffusion, dielectric relaxation, etc.). Dielectric relaxation differs from dynamic relaxation in that the dielectric technique is sensitive to orientational effects only, and not to positional effects. Therefore, it is *blind* to solid-like correlations if these ones are not associated to dielectric orientational effects.

The theory of viscoelasticity has been elaborated to model initially the viscoelastic properties of diluted polymer solutions. In the frame of diluted solutions, the viscoelastic curves were interpreted as resulting essentially from intramolecular interactions. The intermolecular interactions were neglected. This model has been then extended to concentrated solutions and unentangled polymer melts on the basis of the similarities displayed by their rheological behaviours, in particular because they exhibit similar terminal flow behaviour. Now, the consideration of the boundary conditions reveals that the shear elasticity may be actually finite (no-vanishing) in the liquid state and at a macroscopic scale, far away from any phase transition.

Indeed, by optimising the interaction of the fluid on the surface and/or lowering the stress during the dynamic relaxation measurement, different groups using different set-ups have identified so far hidden macroscopic solid-like correlations in simple liquids at a micron scale by Derjaguin, in unentangled polymer melts at several tens of microns and several hundreds microns by Martinoty and by Mendil, respectively, and, in glass formers by Noirez and Orrit. These measurements were carried out away from any phase transition. Their generic character, the macroscopic length scales involved and the variety of the substrates probed in these studies indicate that this elasticity originates likely from long-range intermolecular interactions and certainly not from a huge surface-induced effect. The notion of finite (non-vanishing) shear elasticity in the liquid state is new. The elasticity is measured at a macroscopic scale and is not related to the molecular dimensions. It originates from long-range intermolecular interactions. Therefore, it invalidates any attempt of interpretation in terms of single molecular description in contrast to the theory of viscoelasticity. This concept also differs fundamentally from the paradigm that correlations in the liquid state are restricted to short range. The identification of long-range solid-like correlations is however coherent with the description in terms of cohesion used in simple liquid theory to illustrate the high degree of connectivity of them.

Studies as a function of the thickness (from the millimetre down to the micron scale) have shown that the solid-like component is reinforced by decreasing the length scale. The size dependence of the elasticity gives a new perspective for the analysis of fluid properties. It is absent from current theoretical concepts describing fluids far away from any phase transition. The

concept of a length scale for cooperative motions has been developed in the frame of the description of dynamics of the glass transition but it is restricted to vicinity effects [104, 105]. The low frequency elasticity is observable far above any phase transition (more than a hundred degrees above T_g) and far above the molecular scales involved in confined geometry or surface induced effects. It cannot be interpreted in terms of induced effects due to the surface or the confinement. These effects are known to occur at a molecular scale or a multiple of that, and are identifiable by a modification of the thermodynamical state, mainly a slowdown of the dynamics, observable by a notable change of transition temperatures [96–98].

Interestingly, Bazant [99] considers the generic existence of macroscopic size effects to explain size scaling strengths and fracturing in the rheology of cements and extends its model to granular liquids. Finally, a statistical model due to Volino [106] predicts that long-range elastic correlations are a generic property of the liquid state. Furthermore, these elastic properties are predicted to be dimension-dependent, which is in agreement with the dimensional character observed for the low frequency elasticity [69–71, 76, 87, 88]. Future fascinating developments are expected facing the identification of these so far neglected macroscopic elastic properties. A challenge will be to marry these last ones to nanoscale properties, the dimensional bulk properties to surface induced effects, and to better explain and predict fluidic behaviours in the frame of a multiscale approach.

References

1. P.W. Atkins, *Physical Chemistry* (Oxford University Press, Oxford, 1998)
2. M. de Podesta, *Understanding the properties of Matter* (UCL Press, London, 1996)
3. D. Tabor, *Gases, Liquids and Solids: And Other States of Matter* (Cambridge University Press, Cambridge, 1990)
4. B. Cabane, R. Vuilleumier, C.R. Geosci. **337** 159 (2005)
5. J.D. Kirkwood, P.L. Auer, J. Chem. Phys. **19** 281 (1951)
6. P. Dontula, C.W. Macosko, L.E. Scriven, J. Rheol. **49** 807 (2005)
7. R.G. Larson, Rheol. Acta **31** 213 (1992)
8. M.D. Graham, Chaos **9** 154 (1999)
9. G.S. Beavers, D. D. Joseph, J. Fluid Mech. **81** 265 (1977)
10. C. Pujolle-Robic, L. Noirez, Nature **409** 167 (2001)
11. S. Goldstein, Ann. Rev. Fluid. Mech. **1** 1 (1969)
12. C.L. M.H. Navier, *Mémoire sur les lois du mouvement des fluides*, Mémoires de l'Académie Royale des Sciences de l'Institut de France VI, 389 (1823).
13. H.-W. Hu, G.A. Carson, S. Granick, Phys. Rev. Lett. **66** 2758 (1991)
14. A. Levent Demirel, S. Granick, J. Am. Chem. Phys. **115** 1498 (2001)
15. Y. Zhu, S. Granick, Phys. Rev. Lett. **93** 096101-1-4 (2004)
16. J.N. Israelachvili, D. Tabor, Nature **241** 148 (1973)
17. J.N. Israelachvili, M. Tirrell, J. Klein, Y. Almog, Macromolecules **17** 204 (1984)
18. B. Bhushan, J. Israelachvili, U. Landman, Nature **374** 607 (1995)

19. A. Maali, T. Cohen-Bouhacina, G. Couturier, J.P. Aimé, *Phys. Rev. Lett.* **96** 861051 (2006)
20. J.D. Anderson, *Phys. Today* **58**(12) 42 (2005)
21. H. Schlichting, *Boundary Layer Theory* (McGraw Hill, New York, 1979)
22. O.I. Vinogradova, *Langmuir* **14** 2827 (1998)
23. O.I. Vinogradova, R.G. Horn, *Langmuir* **17** 1604 (2001)
24. G. Reiter, R. Khanna, *Phys. Rev. Lett.* **85** 2753 (2000)
25. F.W. Smith, *Wear* **2** 250 (1959)
26. G. Reiter, A.L. Demirel, S. Granick, *Science* **263** 1741 (1994)
27. G. Luengo, J. Israelachvili, S. Granick, A. Dhinojwala, *Wear* **200** 328 (1996)
28. J.P. Hansen, I.R. McDonald, *Theory of Simple Liquids* (Academic Press, San Diego, CA, 1991)
29. P.A. Egelstaff, *An Introduction to the Liquid State* (Oxford University Press, Oxford, 1994)
30. S. J. Henderson, R. J. Speedy, *J. Phys. Chem.* **91** 3062 (1987)
31. P.E Rouse, *J. Chem. Phys.* **21** 1272 (1953)
32. N. Nemoto, J. L. Schräg, J. D. Ferry, *Polym. J.* **7** 195 (1975)
33. M. Doi, S.F. Edwards, *The Theory of Polymer Dynamics* (Clarendon Press, Oxford, 1986)
34. G. Marrucci, F. Greco, G. Ianniruberto, *Curr Opin Colloid Interface Sci* **4** 283 (1999)
35. G. Marrucci, G. Ianniruberto, *J. Non-Newtonian Fluid Mech.* **82** 275 (1999)
36. J.D. Ferry, *Viscoelastic Properties of Polymers* (Wiley, New York, 1980)
37. M. Guenza, *Phys. Rev. Lett.* **88** 25901 (2002)
38. M. Guenza, *Macromolecules* **35** 2714 (2002)
39. J.D. Kirkwood, J. Riseman, *J. Chem. Phys.* **16** 565 (1948)
40. P.E. Rouse, *J. Chem. Phys.* **21** 1271–1280 (1953)
41. Y-H Lin, *Polymer Viscoelasticity, Basics, Molecular Theories and Experiments* (World Scientific, Singapore, 2003)
42. V. L. Jarkko, J. Saarinen, K-E Peiponen, E.M. Vartiainen, *Kramers-Kronig Relations in Optical Materials Research* (Springer, Heidelberg, 2005)
43. R de L. Kronig, *J. Opt. Soc. Am.* **12** 547 (1926)
44. H.A. Kramers, *Atti Cong. Intern. Fisica*, (Transactions of Volta Centenary Congress) Como, **2** 545 (1927)
45. P.D. Olmsted, P.M. Golsbart, *Phys. Rev. A* **46** 4966 (1992)
46. P.D. Olmsted, *Curr. Opin. Colloid Interface* **4** 95 (1999)
47. M.E Cates, S.M. Fielding, *Adv. Phys.* **55** 799 (2006)
48. J. Vermant, *Curr. Opin. Colloid Interface Sci.* **6** 489 (2003)
49. J.K.G. Dhont, W. J. Briels, *Rheol. Acta* **47** 257 (2008)
50. K. Osaki, M. Kurata, *Macromolecules* **13** 671 (1980)
51. C.M. Vrentas, W.W. Graessley, *J. Rheolog.* (NY) **26** 359 (1982)
52. J. C. Haley, T.P. Lodge, *Colloid Polym. Sci.* **282** 793 (2004)
53. J.A. Pathak, R.H. Colby, G. Floudas, R. Jérôme, *Macromolecules* **32** 2553 (1999)
54. J.C. Haley, T.P. Lodge, Y. He, M.D. Ediger, E.D von Meerwall, J. Mijovic, *Macromolecules* **36** 6142 (2003)
55. X. Yang, SQ Wang, A. Halasa, *Rheol. Acta* **37** 415 (1998)
56. W. Kuhn, F. Grün, *Kolloid-Zeitschrift* **101** 248 (1942)
57. J. Bastide, J. Herz, F. Boué, *J. Phys.* **46** 1967 (1985)

58. Boué F., Bastide J., Buzier M., Colette C., Lapp A., Herz J. *Prog. Colloid Polym. Sci.* **75**, 152 (1982)
59. Wang S.Q., Ravindranath S., Boukany P., Olechnowicz M., Quirk R.P., Halasa A., Mays J. *Phys. Rev. Lett.* **97**, 1878011 (2006)
60. G. Marrucci, F. Greco, G. Ianniruberto, *Curr. Opin. Colloid Interface Sci.* **4** 283 (1999)
61. U. Pschorn, E. Roessler, H. Sillescu, S. Kaufmann, D. Schaefer, H.W. Spiess, *Macromolecules* **24** 398 (1991)
62. P. Sotta, B. Deloche, J. Herz, A. Lapp, D. Durand, J.-C. Rabadeux, *Macromolecules* **20** 2769 (1987)
63. M. Kehr, N. Fatkullin, R. Kimmich, *J. Chem. Phys.* **127** 84911 (2007)
64. R. Graf, A. Heuer, H.W. Spiess, *Phys. Rev. Lett.* **80** 5738 (1998)
65. K. Nakajima, H. Watabe, T. Nishi, *Polymer* **47** 2505 (2006)
66. P.P. Dra, S-Q. Wang, *Phys. Rev. Lett.* **75** 2698 (1995)
67. S.Q. Wang et al. *Phys. Rev. Lett.* **97** 187801-1-4 (2006)
68. P.E. Boukany, S-Q Wang, *J. Rheol (NY)* **51** 217 (2007)
69. B.V. Derjaguin, U.B. Bazarov, K.T. Zandanova, O.R. Budaev, *Polymer* **30** 97 (1989)
70. B.V. Derjaguin, U.B. Bazarov, Kh.D. Lamzhapova, B.D. Tsidypov, *Phys. Rev. A* **42** 2255 (1990)
71. D. Collin, P. Martinoty, *Phys. A* **320** 235 (2002)
72. H. Mendil, P. Baroni, *Europhys. Lett.* **72** 983 (2006)
73. H. Mendil, P. Baroni *Eur. Phys. J. E* **19** 77 (2006)
74. P. Baroni, H. Mendil, L. Noirez, *Méthode de détermination des propriétés dynamiques d'un matériau fluide ou solide déformable*, patent n°05 10988 (27/11/2005)
75. S-Q Wang, S. Ravindrantah, Y. Wang, P. Boukany, *J. Chem. Phys.* **127** 64903 (2007) The authors observe (paragraph II) the low frequency solid-like response on following the full wetting protocol developed on [36]
76. L. Noirez, P. Baroni, H. Mendil-Jakani, *Hidden Solid-like Correlations in Liquid Polymers and Glass Formers*, (*Polymer International* 2009) **58** 962 (2009)
77. R. Zondervan, T. Xia, H. van der Meer, C. Storm, F. Kulzer, W. van Saarloos, M. Orrit, *Proc. Natl. Acad. Sci. U.S.A.* **105** 4993 (2008)
78. Y. Chushkin, C. Caronna, A. Madsen, *Eurphys. Lett.* **83** 36001 (2008)
79. H.P. Kavehpour, G.H. McKinley, *Tribol. Lett.* **17** 327-335 (2004)
80. T. Nakagawa, *Bull. Chem. Soc. Jpn.* **27** 133 (1954)
81. G. Harrison, *Rheol. Acta* **13** 28 (1974)
82. E. Andrade, *Nature* **125** 309 (1930)
83. H. Eyring, *J. Chem. Phys.* **3** 492 (1935)
84. E. Lauga, *Microfluidics: The No-Slip Boundary Condition*, ed. by C. Tropea, A. Yarin, J.F. Foss. Chap. 19 *Handbook of Experimental Fluid Dynamics* (Springer, New York, 2007)
85. T. Young, *Philos. Trans. Soc., Lond.* **95** 65 (1805)
86. P.G. de Gennes, F. Brochard-Wyart, D. Quéré, *Gouttes, Bulles, perles et ondes* (Belin, Paris, 2005)
87. J.L. Gallani, L. Hilliou, P. Martinoty, *Phys. Rev. Lett.* **72** 2109 (1994)
88. P. Martinoty, L. Hilliou, M. Mauzac, L. Benguigui, D. Collin, *Macromolecules* **32** 1746 (1999)
89. Mendil, L. Noirez, P. Baroni, *Europhys. Lett.* **72** 983 (2005)

90. L. Noirez, Phys. Rev. E **72**, 051701 (2005)
91. H. Mendil, I. Grillo, P. Baroni, L. Noirez, Phys. Rev. Lett. **96** 077801–4 (2006)
92. R.H. Colby et al., Europhys. Lett. **54** 269–274 (2001)
93. E.P. Choate, Z. Cui, M.G. Forest, Rheol. Acta **47** 223 (2008)
94. H.P. Kavehpour, G.H. McKinley, J. Non-Newtonian Fluid Mech. **124** 1 (2004)
95. H.P. Kavehpour, G.H. McKinley, Tribol. Lett. **17** 327 (2004)
96. K. Dalnoki-Veress, J.A. Forrest, C. Murray, C. Gigault, J.R. Dutcher, Phys. Rev. E **63** 31801-1-10 (2001)
97. J.A. Forrest, R.A.L. Jones, *Polymer Surfaces, Interfaces and Thin Films*, ed. by A. Karim, S. Kumar (World Scientific, Singapore, 2000)
98. P.G. de Gennes, Eur. Phys. J. E **2** 201–205 (2005)
99. Y. Heo, R.L. Larson, Macromolecules **41** 8903 (2008)
100. P. Joseph, PhD thesis, Paris VI University (2006)
101. L. Léger, E. Raphael, H. Hervet, in *Advances in Polymer Science, Polymer in Confined Environments*, vol. **138**, ed. by S. Granick (Springer, Berlin, 1999), pp. 185–225
102. J.L. Barrat, L. Bocquet, Faraday Discuss **112** 119–127 (1999)
103. Z.P. Bazant, *Scaling of Structural Strengths* (Hermes Penton Science, London, 2002)
104. C. Bennemann, C. Donati, J. Bashnagel, S. Glotzer, Nature **399** 246 (1999)
105. G. Biroli, J.-P. Bouchaud, Europhys. Lett. **67** 21 (2004)
106. F. Volino, Ann. Phys **22**(1) 2 (1997)

Combining Atomic Force Microscopy and Depth-Sensing Instruments for the Nanometer-Scale Mechanical Characterization of Soft Matter

Davide Tranchida and Stefano Piccarolo

Summary. Complex materials exhibit a hierarchical structure where a gradient of features on nanometer scale is induced by the synthetic route eventually enhanced by the loading condition. The nanometer scale at which individual components arrange, determining their properties, is a current challenge of mechanical testing. In this work, a survey on nanoindentation is outlined based on the comparison of results obtained by Atomic Force Microscopy and Depth-Sensing Instruments and their combination. An Atomic Force Microscope equipped with a Force Transducer gives indeed the possibility to scan the sample surface in contact mode, thereby allowing one to choose a suitable position for the nanoindentation, as well as imaging the residual imprint left on the sample. The analysis of the applied load vs. penetration depth curve, also called force curve, shows the limitations of current approaches to determine elastic moduli of compliant viscoelastic materials. Significant deviations from the expected values are observed even after optimizing testing conditions, so as to minimize the artifacts like viscoelastic effects or pile-up. As rigorous approaches are yet to be applied to the interpretation of force curves accounting also of viscoelastic material behavior, an empirical calibration recently proposed by the authors is verified against a set of data on model samples spanning a range of moduli, typical of compliant materials and close to each other, so as to challenge the resolution potential of this method, as well as others in use in the literature.

Key words: Atomic force microscopy, Depth-sensing instrument, Nano-indentation, Soft materials, Polymers, Elastic Young's modulus, Mechanical properties.

7.1 Introduction

Biological samples exhibit a hierarchical structure with an heterogeneous distribution in shape and composition of its constituent building blocks. Hierarchy is also observed in synthetic and natural polymers when properly processed, so as to give rise to an heterogeneous distribution of morphologies.

As a consequence of the hierarchical nature of such materials mechanical heterogeneity is expected to exist at multiple length scales with characteristic features down to the nanometer scale [1, 2]. Enhanced spatial and displacement resolutions are needed to bridge morphology and mechanical properties highlighting the effect of each assembly of micro- and nano phases and their individual contribution to the overall mechanical response [4–13]. Nanoindentations performed with Depth-Sensing Instruments (DSI) can provide a very high resolution, with applied loads down to $10\mu\text{N}$ and displacements of few tens of nanometers [14, 15]. Although the analysis of force curves can repeatedly provide a reliable estimate of the Young's modulus of metals and ceramics, the technique is still affected by several difficulties when applied to polymers [16, 17] or soft matter, in general. Indeed, the common method to analyze the force curve, the Oliver and Pharr (O&P) procedure [14, 15], fails when trying to evaluate the Young's modulus of polymers, and results in gross overestimations, even up to three times [17, 18] when compared with the results obtained in macroscopic tests. Viscoelastic behavior is often thought to be the cause of this failure.

This observation clearly implies that the O&P procedure, which fits the unloading curve with a power law relation since related to the contribution of elastic recovery alone, cannot be applied in the case of polymers. Any attempt of its use results in an unreasonably high or even negative slopes, as demonstrated by the theoretical viscoelastic analysis of Ting [19]. Even though the “nose” is not always observed, Cheng and Cheng [20] showed by numerical simulations of the response of viscoelastic materials, experimentally verified by the authors [17], that the contact stiffness, i.e., the slope of the unloading curve evaluated at the maximum load, was independent of the loading history in displacement-controlled experiments, provided that the loading rate was high enough. If one recalls that different loading rates lead to different applied load needed to achieve the same penetration depth, which is indeed expected for viscoelastic materials, then different contact depths are estimated by the O&P procedure (see (7.4) in the following). On the contrary, Cheng and Cheng [20] found by numerical simulations that the contact depth obtained by nanoindentations performed at different loading rates, but the same penetration depth was identical. On this basis, they concluded that the O&P procedure must lead to significant errors in determining the contact depth and consequently, Young's modulus.

In order to overcome these limitations and to include viscoelasticity into the analysis, several approaches have been proposed recently [21–25]. Viscoelastic contact mechanics models have been developed in the past, but their application has to face serious experimental and analytical difficulties. For this reason, the analysis is often limited to the holding period, i.e., when either the load or the displacement is kept constant after the loading and before the unloading part, which resembles, respectively, a creep or a relaxation experiment. Creep data have been fitted to a simplified 3-element (Maxwell or Voigt) or 4-element (combined Maxwell–Voigt) models [21], sometimes adding

elements to take also the plasticity into account [22]. Simplified viscoelastic contact models have also been used [23–25].

The aim of this chapter is the description of a combined approach, using a DSI-based system in conjunction to Atomic Force Microscope, (AFM), in order to exploit the characteristics of both experimental procedures, as highlighted in the next paragraph. Three different procedures, used for the analysis of the experimental data, will be described in a good detail, and the contact mechanics basis of each one will be analyzed together with the expected causes of their limitations.

Despite these flaws, mainly due to the onset of viscoelastic behavior, a prominent characterization of the mechanical properties of nanostructured materials will be shown. In particular, three procedures are tested against a set of homogeneous model samples whose morphologies are distinguished by a suitable solidification procedure and whose properties differ only slightly, in relation to their morphology; and are also determined by macroscopic testing methods. Finally, two situations of graded mechanical properties, surface modification by means of either UV irradiation or deposition of a hard coating, are discussed.

7.2 Determining Elastic Modulus of Compliant Materials from Nanoindentations

Commercial, stand-alone nanoindenters are the response to the requirement of measuring mechanical properties on a small scale when testing a surface, a thin film, multilayered structures, and graded materials. Standard indentation is unsuited when a very high lateral resolution is required. In these cases, optical microscopy cannot be applied any more, and other visualisation methods such as atomic force microscopy or scanning electron microscopy must be used. However, the localisation of tiny indents proves to be a time-consuming or even impossible undertaking. From the load-displacement data of DSI, many mechanical properties such as hardness and elastic modulus can be determined without imaging the indentation, and the nanoindenter has also been used to measure the fracture toughness and fatigue properties of ultrathin films, [26,27] which cannot be measured by otherwise, conventional indentation tests. With a tangential force sensor, even nanoscratch and wear tests can be performed at ramping loads [28–30]. Although very high resolutions are obtained by DSI with respect to the measured penetration depth, load sensitivity is sufficient when working with hard materials, as for example metals and ceramics; but it represents a limitation on soft matter, i.e., polymers and biologic samples, causing too high penetration depths and thus, reducing lateral resolution. This severely hampers the possibility to study mechanical properties of these systems on a true nanometer scale.

An alternative is offered by optical lever-based systems [31,32], where the load is applied by cantilever bending – the latter being designed to sense low

loads for very compliant materials or larger loads, in the order of a lower bound achievable in standard DSI, adopting a stainless-steel cantilever with a glued diamond tip.

In this regard, in particular, the AFM can be used to collect force curves for nanoindentations and wear tests. As these instruments are obviously optimized to trace surface topography, the measurement of the force and the relative motion between the tip and the sample with the required accuracy needs several calibrations and preliminary determinations in order to turn a force curve into a reproducible plot of force vs. penetration depth [31–33]. Moreover, closed-loop AFM scanners are often needed to minimize the artifacts arising from the piezoelectric scanner creep and hysteresis. These methods give the possibility of true nanometer scale indentations on soft matter due to the low loads which can be varied in a broad range, from nN up to tens of μN , just by changing the cantilever elastic constant [34–36]. A vast amount of literature is now available discussing several of the various calibrations needed and analysing samples with Young's moduli in the range of few kPa up to few GPa [32, 37–44].

An obvious advantage of the AFM is that it can provide in situ high-resolution images of sample topography before and after the nanomechanical tests, and even more sophisticated morphological information, e.g., by phase imaging, culminating in a very powerful tool for mapping mechanical properties with a very good lateral resolution of inhomogeneous materials when one recalls the low loads and consequently, low penetration depths used. More advantages and information can even be obtained by using electrostatic force transducer systems, combined with an AFM [45]. Indeed, it is possible to precisely adjust vertical load, measure vertical displacement and friction coefficient during the wear experiment, and even measure the residual wear depths of the wear trace by not inflicting a further deformation as it happens when the trace is scanned applying non-negligible loads [46–55].

In this case, the load range used for indentation is intermediate between the AFM and the DSI nanoindenters, as shown in Fig. 7.1, in the case of a compliant sample, a poly (propylene) solidified from the melt at a relatively low cooling rate of 2.5 K s^{-1} [56]. It is interesting to note that the slope of the force curves in the case of a standard AFM, and the AFM with the electrostatic force transducer is nearly the same, though the curves are translated vertically. This is due to the use of indenters with different geometries – the sharp AFM tip causing a larger penetration depth at the same load with respect to that by the AFM with a force transducer, FT (hereafter as AFM–FT), with a Berkovich indenter [57].

A fundamental advantage of the AFM–FT over a common DSI system is that it is possible to collect images of the residual indents. For example, Fig. 7.2 shows the imprint left behind the indenter after a very slow (on the left) and a fast nanoindentation (on the right). In the first case, a huge pile-up, [58–60] i.e., material bulging out of the surface, can be noted. This affects the contact mechanics, as the contact area is much larger than the expected

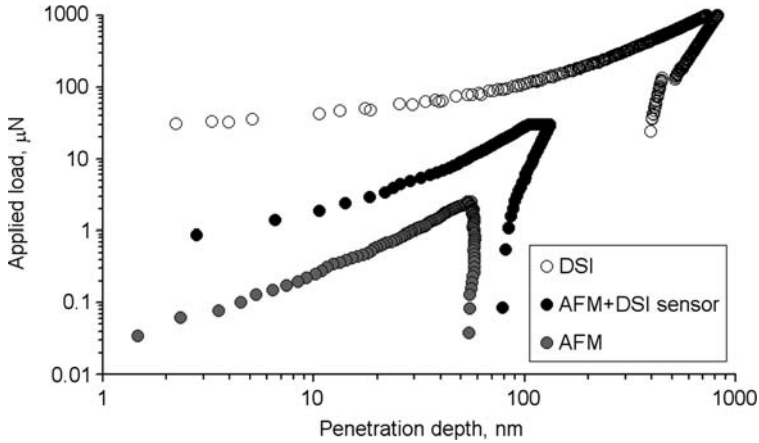


Fig. 7.1. Comparison of force curves obtained by nanoindentations on a isotactic poly(propylene) sample as obtained by different instruments. The AFM curve refers to nanoindentations with a standard sharp AFM tip, while the indentations performed with DSI sensors were performed with the much blunter Berkovich indenter

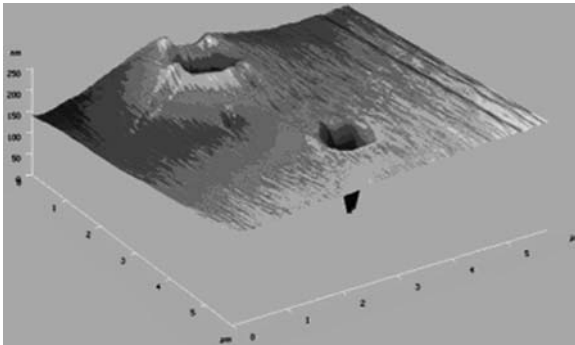


Fig. 7.2. Two residual imprints left behind the nanoindentation of a poly(propylene) sample. The left indentation was performed with a very low loading rate and shows a significant pile-up, unlike the one on the right which was performed with a high loading rate. Reproduced from [17] with permission of ACS

without the pile-up, but cannot be recognized from the force curve alone. Rather, one would observe the ability of the sample material to withstand a high load, though this is only an artifact due to the increased contact area, thus predicting an elastic modulus larger than expected. It may thus happen that the conclusions drawn only from the force curves without being able to access the residual imprints, are erroneous.

In order to produce an image of the indent, a tiny load must be applied to the sensor-producing scratches that are severely modifying the texture of the

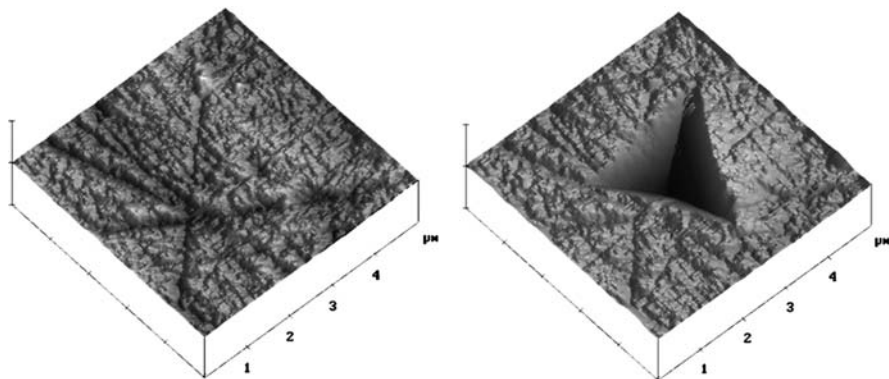


Fig. 7.3. Images of an enamel surface, taken using the nanoindenter tip to scan the surface in contact mode, showing the residual imprint of the nanoindentation. Vertical range is 500 nm. Reproduced from [46] with permission of IOP

imprint obtained after indentation. Such loads are incomparably larger than the perturbations applied by an AFM tip scanned over the sample in a noncontact mode [61]. Therefore, though such visualization tools might help imaging the indent for hard materials, care must be taken to evaluate the effective residual imprint developed on soft matter after an indentation experiment.

Another example is given in Fig. 7.3 and 7.4, showing the ability of the AFM–FT to identify the place where an indentation is performed when compared with some sign on the sample, Fig. 7.3, [46] or to locate the exact position of one indentation with respect to others, when one is interested on mapping the nanomechanical properties. Indeed, Fig. 7.4 shows four indents performed at the distance of $2\mu\text{m}$ from each other on a poly (carbonate) sample [62].

It is worth to remember that, being a scanning probe instrument, the problem of tip convolution during imaging applies in the case of AFM–FT as well [63–65]. This could be a major problem when trying to distinguish very small features on the sample as the Berkovich tip is very blunt when compared with standard AFM tips; however, it can be neglected to a certain extent when imaging the residual imprints since the sample elastically recovers more in the vertical direction than in the lateral one during the unloading; and therefore, the imprints are normally much blunter than the Berkovich tip itself [66, 67].

For comparison, the peculiar contact mechanics of AFM nanoindentation has been extensively studied in the past. In the case of rubber samples, where the elastic behavior is extensive, the indentation is completely reversible and the residual imprint is not appreciable on the scale the AFM can image [37]. For other glassy and semi-crystalline polymers, appropriate testing conditions, e.g., high loading rates, can be identified resulting in residual imprints up to five times smaller than the maximum penetration depth under load, showing a significant elastic recovery [37].

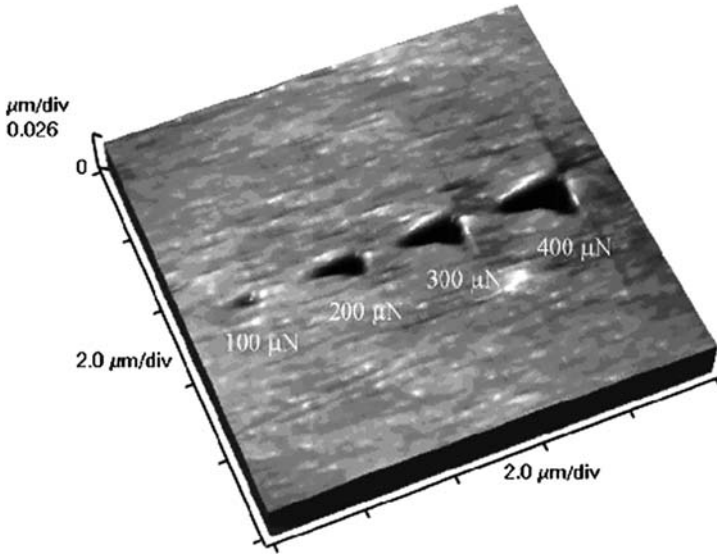


Fig. 7.4. Residual indents after a series of nanoindentation performed with increasing load and space of $2\ \mu\text{m}$, on a PC sample. Reproduced from [62] with permission of Elsevier

For the latter systems, the mechanical behavior in the loading portion of the force curve is dominated by an elastic response, with a close relationship of the penetration depth on the material's macroscopic elastic modulus at constant load. The applied load scales with the penetration depth following a power relation with an exponent equal to 1.5, similarly to rubbers; and it obeys elastic contact models like the one from Sneddon [68], which assumes that the tip can be described as a paraboloid with a certain radius of a curvature.

This elastic-like behavior can be attributed to several distinct and probably, synergetic phenomena: the large elastic range [69], typical of polymeric materials which, accompanied by the confinement determined by the volume around the indentation, induces side constraints which enforce the deformation especially in the case of a compression geometry, where the absence of cavitation significantly extends the elastic range [70, 71]. Also the use of a sharp indenter, like the tip attached to AFM cantilevers, takes advantage of size effects to enhance the elastic behavior [37, 72]. Indeed, even though a sharp indenter induces high stresses in the volume tested, the stress field is rapidly changing in space, and only a small volume of the material experiences a stress higher than the macroscopic yield stress. If such volume is smaller than a critical value, it will not give rise to the onset of plasticity and therefore, the indentation is dominated by elastic deformations [73–76].

In this case, the use of the Berkovich indenter, much blunter when compared with the AFM tip, does not allow to exploit size-scale effects. For example, Table 7.1 shows the ratio of the stress needed to induce plasticity,

Table 7.1. Stress concentration beneath typical indenter tips in dependence of material behaviour [72]

	AFM tip – polymers	Berkovich – polymers	Berkovich – metals
Size scale (σ_0/σ_Y)	5.43	1.02	2.90

σ_0 , and the bulk yield stress of the material, σ_Y , calculated after Bushby and Dunstan [72]. In this simplified model, the Young's modulus and the yield stress of the material, as well as the indenter radius and the radius of the critical volume needed to induce plasticity, are taken as input parameters allowing one to evaluate the stress that the material can hold before the yield arises. Table 7.1 shows the enhancement of material strength in the case of AFM nanoindentations on polymers, while in the case of polymers indented by a Berkovich tip the enhancement is negligible. This implies that plasticity is reached soon and on a larger volume during AFM–FT nanoindentations, and the loading curve obtained when using the Berkovich indenter at relatively high loads should be analyzed with elastic plastic models. However, this discussion is beyond the scope of this chapter and the interested reader can find some more information in [77]. A common approach to extract the elastic modulus from force curves consists then in analyzing the unloading curves, with the assumption that the deformation is, in this case fully elastic.

7.3 Determining Elastic Modulus of Compliant Materials from Nanoindentations

The nanoindentation technique is nowadays a well-assessed technique to measure elastic modulus of metallic and ceramic materials, even though improvements are continuously introduced. Indeed, the mechanical problem of the contact between two bodies is extremely difficult to handle in the case of elastic-plastic, viscoelastic, viscoelastic-plastic behaviors. Different procedures, obtained from theoretical derivations and finite elements simulations, were proposed in the past in order to evaluate elastic modulus of the sample. Among them, the most widely used methods are those from Oliver and Pharr [14] and Loubet [78,79], both based on the hypothesis that the unloading (and the resulting recovery of penetration depth) takes place entirely in the elastic range. With this assumption, it is possible to use the theoretical solution of the problem of a rigid body indenting a perfectly elastic half-space, provided by Sneddon [68]. An account of these two methods will be provided in this background part.

Despite these noticeable improvements, nanoindentation on soft matter is still a quite difficult subject and this holds true in particular for polymers, due to all the experimental issues posed by their unique mechanical behavior. In this and in the following paragraph, the effect of several experimental parameters is analyzed and a theoretical background is first presented

in order to identify the main experimental problems. A short albeit, detailed analysis is given afterwards, showing either the effects of these issues or the resulting limitations. They include an indentation size-scale (ISE) effect, for which sometimes, the elastic modulus and the hardness evaluations show to be dependent on the penetration depth [14, 80–83], and an inaccuracy in the measurements when following the original procedures recommended by Oliver and Pharr and by Loubet.

Once all these aspects are examined, the evaluation of elastic modulus on a nanometer scale is attempted, thereby trying to get accurate, reliable, and reproducible results. Besides testing glassy polymers which have become a standard in this kind of analysis, the measurement of elastic properties will be shown also on a series of poly (propylene) samples with varying crystallinity. The peculiarity of these samples relies on the fact that a special apparatus allows to obtain homogeneous samples with controlled morphologies, thereby tuning the mechanical properties on the basis of the different quenching conditions adopted, and thus, a metallurgical approach adapted to the peculiar thermal properties of polymers [56]. The samples analyzed in this chapter were indeed prepared with cooling rates in the range of $2.5\text{--}888\text{ K s}^{-1}$, thus obtaining semi-crystalline samples at low cooling rates and completely mesomorphic samples at high cooling rates. Samples with both semi-crystalline and mesomorphic phases were obtained at intermediate cooling rates, so that a series of samples with changing morphology and related mechanical properties was obtained. The ability to discriminate the mechanical properties of the different morphologies is a fundamental preliminary test in order to perform nanoscale mapping of properties on inhomogeneous samples.

The problem of determining the distribution of stress within an elastic half space when it is deformed by a normal pressure that was exerted against its boundary by a rigid punch dates back to 1885, to Boussinesq [84]. Using the method of potential theory, Boussinesq derived a solution for the problem, but this theoretical solution did not lend itself to practical computations; and numerical results based on his solution were derived only in the case of flat-ended cylindrical punch and conical punch = [85, 86].

Subsequently, Hertz found a solution in the case of contact between two spheres, valid only at small displacements [87, 88]. This solution can lead to the case of a sphere indenting a half space, just setting one of the curvature radii equal to infinite. The Hertz equation relating applied load L , penetration depth p , sphere curvature radius R , and reduced elastic modulus E_r is

$$E_r = \frac{3}{4} \frac{L}{R^{0.5} p^{1.5}}, \quad (7.1)$$

where E_r depends on the elastic modulus and Poisson ratio of both the sample and the indenter materials.

After these two pioneering works, several alternative solutions were derived.

More recent works rely on the use of Hankel transforms, which are particularly suitable for the study of axisymmetric problems. The most important

among these solutions is the one from Sneddon [68], relating penetration depth and shape of the indenter and, eventually, penetration depth and contact area. In the case of a conical punch, the load-penetration depth is given as

$$L = \frac{2E_r}{\pi \tan \alpha} p^2, \quad (7.2)$$

while in the case of a paraboloidal indenter with curvature radius at the apex R , the relation becomes

$$L = \frac{4}{3} E_r R^{0.5} p^{1.5}. \quad (7.3)$$

Note that this relation is similar when compared with Hertz's one for a spherical indenter, and the latter comes back to Sneddon's one for a paraboloid in the limit of small displacements.

Several phenomenological approaches have been developed to measure mechanical properties from indentation load-displacement data, most of which have focused on the elastic modulus, E , and the hardness, H . Central to these approaches are the methods by which experimentally measurable quantities such as the indentation load, P , the indenter penetration depth, h , and the indentation contact stiffness, $S = dP/dh$, are related to the projected contact area, A , and the elastic constants of the material, E and ν (Poisson's ratio). In one form or another, most methods make use of the relation

$$S = \frac{2}{\sqrt{\pi}} \frac{E}{(1 - \nu^2)} \sqrt{A}. \quad (7.4)$$

This fundamental equation has its roots in elastic contact theory, specifically, the analyses of Love and Sneddon [68, 85], for contact of an isotropic, elastic half-space by rigid indenters of various geometries as mentioned above. Although originally derived for indentation by a rigid cone, Bulychev [89] showed that (7.4) also applies to spherical and cylindrical indenters and speculated that it may hold for other geometries as well. Subsequently, Pharr, Oliver, and Brotzen [90] showed that the equation applies to any indenter that can be described as a solid of revolution of a smooth function. More recently, Cheng and Cheng [91] have suggested that (7.4) is even more broadly applicable than it was previously thought, applying to elastic-plastic as well as purely elastic contact.

Two important results of Sneddon's solution are the expressions for the penetration depth, the contact radius, a , and the indentation load. The penetration depth is given by

$$p = \frac{\pi}{2} h_c = \frac{\pi a}{2 \tan \theta}, \quad (7.5)$$

where θ is the half-opening angle of the cone indenter. In the case of nanoindentations performed with a pyramidal indenter, as the Berkovich one, θ is the equivalent, half-opening angle of the cone that guarantees the same ratio of volume displaced and penetration depth obtained. A very important result is that (7.5) shows that the ratio of the contact depth, h_c , and the total depth is constant, and equal to $2/\pi$. The expression for the load is

$$L = \frac{\pi}{2} \frac{E}{(1 - \nu^2)} a h_c = \frac{\pi}{2} \frac{E}{(1 - \nu^2)} \frac{a^2}{\tan \theta}, \quad (7.6)$$

which illustrates the parabolic dependence of the load on contact depth. Combining (7.6) and (7.7) and noting that $S = dL/dh$ and $A = \pi a^2$ yield the fundamental relation given by (7.4).

It can be noted that (7.4) contains an unknown quantity, the contact area. Moreover, the load as measured from the instrument can be different from the true one applied to the sample because of instrument compliance, i.e., the amount of instrument displaced when applying a certain load. This displacement becomes more important when indenting hard materials, such that a high load must be applied resulting in small displacement of the sample.

Recently, Oliver and Pharr [14, 15] rearranged these results in order to extract information about mechanical properties of a sample through indentation tests. The fundamental assumption is that during loading a material shows elastic-plastic behavior, but on unloading the force curve collected during a nanoindentation records only the elastic recovery of the material thus obeying Sneddon's relations.

Following the previous remark, two calibrations have to be performed in the O&P procedure. The first one regards the evaluation of instrumental compliance. The second calibration involves the area function, i.e., the evaluation of the contact area A in (7.4) from the contact depth that is estimated from the unloading part of the force curve.

Loubet and Hochstetter et al. [78, 79] pointed out the relevance of viscoelastic behavior on the unloading portion of the force curve and recognized with others [57] that fast indentations are needed to minimize such effects which otherwise, preclude the use of the O&P approach to polymers.

The plastic depth h_r (which is not equal to the contact depth under full load) is defined by the following equation, where L is the applied load, S the contact stiffness, and h_t the total depth corresponding to L :

$$h_r = h_t - (L/S). \quad (7.7)$$

The plastic depth is then calculated from h_t after measuring the corresponding load, L , and contact stiffness, S .

For a stiff indenter made of diamond and for a material having a constant Young's modulus, like fused silica, a linear relationship between contact stiffness S and plastic depth h_r is experimentally observed [79] justifying this approach. The assumption implicit in this method is that the linear relation

between S and h_r applies so far as the indenter with its conical shape is in contact with the material. By extrapolating the linear trend to zero-contact stiffness, one obtains a negative value of plastic depth. The absolute value of which, h_0 , is assumed to represent the tip defect, i.e., the missing portion of the tip, representing the deviation of the indenter from a perfect cone. Loubet [78, 79] suggested to add the tip defect h_0 to the plastic depth h_r in order to obtain the contact depth h_c . An empirical correction factor α , of 1.2 was found to represent well all the experimental conditions tested:

$$h_c = \alpha (h_r + h_0). \quad (7.8)$$

On this basis, the contact area is calculated by

$$A = \beta \alpha^2 (h_r + h_0)^2, \quad (7.9)$$

where β is the tip-shape factor equal to 24.56 in the case of a Berkovich indenter.

The hardness can then be calculated by

$$H = L/A. \quad (7.10)$$

Concerning the elastic modulus calculation, the linear relationship between contact stiffness and plastic depth may be expressed, after ref [79], as

$$S = B(h_r + h_0). \quad (7.11)$$

From Sneddon's equation (7.4),

$$S = 2E_r \sqrt{A/\pi}. \quad (7.12)$$

So that, from (7.9) and (7.12), one obtains

$$E_r = \frac{\sqrt{\pi} B}{2\alpha\sqrt{\beta}}. \quad (7.13)$$

Few examples showing this analysis are reported in Fig. 7.5, where plastic depths are plotted against measured contact stiffness for a variety of polymers where the slopes, B , are correctly related to the materials of Young's moduli. The same plot for fused silica, shown in Fig. 7.6, is particularly interesting to validate the technique. Indeed, the same nanoindentations on fused silica are used as a reference either to calibrate the area function according to the O&P procedure or h_0 in the case of the method suggested by Loubet. Unfortunately, the value of h_0 which should account for the so-called tip defect, takes different values for the four polymers shown in Fig. 7.5 and for the fused silica data of Fig. 7.6.

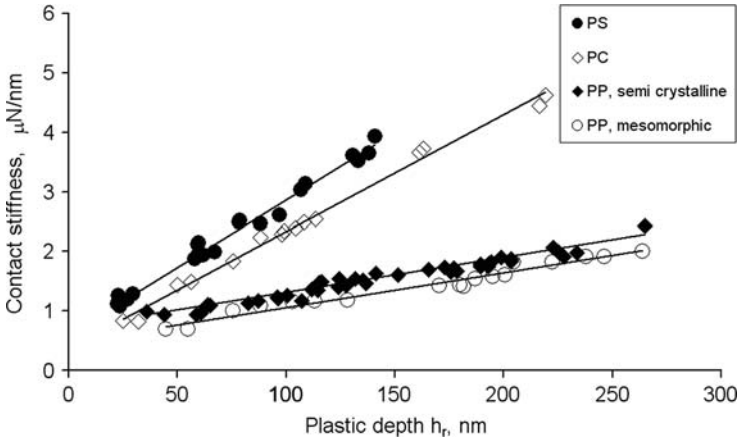


Fig. 7.5. Plots of (7.11), used to evaluate the Young’s modulus of the sample (E_r) and the indenter tip defect (h_0), for a broad range of material stiffness and behavior

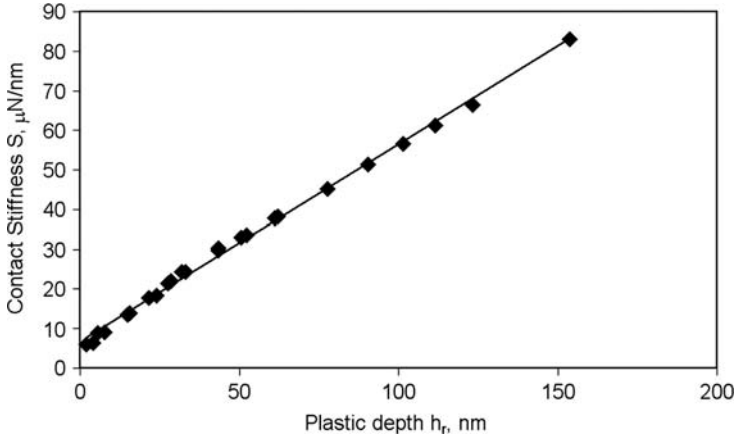


Fig. 7.6. Plots of (7.11), used to evaluate the Young’s modulus of the sample (E_r) and the indenter tip defect (h_0), for a fused silica sample. Note the difference in the ordinate with respect to Fig. 7.5, for a hint of the very different experimental conditions when analyzing polymers

7.4 Modulus Estimate of a Challenging Set of Samples

The procedures suggested by Oliver and Pharr and by Loubet are compared in the following paragraphs by analyzing a challenging series of samples, obtained from different polymers or, for the same material, differing for the latitude of morphology obtained upon cooling from the melt. We will observe that the O&P procedure correctly scales the Young’s moduli of polymers with very different mechanical properties, but does not provide a reliable order of magnitude of their values. A phenomenological correction factor to

the O&P procedure has therefore been introduced for this purpose [18] and its application to the polymers studied in this work is also outlined in the following.

Assuming the tests are performed so as to minimize viscoelastic contributions, as discussed above, the desired correction factor can be first defined as the ratio of the elastic modulus of the sample and the one measured from the O&P procedure

$$\chi \equiv \frac{E_{\text{sample}}}{E_{\text{O\&P}}}. \quad (7.14)$$

Therefore, χ accounting of the overestimate the O&P procedure gives rise was found to be dependent, at fixed loading rate, on penetration depth and on the sample elastic modulus itself. A convenient form for χ was given by

$$\chi = 1 + \frac{ap^b}{E}. \quad (7.15)$$

For metals and ceramics, i.e., materials with high moduli, χ reduces to unity and the correction can be neglected. The power law dependence on penetration depth also removes the so-called Indentation Size scale Effect, ISE. The existence of ISE is well known and it has been justified for metals as due to dislocations movement [80–83]. The origin of ISE for polymers is more debatable and sometimes, takes unexpected forms. For example, Fig. 7.7 shows the variation of elastic modulus with penetration depth in the case of ultra-high,

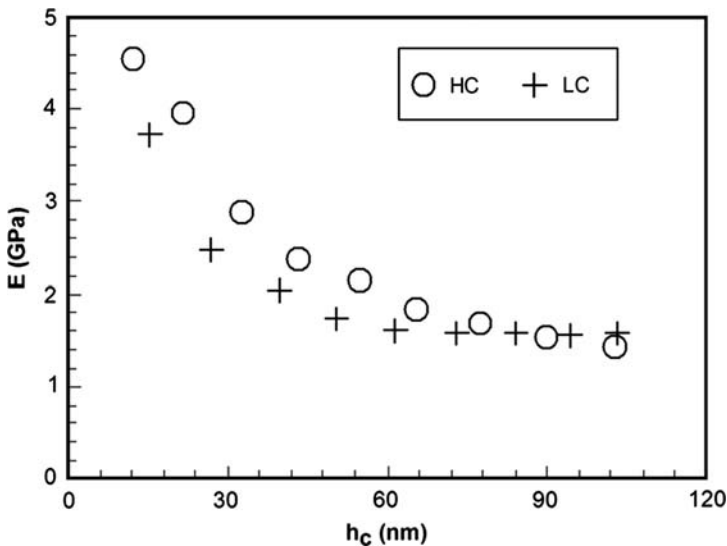


Fig. 7.7. Indentation of size-scale effect, showing a decreasing elastic modulus evaluation with an increasing penetration depth, in the case of two UHMWPE samples with different crystallinity. Reproduced from [47] with permission of Elsevier

molecular weight poly(ethylene) samples, nanoindented with an AFM–FT [47]. Two different size-scale effects can be noticed for high (HC) and low (LC) crystallinity samples, though the “asymptotic” values converge to the same plateau. The force curves, however, exhibit differences between the two samples – not in harmony with the asymptotic trend – neither with the outlook that modulus, depending on morphology, should be expected to vary first of all with crystallinity.

The quantities to be calibrated in (7.15) are the parameters, a and b . Assuming for the moment a new form of (7.15) as

$$\chi = 1 + \frac{B(p)}{E}. \quad (7.16)$$

It is possible to show the dependence of χ on modulus, where penetration depth is now a parameter for the series, while the second factor altogether showing the order of magnitude of the overestimate with respect to the O&P procedure. As the elastic modulus is present with an exponent equal to 1, using several nanoindentations performed at different penetration depths on a single polymer of known modulus which is used as a calibration reference, allows a fit of the dependence of the parameter B on penetration depth, thus providing the calibration of the a and b parameters. For the case of Fig. 7.8, a and b are respectively 1.134×10^{10} and -0.485 with penetration depth and modulus measured in nm and GPa.

A comparison of the Young’s moduli evaluated by means of the above-mentioned three procedures is shown in Fig. 7.9, in the case of amorphous glassy polymers; poly(carbonate) and pol(ystyrene); semi-crystalline ones;

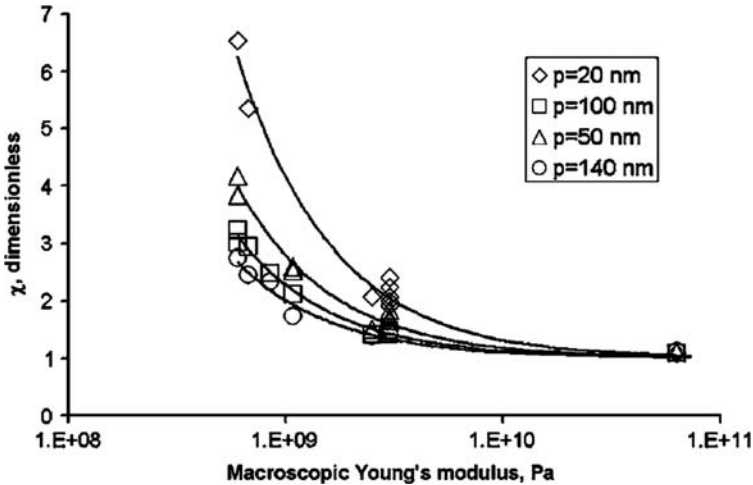


Fig. 7.8. Fit of experimental data following (7.16), in order to evaluate the calibration parameters a and b of (7.15). See text for more details. Reproduced from [18] with permission of AIP

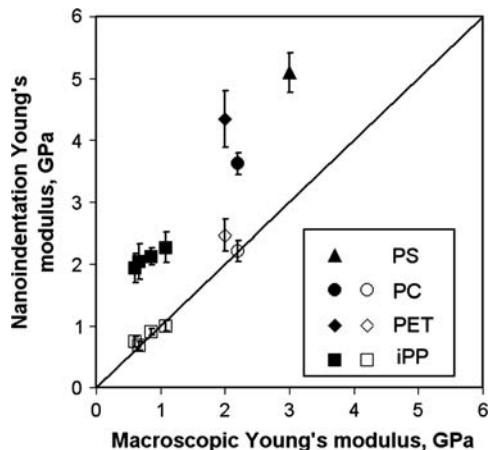


Fig. 7.9. Comparison between Young's moduli obtained by macroscopic tests and by the O&P procedure, with (empty series) and without (filled series) the χ correction factor. Reproduced from [18] with permission of AIP

poly(ethylene), poly(ethylene terephthalate), and poly(propylene); and a mesomorphic one, poly(propylene) fast; cooled from the melt at 888 K/s.

Figures 7.8 and 7.9 clearly show that the O&P procedure overestimates the modulus of the materials examined in this work, in a range of 170–300% times. This failure is commonly attributed to pile-up or viscoelastic effects [17].

In general, a pile-up can be accounted for once a scan of the deformed surface after the nanoindentation can be performed. The real contact area would be the sum of the one predicted by the O&P procedure, which takes into account only sink-in phenomena, and an additional term due to pile-up. Phenomenologically, one observes that pile-up occurs at the edges of the imprint giving rise to an arc around the indent edges. Zhou et al. showed that both the contact area from O&P as well as the pile-up can be expressed as a function of the contact radius [92], thus obtaining an expression for the total contact area which can be evaluated only from the contact depth as estimated by the O&P procedure. On the basis of this approach, the Young's modulus of the sample would approximately be 70% of the value obtained from the O&P procedure. Unfortunately, much larger overestimates are usually observed by even minimizing the pile-up contributions, e.g., adopting very shallow penetration depths, or viscoelastic effects using large loading rates [17, 18].

The reason for this failure has to be found in the peculiar, mechanical behavior of polymers: viscoelasticity changes the nanoindentation contact mechanics with respect to the elastic one and the unloading exponent is always larger than 2, even at very high indentation rates as Fig. 7.10 shows. A quadratic relationship between load and penetration can be considered as a bound for elastic behavior, as suggested by Sneddon; and an exponent larger than 2 would point out that O&P, as well as any other procedure derived from

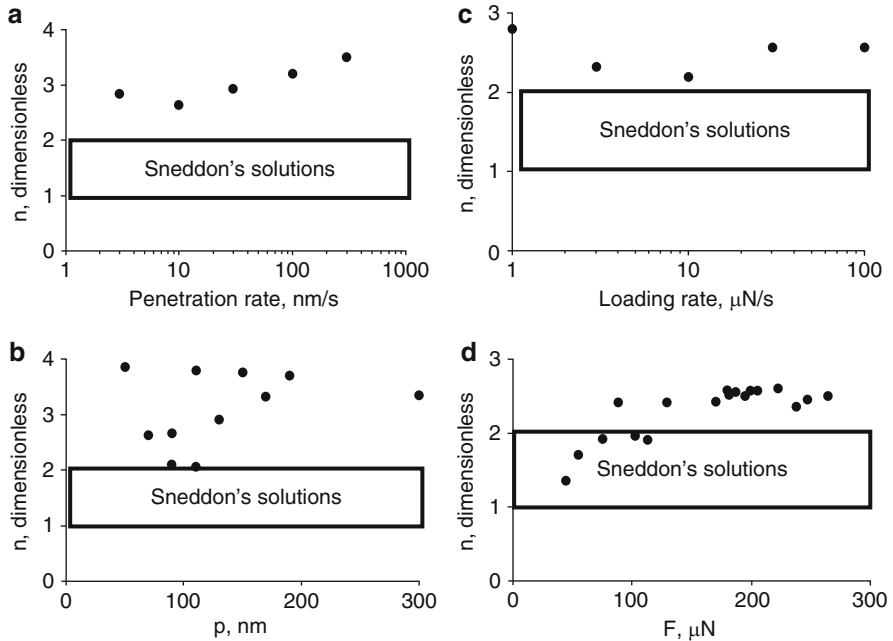


Fig. 7.10. Unloading exponents obtained from force curves on an iPP mesomorphic sample under displacement (A, B) and load (C, D) control. Dependence on penetration rate (A) or loading rate (C) and on penetration (B) and load (D) is shown and compared with the typical range expected on the assumption that an elastic contact model holds. Reproduced from [17] with permission of ACS

elastic contact mechanics, cannot be applied to interpret the load dependence on penetration on unloading, even though the typical “viscoelastic nose” is not observed in the force curve.

Another possible explanation of such deviations could be related to the occurrence of strong adhesion between the diamond tip and the polymeric sample implying a larger contact area, similarly to the above discussion about the pile-up, as shown in Fig. 7.11.

Although this explanation can be debatable in the case of polyolefin samples, where adhesion should be negligible, Grunlan et al. reported a good improvement when using indenters made of a different material, namely tungsten [93]. Figure 7.12 shows the measurements in the case of low-density poly(ethylene) and poly(styrene). The overestimation of the elastic modulus using diamond indenters is in line with the previous observations summarized in Fig. 7.9, while use of a tungsten indenter gives rise to a more reliable measurement.

A final remark, particularly important when analyzing ultrathin coatings, involves substrate effects [70, 94–97]. This consists, unlike the ISE, in an increasing modulus evaluation with increasing penetration depth in the case

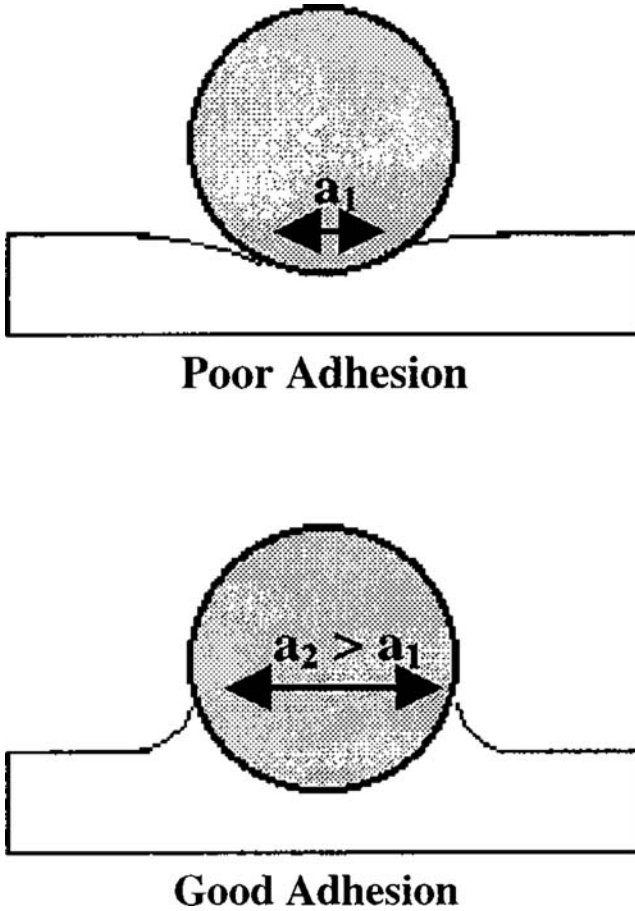


Fig. 7.11. Schematic diagram showing differences in contact size between situations with poor tip–sample adhesion and good tip–sample adhesion. Reproduced from [93] with permission of AIP

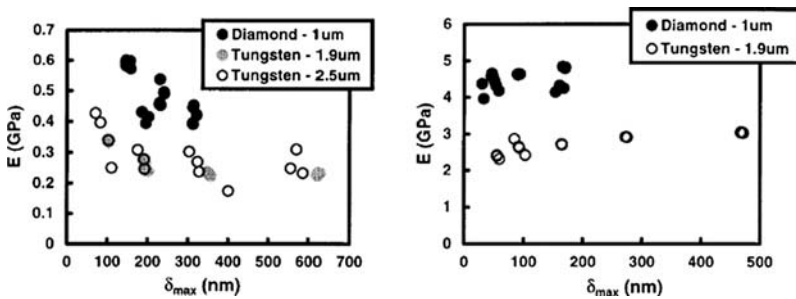


Fig. 7.12. Elastic modulus as a function of maximum penetration depth for indents into bulk LDPE (on the *left*) and PS (on the *right*), using a diamond tip and two tungsten tips with different radii. Reproduced from [93] with permission of AIP

of soft sample/hard substrate, and it is due to the fact that the stress field extends to the substrate biasing the measurement, since the test is referred to the layered system.

Indeed, the indentation response of a thin film on a substrate is a complex function of the elastic and plastic properties of both the film and the substrate. Numerous investigators [98–103] have used both experimental and theoretical methods to study the problem of extracting “true” film properties from nanoindentation data for film/substrate composites. Due to the severe constraint imposed on the plastic deformation of the film by the relatively undeformable substrate, the material piles up on the sides of the indenter leading to erroneously high, apparent hardness. The strain gradient theory, proposed by Chen and Wang [104,105] was used to study the substrate effect in the case of Al/glass systems, and it has been described well by numerical simulations [106]. An alternative mathematical approach was reported by Schewarzer et al. [107], which allows to find the exact solution for the displacement and stress distribution in the case of a coated half-space under Hertzian load. The case of a stiff coating on a soft substrate and the corresponding inverse case have been studied as well [94–97].

A common rule of thumb prescribes the use of penetration depths smaller than 10% of the total film thickness. This rule is sometimes criticized by several authors [70], suggesting the use of even smaller penetration. However, Fig. 7.13 shows the elastic modulus evaluated from nanoindentations

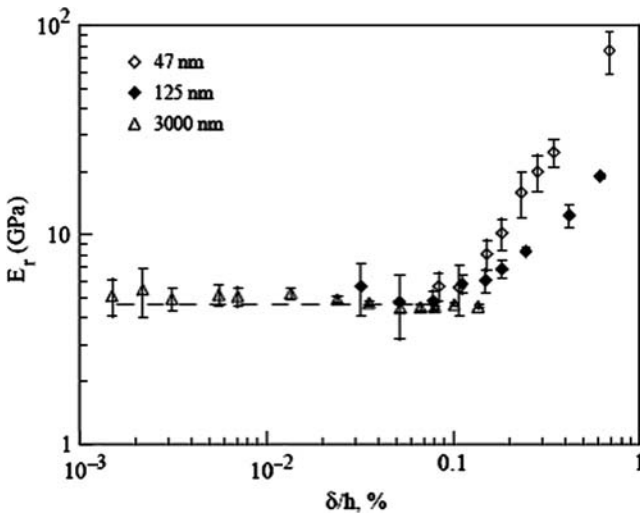


Fig. 7.13. Dependence of the reduced contact modulus on the indentation depth, for dipentaerythritol pentaacrylate monomer, crosslinked by 5%.-wt photoinitiator, 1-hydroxy-cyclohexylphenyl-ketone, showing a roughly constant “substrate effect” behavior for a broad range of film thickness. The abscissa represents the ratio of penetration depth and film thickness. Reproduced from [55] with permission of Elsevier

performed at different penetration depths on a coating deposited with a broad range of thicknesses. In all cases considered, the Young's modulus starts to increase significantly after one-tenth of the total thickness.

The surface mechanical properties are often difficult to study due to the inherently small thickness of the layer which bears different properties with respect to the bulk substrate. Nanoindentation obviously represents a good tool to this aim. Poly (carbonate) is a typical example of a polymer, whose toughness and transparency coupled with low density makes it a widely used structural material. Mailhot et al. [108] studied the effect of UV irradiation of PC, showing an increase of mechanical properties upon irradiation, probably due to chain reorientation and the onset of hydrogen bonds which give rise to an increase of cohesion. Besides, the effect of surface modification by thin ceramic layers on UV-irradiation response was also studied, and it was made possible only by the use of AFM-FT nanoindentations. Figure 7.14 shows the improvement of mechanical properties when coating PC with ZnO and Al_2O_3 , or with a combination of these oxides. The use of Al_2O_3 stems on its high wear resistance, and the steeper force curve clearly records this feature with respect to a ZnO layer of the same thickness. However, alumina is photocatalytically inactive and therefore, it is preferred to couple with a thin layer of ZnO. This material indeed absorbs UV though it even promotes photochemical reactions for the underlying polymeric system at the polymer interface; therefore, a layer sequence of $\text{Al}_2\text{O}_3/\text{ZnO}/\text{Al}_2\text{O}_3$ gives the best combination of photo-protection and wear-resistance enhancement. In addition, the ceramic layers also act as a barrier toward oxygen. The differences among the various systems

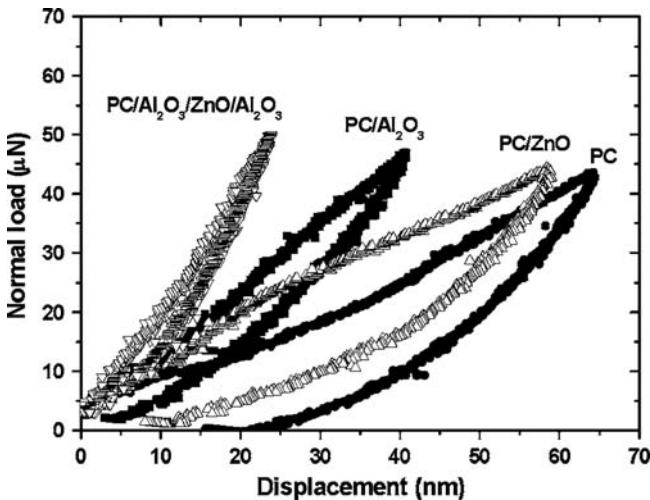


Fig. 7.14. Nanoindentation force curves of PC with and without single- and three-layer coatings. Thickness of each layer is 50 nm. Reproduced from [108] with permission of AIP

are clearly discerned by nanoindentations, which proved that not only the mechanical properties were enhanced, but the sensitivity to UV irradiation of the polymer substrate was negligible.

A further example of a challenging problem, where a very high sensitivity on determining the mechanical properties on the nanoscale is requested is given by sets of polymer samples with different morphologies. This can find a suitable application on mapping the mechanical properties of items where the complex interplay of pressure, temperature gradients, and shear stress gives rise to inhomogeneous samples with a 1D gradient of the morphology developed. Morphology changes can take place over relatively small distances making nanoindentations as the ideal tool for the characterization of the spatial assembly of the morphology developed and its relation to the processing conditions adopted during solidification.

To develop such a tool, it is necessary to understand if nanoindentations can discriminate slight differences in mechanical properties as determined by differences brought by morphological changes alone. For this purpose, a set of model samples was prepared so as to ensure a homogeneous morphology throughout a volume suitable for nanomechanical as well as macroscopic tests to be used as reference. The morphology developed [56, 109] was tuned with the metallurgical approach which was already described so as to mimic the structure developed in a real processing route, i.e., experiencing similar solidification conditions, albeit in a real system they may eventually change in space; such changes being represented by a set of different samples with stepwise changing morphologies. These samples have been studied in the past by means of AFM nanoindentations, showing that it was possible to accurately distinguish among different morphologies, from semi-crystalline to amorphous or mesomorphic, obtained on solidification from the melt at different cooling rates [109]. It is worth noticing that the contact mechanics during AFM nanoindentations sensibly differs from the case of the AFM-FT experiments.

The same samples were tested by the AFM-FT and the results are reported in Fig. 7.15 as a plot of cooling rate, i.e., structure evolution vs. nanoscale elastic modulus, as determined by several procedures: the approach due to Loubet, the O&P, and the phenomenological correction factor as discussed above. Moreover, they are compared with the macroscopic values of the elastic modulus measured with a standard, tensile-testing machine. The comparison shows the pitfalls and limitations of the methods so far discussed to extract mechanical properties from the indentations on polymer samples. The slight differences in elastic moduli are well described with an agreement quite satisfactory in the case of the Loubet et al. procedure and of the phenomenological correction to the O&P method. These results point out the possibility to discriminate subtle structural features and the related slight differences of mechanical properties, thus further opening the potential of nanoindentations once an instrument with suitable lateral resolution is available. So far, both the AFM-based cantilever instruments and the stiffer-sensor DSI bear several limitations when nanoindentation tests on compliant viscoelastic materials are

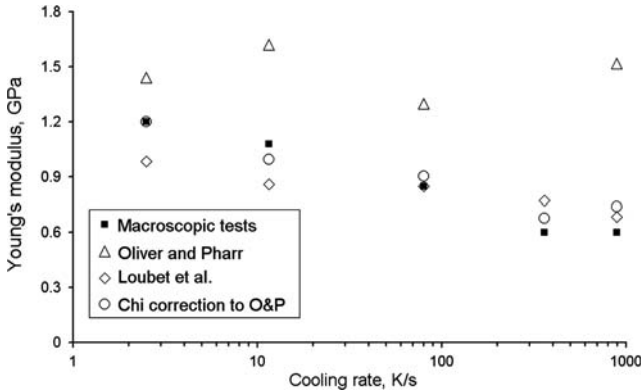


Fig. 7.15. Young's modulus values obtained from nanoindentations, as evaluated from the procedures described in [14] and [79], and after a further calibration (*circles*) [18], showing a good agreement with values determined from macroscopic tests (*black symbols*) on sets of samples of fine-tuned morphology and crystallinity

performed. Depending on the application, one is forced to choose either the larger lateral resolution offered by the first or the better accuracy provided by the latter family of instruments.

References

1. P. Fratzl, R. Weinkamer Prog. Mater. Sci. **52** 1263 (2007)
2. M.A. Meyers, P.Y. Chen, A.Y.M. Lin, Y. Seki Prog. Mater. Sci. **53** 1 (2008)
3. F.J. Baltà-Calleja Adv. Polym. Sci. **66** 117 (1985)
4. A. Flores, F.J. Baltà-Calleja, D.C. Bassett J. Polym. Sci. B **37** 3151 (1999)
5. F.J. Baltà-Calleja, A. Flores, F. Ania, D.C. Bassett J. Mater. Sci **35** 1315 (2000)
6. A. Flores, F.J. Baltà-Calleja, G.E. Attenburrow, D.C. Bassett Polymer **41** 5431 (2000)
7. S. Fakirov, B. Krasteva J. Macromol. Sci. Phys. B **39** 297 (2000).
8. Y.P. Zheng, A.F.T. Mak, A.K.L. Leung J. Rehabil. Res. Dev. **38** 487 (2001)
9. L.A. Setton, V.C. Mow, F.J. Muller, J.C. Pita, D.S. Howell J Orthopaed Res **12** 451(1994)
10. K.A. Athanasiou, M.P. Rosenwasser, J.A. Buckwalter, T.I. Malinin, V.C. Mow J. Orthopaed. Res. **9** 330 (1991)
11. J.H. Kinney, S.J. Marshall, G.W. Marshall Crit. Rev. Oral Biol. Med. **14** 13 (2003)
12. P.E. Riches, N.M. Everitt, D.S. McNally J. Biomech. **33** 1551 (2000)
13. A.J. Rapoff, R.G. Rinaldi, J.L. Hotzman, D.J. Daegling Am. J. Phys. Anthrop. **135** 100 (2008)
14. W.C. Oliver, G.M. Pharr J. Mater. Res. **7** 1564 (1992)
15. W.C. Oliver, G.M. Pharr J. Mater. Res. **19** 3 (2004)
16. D. Tranchida, S. Piccarolo Macr. Rap. Commun. **26** 1800 (2005)

17. D. Tranchida, S. Piccarolo, J. Loos, A. Alexeev *Macromolecules* **40** 1259 (2007)
18. D. Tranchida, S. Piccarolo, J. Loos, A. Alexeev *Appl. Phys. Lett.* **89** 171905 (2006)
19. T.C.T. Ting *J. Appl. Mech.* **33** 845 (1966)
20. Y.T. Cheng, C.M. Cheng *J. Mater. Res.* **20** 1046 (2005)
21. L. Cheng, X. Xia, L.E. Scriven, W.W. Gerberich *Mech. Mater.* **37** 213 (2005)
22. M.L. Oyen, R.F. Cook *J. Mater. Res.* **18** 139 (2003)
23. C.C. White, M.R. Van Landingham, P.L. Drzal, N.K. Chang, S.H. Chang *J. Polym. Sci. B* **43** 1812 (2005)
24. M.R. Van Landingham, N.K. Chang, P.L. Drzal, C.C. White, S.H. Chang *J. Polym. Sci. B* **43** 1794 (2005)
25. C.A. Tweedie, K.J. Van Vliet *J. Mater. Res.* **21** 1576 (2006)
26. M. Fujikane, M. Leszczyski, S. Nagao, T. Nakayama, S. Yamanaka, K. Niihara, R. Nowak *J Alloy Comp.* **450** 405 (2008)
27. B. Cappella, D. Silbernagl *Thin Solid Films* **516** 1952 (2008)
28. B. Bhushan, X. Li *Int. Mater. Rev.* **48** 125 (2003)
29. S. Umemura, Y. Andoh, S. Hirono, T. Miyamoto, R. Kaneko *Philos. Mag. A* **74** 1143 (1996)
30. S. Lafaye, M. Troyon *Wear* **261** 905 (2006)
31. B. Cappella, G. Dietler *Surf. Sci. Rep.* **34** 1 (1999)
32. H.J. Butt, B. Cappella, M. Kappl *Sur. Sci. Rep.* **59** 1 (2005)
33. D. Tranchida, S. Kiflie, Z. Piccarolo '*Atomic Force Microscope Nanoindentations to Reliably Measure the Young's Modulus of Soft Matter*' in *Modern Research and Educational Topics in Microscopy* (Formatex, Badajoz, 2007)
34. V.V. Tsukruk, V.V. Gorbunov, Z. Huang, S.A. Chizhik *Polym. Int.* **49** 441 (2000)
35. H.W. Hao, A.M. Barò Saenz *J. Vac. Sci. Technol. B* **9** 1323 (1991)
36. N.A. Burnham, X. Chen, C.S. Hodges, G.A. Matei, E.J. Thoreson, C.J. Roberts, M.C. Davies, S.J.B. Tandler *Nanotechnology* **14** 1 (2003)
37. D. Tranchida, S. Piccarolo, M. Soliman *Macromolecules* **39** 4547 (2006)
38. S.A. Chizhik, Z. Huang, V.V. Gorbunov, N.K. Myshkin, V.V. Tsukruk *Langmuir* **14** 2606 (1998)
39. V.V. Tsukruk, Z. Huang, S.A. Chizhik, V.V. Gorbunov *J. Mater. Sci.* **33** 4905 (1998)
40. K. Herrmann, N.M. Jennett, W. Wegener, J. Meneve, K. Hasche, R. Seemann *Thin Solid Films* **377** 394 (2000)
41. M.S. Bischel, M.R. VanLandingham, R.F. Eduljee, J.W. Gillespie, J.M. Schultz *J. Mater. Sci.* **35** 221 (2000)
42. B. Du, J. Liu, Q. Zhang, T. He *Polymer* **42** 5901 (2001)
43. V.V. Tsukruk, A. Sidorenko, H. Yang *Polymer* **43** 1695 (2002)
44. B.D. Beake, G.J. Leggett *Polymer* **43** 319 (2002)
45. B. Bhushan, A.V. Kulkarni, W. Bonin, J.T. Wyrobek *Philos. Mag. A* **74** 1117 (1996)
46. M.E. Barbour, R.P. Shellis *Phys. Med. Biol.* **52** 899 (2007)
47. K.S. Kanaga Karuppiah, A.L. Bruck, S. Sundararajan, J. Wang, Z. Lin, Z.H. Xu, X. Li *Acta Biomater.* **4** 1401(2008)
48. T. Chen, S. Yao, K. Wang, H. Wang *Nucl. Instrum. Methods Phys. Res. B* **266** 3091 (2008)
49. J.D. Torrey, R.K. Bordia *J. Eur. Ceram. Soc.* **28** 253 (2008)

50. M. Nowicki, A. Richter, B. Wolf, H. Kaczmarek *Polymer* **44** 6599 (2003)
51. S. Tajima, K. Komvopoulos *J. Appl. Phys.* **101** 014307 (2007)
52. R. Bandorf, D.M. Paulkowski, K.I. Schiffmann, R.L.A. Kuster *J. Phys. Condens. Matter* **20** 354018 (2008)
53. K.R. Morasch, D.F. Bahr *Thin Solid Films* **515** 3298 (2007)
54. K.I. Schiffmann, A. Hieke *Wear* **254** 565 (2003)
55. K.B. Geng, F.Q. Yang, T. Druffel, E.A. Grulke *Polymer* **46** 11768 (2005)
56. V. Brucato, S. Piccarolo, V. La Carrubba *Chem. Eng. Sci.* **57** 4129 (2002)
57. B. Wolf *Cryst. Res. Technol.* **35** 377 (2000)
58. Z. Xiangyang, J. Zhuangde, W. Hairong, Y. Ruixia *Mater. Sci. Eng. A* **488** 318 (2008)
59. Y.H. Lee, J.H. Hahn, S.H. Nahm, J.I. Jang, D. Kwon *J. Phys. D Appl. Phys.* **41** 074027 (2008)
60. Y.H. Lee, U. Baek, Y.I. Kim, S.H. Nahm *Mater. Lett.* **61** 4039 (2007)
61. R. Garcia, R. Perez *Surf. Sci. Rep.* **47** 197 (2002)
62. T.H. Fang, W.J. Chang *Microel. J.* **35** 595 (2004)
63. J.S. Villarrubia *J. Res. NIST* **102** 425 (1997)
64. D. Tranchida, S. Piccarolo, R.A.C. Deblieck *Meas. Sci. Technol.* **17** 2630 (2006)
65. I. Hiroshi, F. Toshiyuki, I. Shingo *Rev. Sci. Instrum.* **77** 103704 (2006)
66. D. Tranchida, Z. Kiflie, S. Piccarolo *Macromolecules* **40** 7366 (2007)
67. I. Karapanagiotis, D.F. Evans, W.W. Gerberich *Polymer* **43** 1343 (2002)
68. I.N. Sneddon *Int. J. Eng. Sci.* **3** 47 (1965)
69. Z. Bartzczak *Polymer* **46** 10339 (2005)
70. A.C. Fischer Cripps *Nanoindentation* (Springer, New York, 2002)
71. A. Pawlak, A. Galeski *Macromolecules* **38** 9688 (2005)
72. A.J. Bushby, D.J. Dunstan *J. Mater. Res.* **19** 137 (2004)
73. Y. Zhou, P.K. Mallick *Polym. Eng. Sci.* **42** 2449 (2002)
74. R.E. Robertson *J. Appl. Polym. Sci.* **7** 443 (1963)
75. J.A. Roetling *Polymer* **6** 311 (1965)
76. T. Kazmierczak, A. Galeski, A.S. Argon *Polymer* **46** 8926 (2005)
77. D. Tranchida, Z. Bartzczak, D. Bielinski, A. Galeski, S. Piccarolo *Polymer* **50** 1939 (2009)
78. J.L. Loubet, J.M. Georges, J. Meille *Nanoindentation Techniques in Materials Science and Engineering* (ASTM, Philadelphia, 1986)
79. G. Hochstetter, A. Jimenez, J.L. Loubet *J. Macromol. Sci. B Phys.* **38** 681 (1999)
80. A. Carpinteri, S. Puzzi *Eng. Fract. Mech.* **73** 2110 (2006)
81. R.K. Abu Al-Rub *Mech. Mater.* **39** 787 (2007)
82. M. Zhao, W.S. Slaughter, M. Li, S.X. Mao *Acta Mater.* **51** 4461 (2003)
83. D. Chicot *Mat. Sci. Eng. A* **499** 454 (2009)
84. J. Boussinesq *Applications des Potentiels à l'Etude de l'Equilibre et du Mouvement des Solides Elastiques.* (Gauthier-Villars, Paris, 1885)
85. A.E.H. Love *Philos. Trans.* **228** 377 (1929)
86. A.E.H. Love *Q. J. Math.* **10** 161 (1939)
87. H. Hertz *Miscellaneous Papers* (MacMillan, New York, 1896)
88. L.D. Landau, E.M. Lifshitz *Theory of Elasticity* (Pergamon Press, Oxford, 1986)
89. S.I. Bulichev *Zavodsk Lab.* **53** 76 (1987)
90. G.M. Pharr, W.C. Oliver, F.R. Brotzen *J. Mater. Res.* **7** 613 (1992)

91. Y.T. Cheng, C.M. Cheng *J. Appl. Phys.* **84** 1284 (1998)
92. J. Zhou, B. Berry, J.F. Douglas, A. Karim, C.R. Snyder, C. Soles *Nanotechnology* **19** 495703 (2008)
93. J.C. Grunlan, X.Y. Xia, D. Rowenhorst, W.W. Gerberich *Rev Sci Instr* **72** 2804(2001).
94. L. Calabri, N. Pugno, A. Rota, D. Marchetto, S. Valeri *J. Phys. Condens. Matter.* **19** 395002 (2007)
95. S.B. Liu, A. Peyronnel, Q.J. Wang, L.M. Keer *Trib. Lett.* **18** 303 (2005)
96. R. Saha, Z.Y. Xue, Y. Huang, W.D. Nix *J Mech. Phys. Solids* **49** 1997 (2001)
97. T.H. Fang, W.J. Chang *Micro Eng.* **65** 231 (2003)
98. P.J. Burnett, D.S. Rickerby *Thin Solid Films* **148** 51 (1987)
99. P.J. Burnett, D.S. Rickerby *Thin Solid Films* **148** 41 (1987)
100. M.F. Doerner, D.S. Gardner, W.D. Nix *J. Mater. Res.* **1** 845 (1986)
101. B.D. Fabes, W.C. Oliver, R.A. McKee *J. Mater. Res.* **7** 3056 (1992)
102. D. Stone, W.R. LaFontaine, P.J. Alexopoulos *J. Mater. Res.* **3** 141 (1988)
103. T.Y. Tsui, W.C. Oliver, G.M. Pharr *Mater. Res. Soc. Symp. Proc.* **436** 147 (1997)
104. S.H. Chen, T.C. Wang *Eur J Mech A Solids* **20** 739 (2001)
105. S.H. Chen, T.C. Wang *Int. J. Solids Struct.* **39** 1241 (2002)
106. S.H. Chen, L. Liu, T. Wang *Acta Mater.* **52** 1089 (2004)
107. N. Schwarzer, F. Richter, G. Hecht *Surf. Coat. Technol.* **114** 292 (1999)
108. B. Mailhot, A. Rivaton, J.L. Gardette, A. Moustaghfir, E. Tomasella, M. Jacquet, X.G. Ma, K. Kornvopoulos *J. Appl. Phys.* **99** 104310 (2006)
109. D. Tranchida, S. Piccarolo *Polymer* **46** 4032 (2005)

Static and Dynamic Structural Modeling Analysis of Atomic Force Microscope

Yin Zhang and Kevin D Murphy

Summary. As a cantilever structure, atomic force microscope (AFM) can be either modeled as a beam, plate or a simple one degree-of-freedom (DOF) system depending on its geometry and application scenario. The AFM structure can experience the deformation shapes of vertical bending, lateral bending, torsion, extension and couplings between these four deformations depending on the excitation mode. As a small structure of micron scale, forces like van der Waals (vdW) force, surface stress, electrostatic force and residual stress can have significant influence on the AFM deflection. When the AFM tip is in contact with the sample surface, different contact mechanics models are needed depending on the tip geometry, AFM operating mode and tip, sample surface material properties. In dynamic mode, the AFM tip-sample surface intermittent contact is a complicated nonlinear dynamics problem. As a powerful tool, the AFM application is already beyond the stage of being used to image the sample surface topography. Nowadays, AFM is used more often to extract the sample materials properties such as Young's modulus, surface energy/adhesion and viscosity. How to properly model the AFM structure with different deformations and their coupling under different forces and the tip-sample surface interaction is vital to linking the experimentally measured data correctly with the sample surface material properties. This chapter reviews the different models concerning the AFM structure (static) deformations, external residual forces modeling, tip-surface contact and the AFM dynamics. This chapter is intended to provide a comprehensive review rather than an in-depth discussion on those models. Because there are too many factors influencing the experimentally measured data during the application of AFM, it is extremely difficult to consider all these factors in one model for AFM if not impossible. Because there are too many factors influencing the AFM deformations/dynamics, it will be extremely difficult if not impossible to link all of the influencing factors to the experimental data. Therefore, in the modeling aspect, certain assumptions must be made to render the problem solvable. One of the major purposes of this chapter is to discuss and analyze the assumptions of those models and by doing so we try to outline the applicability ranges of those models. At the same time by analyzing the assumptions of the models applied to the AFM, the limitations of some models are also presented. Pointing out the limitations of those models which work fine with certain application scenarios is intended to make the applicability ranges of the models clearer and also helps to better interpret the

experimental data. Only the dominant factors should be considered in a model and the other factors must be neglected to have a workable model. However, for different AFM applications the dominant factors are varied and thus transferring the model developed for one AFM application to another one can be inappropriate or even wrong. The analysis on the model assumption thus plays an important role of applying one model developed for certain application scenario to other applications.

Key words: Structural analysis, Modeling, Contact, Adhesion, Hysteresis, Impact, System response.

8.1 Introduction

We start our introduction by quoting the following words by Hofer et al. [1].

In principle, it is the goal of theory to develop a reliable model that can be used to interpret experimental images without resorting to direct simulation. However, this goal has not yet been achieved with either SPM technique. (SPM: scanning probe microscope)

It is an incisive and accurate comment on the SPM technique. Interpreting the experimental data can never, ever be treated as a trivial and simple task. This chapter aims to offer such analyses and review of structural modeling “reliability” on atomic force microscope (AFM), which is a member of SPM/SFM family (SFM: scanning force microscope). The difficulties of the structural modeling and complicated AFM response, especially dynamic one, are also presented at the same time. Structural analysis is the direct link of the sample surface topography/material properties with the measured experimental data, for example, force–displacement data in AFM static operating mode; phase shift, response frequency spectrum, amplitude, etc. in the AFM dynamic operating mode. The sample surface topography or material properties are extracted from those data. Modeling the AFM structure and the forces exerted on AFM is essential to the interpretation of the experimental data and the evaluation of the sample surface material properties. So, in that sense, the words “structural modeling” in the title include not only the modeling AFM geometry structure itself but also the modeling of the forces acting on AFM, which generally are also dependent of the AFM geometry structure. For those forces like van der Waals (vdW), electrostatic force, its magnitude depends on the tip–sample parameters like the gap distance, shape and other material properties. The resulting total force acting on the AFM is thus mingled with those forces. The AFM image is a complex convolution of the tip and the surface parameters. Therefore, the features of the sample surface become diluted by the interaction forces. This kind of convolution is one of physical restrictions for AFM to achieve subnanometer resolution. The convolution is unavoidable but there are good chances to reconstruct rather true image from diluted one using special mathematical methods. Such reconstruction is called

deconvolution, which fundamentally is based on the structural analysis on the tip–sample and the AFM cantilever–sample interaction forces. Other external forces acting on AFM structure can be the forces like surface stress, which is independent of the tip–sample distance. The direct effect of surface stress causes the AFM cantilever to bend and/or rotate [2,3], thus to affect the AFM displacement measurement.

The AFM tip is fixed at or around the free end of the AFM cantilever. Structural analysis of AFM cantilever with different geometries is the fundamental and pivotal part of properly interpreting the experimental data. We first start with the statics and then dynamics in this chapter. The AFM tip is directly responsible for AFM imaging and thus modeling of tip and tip–sample interaction is very important in this analysis. For the tip–sample interactions, we only focus on the physical ones. Chemical and biochemical types of interactions are omitted. The tip–sample contact mechanics and the impact dynamics are the major interaction topics discussed in this chapter. Besides the physical restrictions, the technical restrictions like the tip geometry and feedback control of AFM are also responsible for AFM to achieve subnanometer resolution. With the application of carbon nanotube (CNT) as the probing tip [4–10], the resolution problems concerning the tip geometry are greatly relieved. However, the tip–sample impact dynamics of the intermittent contact can induce very complicated system response, which induces subharmonic motions and later can lead to chaotic motions. The subharmonic periodic and chaotic motions are harmful to the feedback control. Analyzing the AFM structure system response of intermittent contact can offer more insights and will be helpful to the feedback control to improve the AFM measurement accuracy. As a continuous system, the AFM structure sometimes is simplified as one degree-of-freedom (DOF) system. The one DOF model can relatively accurately capture the statics and the AFM non-contact vibration with the excitation frequency around the fundamental resonant frequency. However, with the higher excitation frequency or the coupling of the AFM deformations, or intermittent contact, the one DOF model cannot be accurate. The limitations of one DOF modeling such as its response to different excitation frequencies are also examined.

8.2 Working Principle and Modes

The atomic force microscope (AFM) is one of the family of scanning probe microscopes (SPM). AFM generally measures the vertical and horizontal deflection of the cantilever. Most AFMs today use the optical lever, a device that achieves resolution comparable to an interferometer while remaining inexpensive and easy to use. The optical lever operates by reflecting a laser beam off the back of the cantilever (see Fig. 8.1). The reflected laser beam strikes a position-sensitive photodetector consisting of four side-by-side photodiodes. A differential photodetector signal V_{dif} ($V_{dif} = S \times \delta$, S is called sensitivity

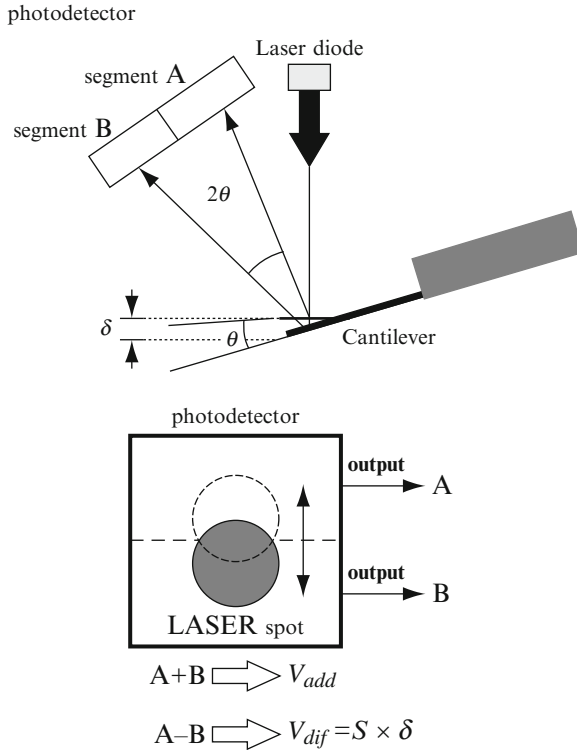


Fig. 8.1. Schematic diagram of laser-beam detection system of a commercial AFM. S is called sensitivity. A , B are the measured photodetector outputs. δ is the actual displacement of some point on cantilever (*Source: Miyatani and Fujihira [11]*)

and δ is the displacement of some point on the cantilever) output is generated and thus the displacement/deflection is measured. The deflection can then be plotted as a function of surface position to give the theoretical image of the surface [1]. The schematic diagram only shows two-dimensional scenario. AFM can be used to obtain the three-dimensional topography of the sample. The AFM horizontal/lateral deflection and twisting/rotation are also measured by the very similar way as shown for vertical deflection [12].

From the contact point of view, the operating mode of AFM can be categorized as two modes: contact and non-contact mode. In contact mode, the tip and sample are in contact, thus the atomic resolution can be achieved. But this mode may be destructive to the samples, in particular for the soft ones due to the shearing/lateral force. The non-contact mode is with frequency modulation and AFM oscillates around its fundamental resonant frequency with constant amplitude without touching the sample surface. In general, the expense of the non-contact mode is low resolution and difficult feedback conditions. From the motion point of view, there are static AFM and dynamic AFM.

Compared with static AFM, dynamic AFM can provide better signal-to-noise ratio and higher resolution in measurement of material and surface properties [12]. From the views of applications and excitations, there are some other modes called pulsed force mode [13, 14], jumping mode [15], tapping mode [7, 16] and torsional resonance mode [12, 17]. In the pulsed force mode, the force–distance curve is obtained for each imaging point and image is extracted from this curve by appropriate processing. In the pulsed force mode [13], the sample is moved up and down on a scanner under a sinusoidal modulation. The sinusoidal modulation frequency is much lower than the AFM resonant frequency and the tip and sample are in intermittent contact. For the jumping mode [15], it is designed to probe the local variation of tip–sample interaction. During the operation of AFM in jumping mode, the jump-in and jump-off phenomena are frequently encountered (see Fig. 8.2 inset). In the jumping mode, the interaction can be recorded and evaluated by only one parameter, the normal force (see Fig. 8.2), which can be displayed as an image in real time. The similarity between the topological image and adhesion image for the sample with uniform composition shows that the adhesion can be correlated with the topography [15]. For the tapping mode (TM) of intermittent contact, the cantilever vibrates as it scans the sample, and its oscillation amplitude is measured and used as feedback control parameter. Compared with continuous

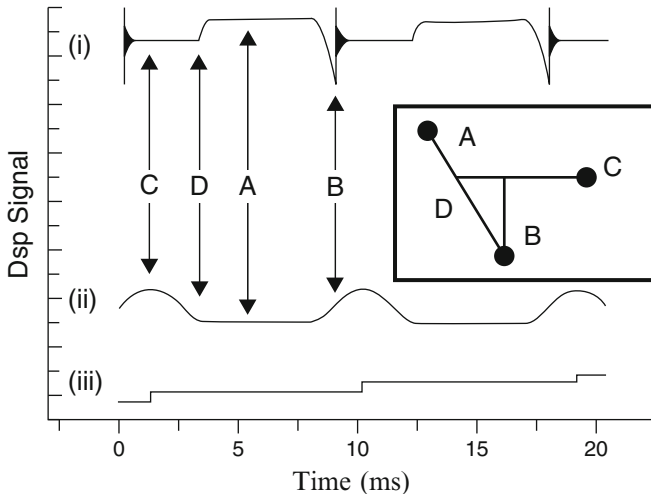


Fig. 8.2. Oscilloscope tracing showing (i) the normal force signal (cantilever deflection), (ii) the piezo signal (sample motion along Z axis), and (iii) the lateral displacement as a function of time. Region A corresponds to the feedback time, Point B to the point where tip and sample separate (“snap off”), point C to the largest tip–sample distance, and point D to the position where tip–sample contact is established (“jump in”). The inset shows a schematic composition of the normal force (i) and the Z piezo motion (ii), which is a force versus distance curve performed at each sampled point (*Source*: de Pabblo et al. [15])

contact mode, because the cantilever tip vibrates vertically, no or little lateral force is exerted on the sample surface in the tapping mode. Consequently, TM causes less damage to the soft sample during the intermittent contact [7] or cantilever structure insensitivity to lateral force gradients [17]. In contrast to the tapping mode in which the normal/transverse oscillation amplitude is exerted, the torsional oscillation amplitude is exerted by the piezoelectric holder for torsional resonance mode (TRM). The tip-sample distance for TRM keeps unchanged, which affects the measured data less compared with TM [17]. The torsional stiffness is about two orders of magnitude larger than lateral bending stiffness, which causes more sample deformation in TRM, thus better contrast for sample image can be generated [17]. For these two reasons, TRM is expected to provide better contrast in the tribological and mechanical properties of the sample near surface region as compared to the TM [17].

8.3 Statics of Atomic Force Microscope Cantilever: Effective Stiffness Approach

The most popular materials for the cantilevers are monocrystalline silicon and silicon nitride (Si_3N_4). Tungsten, nickel and other materials can also be used. There are two basic structural designs for the AFM cantilever: straight rectangular and V-shaped (chevron, triangular) cantilever as shown in Fig. 8.3. The cantilever can be modeled as either beam or plate structure. The effective stiffness approach is the way to find the effective spring stiffness for beam or plate of continuous system. The stiffness found by this approach is suitable for AFM to be used in the (quasi)static mode or the lumped mass-spring model for AFM. In many AFM dynamics applications, the higher modes of AFM bema/plate can be easily excited [19], AFM should be modeled as the continuous system of beam/plate as shown later. However, this effective stiffness approach can directly obtain some key information about the AFM structure (e.g., the fundamental resonant frequency), which is very useful for design and calibration. As shown in Fig. 8.4, there are four possible deformation shapes of an AFM: vertical bending, lateral bending, torsion and extension [12]. The couplings of these four deformations can cause complicated structural analysis. The vertical bending case is the most frequently encountered one and our analysis starts with the uncoupled case. The effective stiffness is only demonstrated for the vertical bending case, which can also be easily applied to the lateral bending case.

Currently, very few AFM can operate with the torsional resonant mode. For the detailed torsion analysis, the reader should refer to the works of Song and Bhushan [12, 19–21]. However, if AFM is with a tip eccentricity, torsion is unavoidable even for the lateral bending case or the vertical bending [21] and we also give a brief discussion on that.

When the concentrated load is applied at the beam end and other effects like torsional deformation and in-plane deflection are ignored, the AFM spring

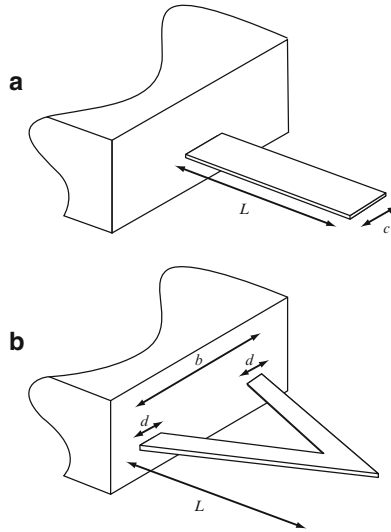


Fig. 8.3. The Schematics of (a) straight rectangular cantilevered structure and (b) V-shaped (triangular, chevron shaped) cantilever structure (Source: Sader and Green [18])

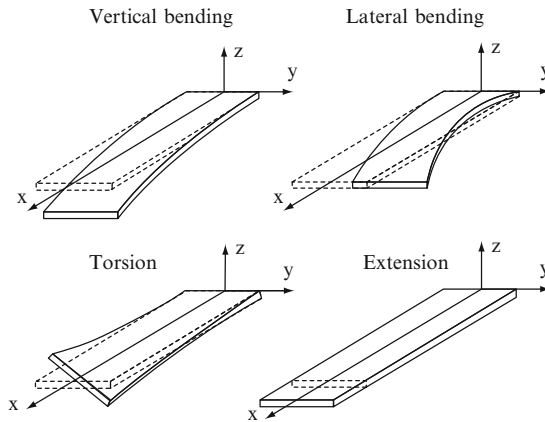


Fig. 8.4. Four deformation shapes of a rectangular cantilever with free-clamped conditions (Source: Song and Bhusan [12])

constant k (effective stiffness) for transverse deflection can be calculated by using the knowledge of mechanics of materials as follows [22, 23]:

$$k = \frac{Et^3w}{4L^3} \tag{8.1}$$

E , t , w and L are the Young’s modulus, thickness, width and length of the AFM beam, respectively. Equation (8.1) is for a straight rectangular cantilever beam structure of continuous system with the concentrated load at the end.

When there is ΔL between the concentrated load and the beam end, the equation above changes as

$$k = \frac{Et^3w}{4(L - \Delta L)^3} \quad (8.2)$$

The spring constant/effective stiffness approach is to transform the beam continuous system to one degree-of-freedom (DOF) system, which only requires measuring the concentrated load and the displacement at the load location. In practice, the difficulty of using the two equations above is that Young's modulus and dimensions of the AFM are needed and difficult to be precisely measured. The microfabrication technique often results in variable stoichiometry, or say, the problem of precisely measuring the dimensions of the cantilever, which in turn will cause considerable variation in elastic properties, especially Young's modulus [24]. The two equations above also implicitly assume that the beam is unflawed and free from surface and residual stresses. In reality, AFM structure is very sensitive to the flaws or microcracks, thus, the direct use of the formula(s) of spring constant above may be problematic for the flawed AFM structure. During the fabrication, the microstructure is exposed to harsh environment. Therefore, the surface stress and residual stress may be left. The residual stress either strengthens or decreases the structures stiffness, or even causes failures like blistering, peeling or buckling [25]. The pragmatic method of obtaining spring constant and the alternative of avoiding measuring Young's modulus and (some) dimensions of the AFM beam are to measure the static AFM beam (tip) displacement due to the (concentrated, known) load or the AFM beam (tip) amplitude response to the excitation frequency [24, 26–30]. By varying the beam thickness and mass attached at the beam tip, Cleveland et al. [26] obtain the resonant frequencies of the AFM beams. The regression of the measured experimental data makes it possible to curve-fit a line, whose slope is the spring constant, and can be projected to predict the unmeasured AFM beam spring constant. Sader et al. [24] improve Cleveland's method of measuring spring constant of the cantilever [26], which requires the information about the effective mass attachment at the cantilever tip to determine the spring constant. Both the methods of Cleveland et al. [26] and Sader et al. [24] are suitable for determining the spring constant of the arbitrary-shaped thin cantilever.

For a V-shaped AFM cantilever, the parallel beam approximation (PBA) method is proposed to model the structure [22, 31]. Although Butt et al. [32] doubt the accuracy of PBA method on the V-shaped AFM cantilever, Sader [31] strictly proves the validity of the PBA method. Applying simple rotation transformation of coordinate system of the skewed beam/plate as shown in Fig. 8.5, Sader [31] proves that the skewed beam/plate possesses the identical deflection function as that of the unskewed one. And by a symmetry analysis, Sader [31] also proves that the spring constant of the two skewed rectangular beams/plates with length (L) and width (\bar{d}) is equivalent to that of the unskewed rectangular beams/plates with length (L) and width ($2\bar{d}$), as if

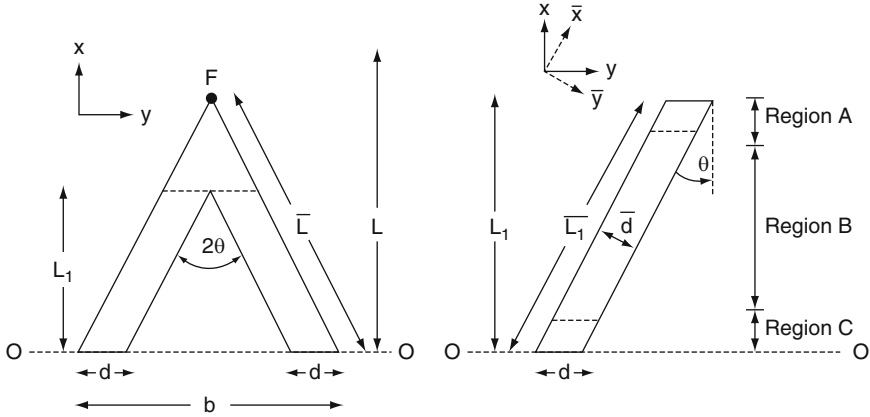


Fig. 8.5. Schematics of a V-shaped cantilevered structure coordinate systems and rotation of the coordinate systems. When the load is applied at the cantilever tip F, its spring constant k is computed by parallel beam approximation (PBA) method as follows

$$k = \frac{E I^3 \bar{d}}{2 L^3} \cos \theta \left\{ 1 + \frac{4 \bar{d}^3}{b^3} (3 \cos \theta - 2) \right\}^{-1}$$

(Source: Sader [31])

two (parallel) unskewed beams/plates with length (L) and width (\bar{d}) combine together into one. That is the adjective “parallel” word coming from. And it should also be noticed that those results above are obtained by assuming torsional effect is negligible and the beam/plate deflection is independent of coordinate y (the width dimension as shown in Fig. 8.5).

When a lateral force acts on its tip, AFM is subject to rotation as shown in Fig. 8.6. Therefore, the torsional effect of AFM beam due to the rotation has to be considered and this effect is widely investigated [33–41]. In the statics of effective stiffness method, torsional spring constant is added as a second DOF. With the consideration of the torsional effect, there is a rather astonishing and interesting comparison conclusion reported by Sader [40] that V-shaped cantilever structure design for AFM is fundamentally flawed. From his theoretical analysis, Sader [40] suggests that the straight rectangular cantilever should be used in preference to V-shaped cantilever for the applications where the effects of lateral forces are to be minimized. Sader [40] concludes that V-shaped cantilever will enhance the lateral forces effect in comparison with the straight rectangular cantilever of identical normal stiffness, which is believed to degrade the AFM imaging performance. The conclusion and suggestion above are independent of whether the cantilever is in contact with a surface (sample) or not.

The bending stiffness of vertical bending is much smaller than the torsional stiffness, therefore the coupling of the vertical bending and torsion is usually ignored unless there is an eccentricity of the AFM tip [21]. The lateral

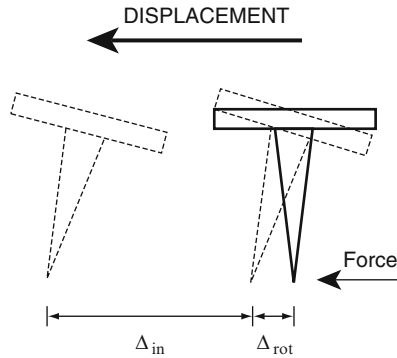


Fig. 8.6. Schematics of in-plane and rotation deformations Δ_{in} is the in-plane deformation, Δ_{rot} is the rotation deformation, the total in-plane deformation is the sum of these two deformation (*Source*: Sader and Green [18])

force exerted on the AFM tip not only generates the torsional moment to let the AFM rotate but also induces the in-plane deformation of the AFM. The in-plane deformation is defined as the deformation parallel to the plane of cantilever (cross-section), that is, left–right direction. The AFM structure is usually assumed to have the strong resistance to the in-plane deformation due to lateral loading and this in-plane deflection is thus assumed negligible. However, Sader et al. [18] demonstrate that this in-plane deformation can strongly contribute to the total deformation, which will affect the AFM imaging performance as a result, particularly for the rectangular cantilever with a high length/width aspect ratio. This is understandable because in the lateral bending and torsion cases (see Fig. 8.4), the bending stiffness of lateral bending is comparable to the torsional stiffness, which will induce the coupling very easily. The coupling effects between the vertical/lateral bending and torsion are studied systematically [12,19–21]. The coupling has significant impact on the frequency response function which determines the AFM imaging. The coupling also results in rich motion pattern of the system [42].

8.4 Electrostatic, Surface and Residual Stress Influence on the Atomic Force Microscope Initial Deflection

Many undesirable external forces can be exerted to AFM structure during the fabrication and operation. In this section, electrostatic force, surface stress and residual stress cases are discussed.

The influence of vdW force is incorporated in Sect. 8.6. These forces affect the AFM measured force–displacement data curve directly. Properly modeling these forces is vital for the correct extraction of material and surface properties from the experimental data.

The AFM cantilever–sample surface or the AFM tip–sample surface can form a capacitance-like structure [43–45]. If the electrical charges accumulate on cantilever and sample, there is a voltage difference between the two. Thus an attractive electrostatic force can be generated. The electrostatic force can be strong enough to cause the pull-in instability for the system [44]. The model by Abdel-Rahman et al. [44] considers the large deformation of the beam, which is the in-plane stretching of the beam and it strengthens the beam stiffness. Because the electrostatic force is inversely proportional to the square of the gap distance, sometimes the electrostatic force at the AFM tip is the dominant one and the electrostatic force acting on other parts of the AFM structure can be ignored. Belaidi et al. [46] discuss six models on the electrostatic force acting on the AFM tip only cases for different geometries of AFM tip. They are the sphere, single charge, knife edge, perfect cone, plane surface and hyperboloid models. If the electrostatic acting on the tip is not dominant, the combination of the models by Abdel-Rahman et al. [44] and Belaidi et al. [46] is necessary to obtain more accurate initial AFM deflection.

The sources of surface stress can be the surface reconstruction, the interactions between a surface and adsorbates or the interactions between a surface and its ambient environment, etc. During the microfabrication of AFM cantilever, the surface stress can arise and the effects due to it have been widely investigated [2,3,11,47,48]. The knowledge of the deflection of AFM cantilever is fundamental to the performance of the instrument and the interpretation of the measurements. Because surface stress acts on the AFM cantilever surface and there is an offset between the AFM structural neutral axis and the surface, bending moment is thus generated and AFM deflects consequently. The (initial) deflection due to surface stress must be considered for the interpretation of the AFM measured data; otherwise the topography interpretation of the sample surface will be distorted. For the surface stress influence on the cantilever deflection, Stoney's formula is widely used to evaluate the effect. Sader [2,3] argues that Stoney's formula is more suitable for the structure without constraint. And another implicit assumption of Stoney's formula is that structure deflects with a uniform curvature. Zhang et al. [49] give three models for the surface stress acting on a cantilever as shown in Fig. 8.7. Because surface stress acts on the cantilever surface and the beam governing equation describes the deflection of the neutral surface, the bending moment and axial force should be both embodied in the governing equation of the beam. Through the analysis [49], it is clear that Stoney's formula only account the bending moment effect. In Stoney's formula, the boundary conditions/constraints have no influence on curvature and the curvature keeps uniform because of the fact that only the bending moment is accounted. If the axial force effect of surface stress is accounted as shown in Fig. 8.7b, c, the curvature of the beam can no longer be uniform and boundary conditions can have very significant influence on the curvature. Zhang and Zhao [50] presented the detailed analysis of boundary conditions influence on the beam curvature. In many AFM applications, different materials may be deposited on the AFM surface to achieve

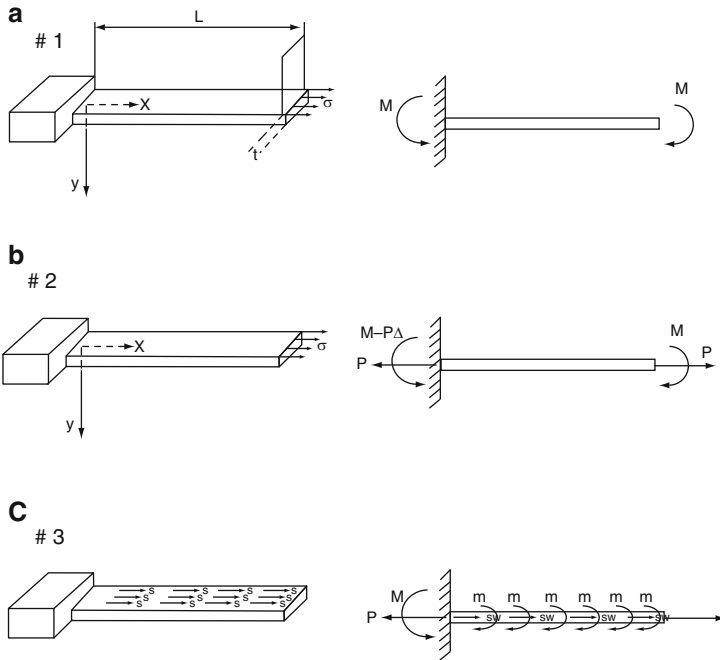


Fig. 8.7. Three models for the surface stress on a cantilever beam. (a) The surface effect is modeled as a bending moment; (b) surface stress is modeled as a bending moment and a concentrated axial force at the end; (c) surface stress is modeled as distributed axial force and bending moment (*Source:* Zhang et al. [49])

certain function, which in essence makes the AFM cantilever beam a composite structure. Because the AFM structure is of micron scale, the damaged interface effect which is negligible in large scale can be very significant. The damaged interface significantly changes the stress/strain inside the different layers, which as a result also changes the structure deflection and curvature [51]. When the AFM width is comparable with its length, the plate model is needed. Sader [2,3] shows that there is strong warping effect of the AFM cantilever plate and the plate deforms with nonuniform curvature with isotropic and homogeneous surface stress acting on it.

Functionally speaking, surface stress is a type of residual stress, which acts on the structure surfaces. During the fabrication of microstructures, various residual stresses are left. The residual stresses distributed nonuniformly in the thickness direction, which can usually be approximated by polynomial series [25]. The residual stresses generate both the axial force and bending moment on the microstructure which leads to tilting and curling [52]. When the AFM cantilever beam is with the tilting deformation, it will have impact on the AFM contact modeling [19] and the axial force due to the residual stresses can also alter the structure stiffness as a whole. One tricky thing about the residual

stresses is that the combination of different residual stresses can generate the same deflection, so the deflection alone cannot uniquely determine the residual stresses state. However, the deflection and curvature together can uniquely determine the residual stresses state [53,54].

8.5 Modeling Tip–Sample Contact

The tip on or around the AFM free end is the main sensing component ultimately responsible for the quality of AFM imaging. There are various choices for AFM tip material and the materials like diamond, silicon with various coatings can be used. For carbon nanotube (CNT), the properties of small tip radii, high mechanical resilience, wear resistance and large length–radius aspect ratio (which can result in significant bending flexibility) etc., make the CNT a good candidate for AFM tip material [4–10]. For small tip radii of CNT, the image with much better resolution can be achieved (see Fig. 8.8 for the comparison). The sharpness of the tip is the key to the image resolution. Snow et al. [8,9] investigate the dynamics of AFM with a CNT tip without contacting the surface. The dynamics model for the tapping mode of AFM with a CNT tip by Lee et al. [6] deal with the tip–surface contact issue. The

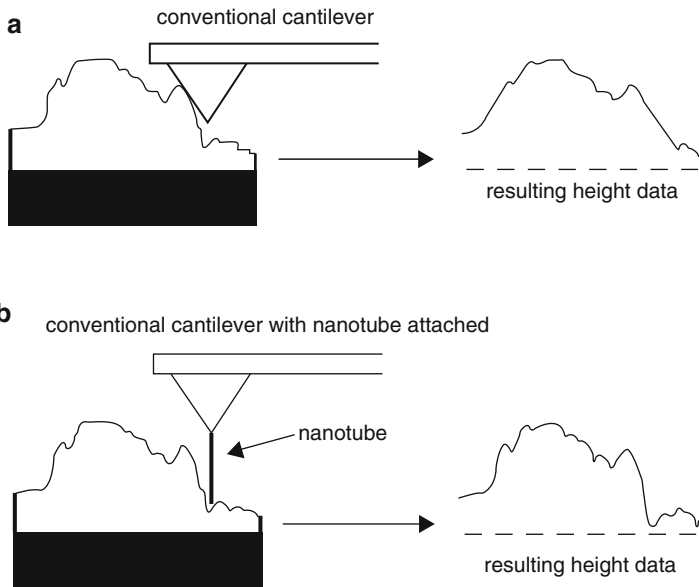


Fig. 8.8. Diagram of cross-sections of AFM tip–sample surface interaction. (a) Demonstration of how an imaging artifact can arise when imaging with a conventional AFM cantilever. (b) Demonstration of how a nanotube tip follows the surface, resulting in a more accurate representation of the surface topography (*Source*: Stevens et al. [10])

normal and lateral forces exerted on the AFM CNT tip by the surface may cause the large deformation, buckling, slip, slide and adhesion to the surface. The CNT tip is modeled as an Euler elastica. During its contact with the surface, the CNT tip experiences the large deformation and buckling [6]. As for most AFM tips which are chunky structures and is thus often modeled approximately as a sphere in contact with the sample, our contact mechanics discussion is on the AFM tip approximated as a sphere.

There are several different contact models which are applied in different scenarios. It is necessary to give a brief description on those contact models here. When the AFM tip is in contact with sample, the adhesion due to the energy difference of the two surfaces can be an important factor affecting the AFM force–displacement measured data curve. Compared with Hertz contact model, the JKR [55] model is a continuum mechanics theory considering the surface/adhesion energy. It relates the contact load F and contact radius a through the adhesion energy γ between the two surfaces as follows:

$$a^3 = \frac{R}{K} \left[F + 3\pi\gamma R + \sqrt{6\pi\gamma RF + (3\pi\gamma R)^2} \right] \quad (8.3)$$

R is effective radius, K is material property, γ is the energy per unit area required to separate the two surfaces. The equation above is for two elastic spheres contact case. The JKR model only considers the short-range adhesion inside the contact area. On the other hand, the DMT [56] model considers the long-range surface force acting outside the contact area. The MD (Maugis-Dugdale) [57] model uses the Dugdale (square well) potential to describe the attractive force. The MYD [58] model considers both the long-range and the short-range force acting both inside and outside the contact area [59]. The JKR theory is applicable to soft, compliant materials with high surface energy. In contrast, the DMT theory is suitable for harder, less compliant materials with low surface energy. The MD theory offers transition solution between the JKR and DMT theories [60]. A dimensionless number, Tabor number (or elastic parameter), is an important parameter differentiating the applicability regimes of the models above [59,61,62]. Physically, Tabor number is the ratio of (the order of) the “neck” height to the equilibrium separation of atoms. Two contacting bodies form a “neck” during contact. The neck height in essence determines the intermolecular forces outside contact area. The JKR model is the one with large Tabor number, which means the “neck” height is large compared with the atom equilibrium separation distance; therefore, the intermolecular forces outside the contact area are little to be ignored. On the other hand, the DMT model is the one with small Tabor number, the intermolecular forces outside the contact are significant. Zhang [61] recently shows the transitions between the above contact mechanics models by analyzing the pressure profiles inside the contact areas and the intermolecular forces influence outside the contact area. It also needs to point out that all the contact models above are also frictionless ones. In reality, AFM moves not only up and down, but also left and right. When a lateral force is exerted on the

AFM tip, the interfacial friction force can be significant [63]. Even in the vertical direction, Morrow et al. [60] also show that the friction can significantly be increased by the adhesive overload and their model considers the asperity effect of rough surface. The contact models discussed above only consider the normal force (vertical direction). The lateral/tangential force is obtained via the sliding contact model [64]. Song and Bhushan [19] give the contact forces in both normal and tangential directions for a tilted AFM cantilever as shown in Fig. 8.9:

In Fig. 8.9, f_n is the normal contact force which can be calculated from the contact models discussed above depending on the application scenarios. f_t is the tangential contact/friction force calculated from the sliding contact model [64]. The friction force given by Song and Bhushan [19] is a finite one. The frictionless contact and sticky contact which assume zero and infinite friction force, respectively, are the lower bound and upper bound of the real cases [63].

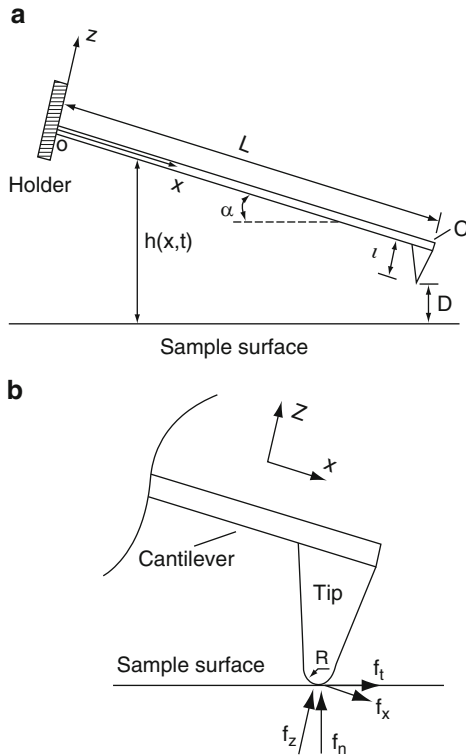


Fig. 8.9. Schematic diagram of a tip-cantilever system tilted to the sample surface with angle of α (a) and the tip-sample interaction (b) (Source: Song and Bhushan [19])

The contact mechanics models discussed above, for example, (8.3) of the JKR model, only establish the AFM tip (approximated as a sphere) contact with the sample surface. When applying such contact mechanics model to the AFM contact problem, the stiffness of the AFM cantilever should be considered. Song and Bhushan [12,20,21] incorporate the force due to the contact of the AFM tip–sample surface into the boundary conditions for the continuous system of the AFM cantilever. While, for the one DOF mass–spring system, the stiffness of the AFM cantilever must be incorporated in the governing equation. Yang [65] establishes such a model incorporating the spring constant of the AFM cantilever as a parameter into the AFM tip–sample contact problem (see Fig. 8.10). Equation (8.3) is the contact radius–load relation. Unlike the case of macroscopic contacts, it is usually not possible to measure the contact radius directly in experiments [66]. Yang [65] introduces another parameter of the transverse (vertical) displacement, which relates the contact load and the spring stiffness of the AFM cantilever. In the AFM measurement, it is the load–displacement curve that is obtained directly, not the contact radius–load curve. Yang et al. [63] establish such load–displacement model for adhesive sticky contact of a rigid sphere. Here the adjective “sticky” means the contact surface having infinite friction and therefore no slippage. The two equations derived by Yang et al. [63] for the load–displacement curve are as follows:

$$a^3 = \frac{R}{K_1} \left[P + 3\pi\gamma\chi R \pm \sqrt{6\pi\gamma\chi RP + (3\pi\gamma\chi R)^2} \right] \quad (8.4)$$

and

$$\delta = \frac{1}{\chi a K_1} \left[P + \pi\gamma\chi R \pm \sqrt{\frac{2\pi\gamma RP}{3} + (\pi\gamma R)^2} \right] \quad (8.5)$$

Here + sign stands for stable equilibrium and – sign for unstable one. R is the sphere radius. P is the applied load and δ is the vertical displacement (see Fig. 8.11). γ is the energy per unit area required to separate the two

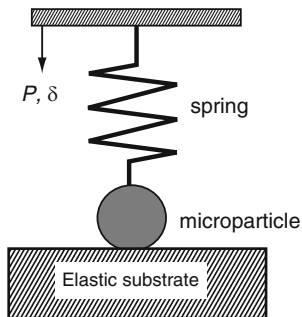


Fig. 8.10. AFM ‘spring–ball’ model – contact between a microparticle and an elastic substrate. Here the spring is to account the stiffness of the AFM cantilever (*Source*: Yang [65])

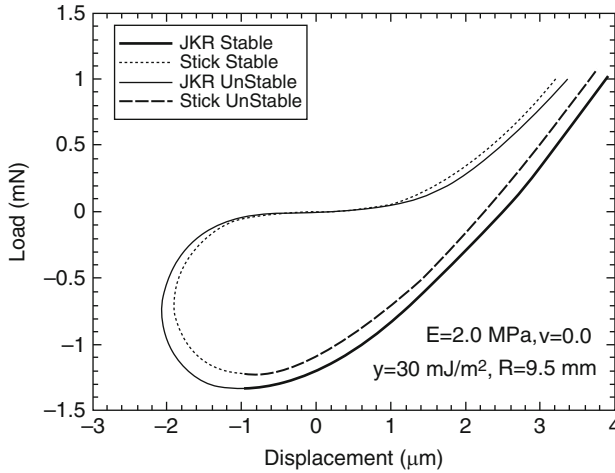


Fig. 8.11. Load–displacement comparison of frictionless contact (JKR) and sticky contact (*Source:* Yang et al. [63])

surfaces. χ is the parameter given as $\chi = 1 - 0.6931k^2 + 0.2254k^4 + \dots$ ($k = \frac{\ln(3-4\nu)}{\pi}$, ν is Poisson ratio). $K_1 = \frac{4E \ln(3-4\nu)}{3\chi(1-2\nu)(1+\nu)}$, E is Young's modulus. Because (8.5) contains the parameter of contact radius a , which is given in (8.4), these two equations are needed together to give the load–displacement relation of sticky contact. Both the stable and unstable solutions derived from (8.4) and (8.5) are compared with those of the JKR frictionless contact theory and shown in Fig. 8.11. It needs to point out that the adhesion hysteresis effect is not included in (8.4) and (8.5).

In all the above contact models dealing with adhesive contact, for example, (8.3) of the JKR model, there is also an implicit assumption that γ (surface energy per unit area) is a constant over the contact area and does not change with either the time of contact or the loading history. Usually this is not true, especially when the AFM is in contact with the viscoelastic materials. Hysteresis is very common in adhesive contact even without any inelastic deformation of the materials [67]. The energy needed to separate the two surfaces is usually greater than that gained by adjoining the two. The difference of these two energies is the reason causing the adhesion hysteresis [68]. The adhesion hysteresis is also the dissipation energy during a cyclic loading (see Fig. 8.12). It can be measured as the area enclosed by the cycle of displacement–load curve [66]. Rolling contact mechanics is another effective method to study the adhesion hysteresis. Rolling of a cylinder on a flat surface can be viewed as the propagation of two cracks. One is closing at the advancing edge and the other is opening at the receding edge. The receding work of adhesion (separating the cylinder from the flat plate) is usually greater than that of advancing one (adjoining the two surfaces). The receding work of adhesion

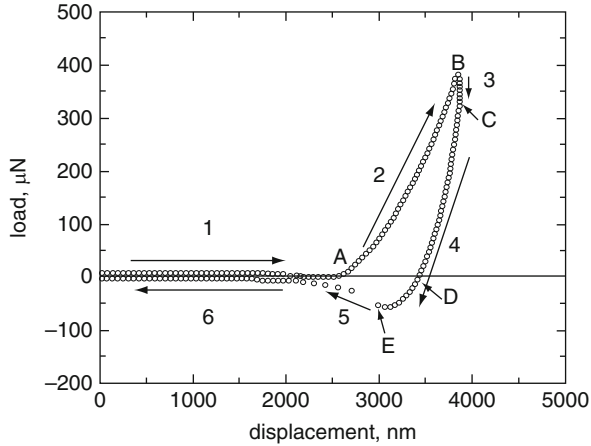


Fig. 8.12. Hysteresis: the load–displacement curve during a typical loading–unloading cycle at low loading rate (*Source:* Giri et al. [66])

and advancing work of adhesion are related by the contact widths of the stationary and rolling cylinder [68]. She and Chaudhury [68] propose a method of measuring advancing and receding works of adhesions, simultaneously and independently. The force needed to separate the two surfaces is called pull-off force. Yang et al. [67] show that the pull-off force is strongly dependent on the contact conditions (friction, ratio of the punch radius to the film thickness, etc.). For their model (the rigid punch on an incompressible elastic film), the contact area is assumed to be kept circular all the time. With the presence of adhesion, the pull-off force is tensile and measuring the pull-off force is also an effective way of finding the adhesion energy.

8.6 Non-Contact Atomic Force Microscope Dynamics: Damping and Influence of Tip–Surface Interaction

Non-contact AFM (NC-AFM) is called frequency modulation AFM (FM-AFM). In NC-AFM, the cantilever is always oscillated at/around its resonance frequency with a constant amplitude. During the measurement, the vibrating tip does not touch the sample surface. With the presence of the tip–sample interaction, the cantilever resonance frequency and oscillation amplitude are changing (due to the tip–surface interaction). The spatial difference of frequency shift can be used as a contrast to image the sample surface. Before we give the description of the AFM frequency shift under the vdW force influence, we would like to first address some characteristics of one DOF oscillator system.

The governing equation of one DOF oscillator holds as follows [69]:

$$\ddot{x}(t) + 2\xi\omega_n\dot{x}(t) + \omega_n^2x(t) = \omega_n^2Ae^{i\omega t} \quad (8.6)$$

x is the oscillator displacement, ω_n is the resonant frequency of undamped system, ξ is the viscous damping factor, $Ae^{i\omega t}$ is the external driving force (A is the amplitude and ω is the driving frequency). The steady-state response of the system response to the harmonic excitation of the above equation can be easily solved as follows:

$$x(t) = \frac{Ae^{i\omega t}}{1 - (\omega/\omega_n)^2 + i2\xi(\omega/\omega_n)} \quad (8.7)$$

The complex frequency response function (or magnification factor) can be defined as follows:

$$H(\omega) = \frac{x(t)}{Ae^{i\omega t}} = \frac{1}{1 - (\omega/\omega_n)^2 + i2\xi(\omega/\omega_n)} \quad (8.8)$$

This complex function $H(\omega)$ contains two pieces of information on the response: the magnitude $\left(|H(\omega)| = \left\{ [1 - (\omega/\omega_n)^2]^2 + 4\xi^2(\omega/\omega_n)^2 \right\}^{-1/2} \right)$ and the phase shift $\left(\phi = -\tan^{-1} \left[\frac{2\xi\omega/\omega_n}{1 - (\omega/\omega_n)^2} \right] \right)$, which are the two most important performance evaluation parameters in AFM. Notice that when $\omega = \omega_n$, $\phi = -\pi/2$; ϕ changes from 0 to $-\pi$ when ω changes from 0 to infinity. However, the phase shift angle ϕ has the dramatic change around $\phi = -\pi/2$ [12, 20, 21]. It is also noticed that the resonance frequency of the damping system decreases to $\omega_d = \omega_n\sqrt{1 - \xi^2}$ (ω_d is the resonant frequency of damped system).

In dynamic contact of tip-sample, there is an impact damage caused for both tip and sample. One way to decrease the impact energy is to immerse the AFM and sample in fluids to increase damping. Also in the application of measuring the properties of biological samples, the AFM is often required to be operated in their native solutions [19]. When AFM vibrates in fluids or the internal friction cannot be ignored as compared with that AFM is operated in vacuum, damping will increase very significantly and the additional mass effect will also brought in. Therefore there will be a resonant frequency shift [see (8.6)]. Sader et al. conduct a series of theoretical and experimental investigation on this issue [27, 29, 30]. In their model, the AFM internal friction/dissipative effect is ignored. The adding of the fluids is expected to increase the system damping (dissipative effect) and increase the system mass, which is often referred to as additional mass. The fluid influence is demonstrated both theoretically and experimentally for the viscous fluids [27], in which the density and viscosity effects on the resonant frequency shifts are well modeled. However, for inviscid fluids, the experimental data do not agree well with the data predicted by their model [27]. And we notice that in some models [27, 29, 30], the damping term does not show explicitly in the equation of

motion. It is implicitly included in the hydrodynamic force term of the driving force. The hydrodynamic force term can be approximated as follows [19]:

$$f_d(x, t) = -\rho_a \frac{\partial^2 w(x, t)}{\partial t^2} - c_a \frac{\partial w(x, t)}{\partial t} \quad (8.9)$$

f_d is the hydrodynamic force, ρ_a is the additional mass density and c_a is the additional hydrodynamic damping coefficient due to fluid. The first term is due to that the AFM induces the fluid motion during its vibration in liquid and the second term is usual viscous damping term. To find the accurate ρ_a is not an easy task. ρ_a has been given quite differently [19]. It is also worth mentioning that c_a may not be a constant. c_a consists of two parts $c_a = c_\infty + c_s$, c_∞ is the damping when the AFM is far away from the sample, which is a constant during the vibration; c_s is the damping when the AFM is very close to the sample to squeeze the fluid in and out of the region between AFM and sample and c_s is dependent on the gap distance between the AFM and sample, therefore it is not a constant [19].

For the tip–surface interaction, the Lennard-Jones (LJ) potential is considered [70]. The LJ potential model includes both the non-retarded dispersive attractive vdW force and the short-ranged repulsive force. If the tip–surface interaction only includes the attractive vdW force, the AFM always shows the softening effect in its frequency–amplitude response as shown in Fig. 8.13. With the presence of both attractive and repulsive interaction forces, the AFM can have either softening or hardening response.

For single DOF mass–spring oscillator model for AFM under external excitation, the following equation of motion is given [71]:

$$m \frac{d^2 z}{dt^2} + \frac{m\omega_o}{Q} \frac{dz}{dt} + k(z - u_o + h) + \frac{dV(z)}{dz} = kl \cos(\omega t) \quad (8.10)$$

Here m , k are effective mass of AFM cantilever and tip and spring constant of AFM cantilever, respectively. z is the total displacement of the oscillator. Q is the cantilever quality factor. ω_o is the AFM (fundamental) resonant frequency, u_o is equilibrium displacement, around which the external excitation is exerted. $V(z)$ is the tip–sample interaction energy, which is Lennard-Jones potential energy in the model of Sasaki and Tsukada [71]. l , ω are external excitation amplitude and frequency, respectively. u_o , ω are used as the parameters to study the AFM response frequency shift. Sasaki and Tsukada [71] use both numerical analysis and perturbation method to analyze (8.10). As u_o increases, the very similar softening effect to that of the continuous model by Rützel et al. [70] is shown. Their softening effect is due to the Lennard-Jones potential. Rützel et al. [70] show the stable and unstable responses of the steady-state motion in the frequency–amplitude and frequency–phase angle curves (see Fig. 8.13). The stability of the different portions of response curves can be determined either by investigating the nature of the singular point(s) or by superposing a perturbation on the steady-state motion [72].

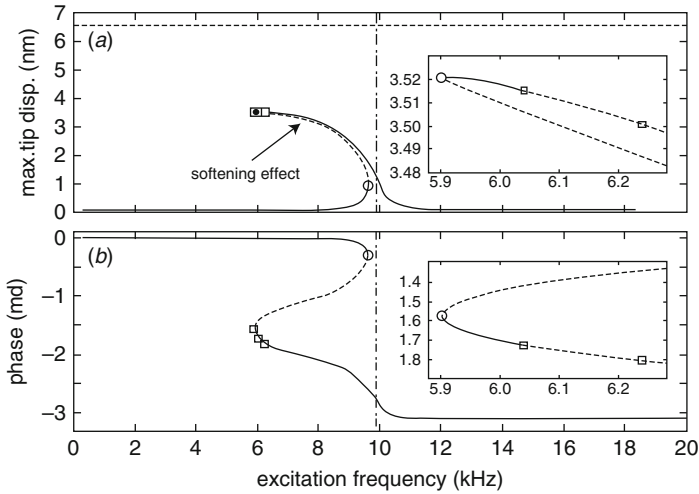


Fig. 8.13. The softening effect shown by the tip amplitude response versus the excitation frequency. The solid line and dashed line indicate the stability and instability separately. The softening is due to the attractive force of the Lennard-Jones potential, or say, van der Waals (vdW) force. With the repulsive force of the Lennard-Jones potential, the hardening effect can also be induced (*Source: Rützel et al. [70]*)

Hölscher et al. [73] differentiate the self-driven (self-excited, self-sustaining) cantilever from external-driven cantilever for AFM. It is not that difficult to conclude from the paper by Hölscher et al. [73] that for the self-driven system, the driving force is (the AFM cantilever) displacement-dependent; for the external driven system, it is displacement-independent. In a broader sense, the self-driven system can also include the system that the driving force is displacement-independent but the system with negative damping [72]. For both the TM and TRM, if the cantilever is driven to vibrate at frequencies above the fundamental resonant frequency, it would have less motion (high apparent stiffness), leading to higher sample deformation and better contrast [17] or the enhanced sensitivity to the force gradients in NC-AFM operating mode [74]. But for the self-driven system, if the system resonant frequency and external excitation are not commensurable, the system total response motion (the sum of free-oscillation and steady-state motion) may be aperiodic [72]. So if the cantilever is driven with an excitation frequency much higher or lower than the AFM fundamental resonant frequency, one should pay attention to that such scenario of aperiodic motion may happen.

For a continuous system, if the one-mode analysis is used, the system is equivalent to the one DOF system [75]. In their model, the sample–tip is still Lennard-Jones potential and during vibration of the AFM cantilever, the force due to the LJ potential keeps continuous. For the undamped one DOF system with no external excitation, the system is Hamiltonian (conservative). Ashhab

et al. [75] treat the damped system with external excitation system as perturbed Hamiltonian system by assuming the very small damping coefficient and excitation amplitude. Ashhab et al. [75] use Melnikov function method to do the time-periodic perturbation of a Hamiltonian system. A surface is generated from their solution, which differentiates the regions of chaos and no chaos. And it is worth mentioning that the shortcoming of modeling the continuous AFM cantilever system as one DOF system is that only one resonant frequency exists. For a continuous system, there are infinite resonant frequencies. So, using the spring constant method to model the AFM cantilever as one DOF system, one need to restrict the research about the AFM response around the fundamental resonant frequency. For those AFM acoustically excited at different very high frequencies, usually ultrasonic ones [76], the AFM must be modeled as a continuous system, otherwise it is impossible to model its response at difference high frequencies.

For a continuous system, the equation of motion of an AFM cantilever under the tip–surface interaction of LJ potential is given as follows [70]:

$$\begin{aligned} &\rho A \ddot{u}(x, t) + EI[u''''(x, t) + w^*''''(x, t)] \\ &= \left[\frac{-\frac{A_1 R}{180(Z-w^*(L)-u(L,t)-Y \sin \Omega t)^8}}{+\frac{A_2 R}{6(Z-w^*(L)-u(L,t)-Y \sin \Omega t)^2}} \right] \times \delta(x - L) + \rho A \Omega^2 Y \sin \Omega t \end{aligned} \quad (8.11)$$

u , ρA , EI , L are the transverse displacement from the equilibrium displacement of w^* , mass per unit length, bending stiffness and length of the beam, respectively. $\ddot{u} = \frac{d^2 u}{dt^2}$ and $\left\{ -\frac{A_1 R}{180[Z-w^*(L)-u(L,t)-Y \sin \Omega t]^8} + \frac{A_2 R}{6[Z-w^*(L)-u(L,t)-Y \sin \Omega t]^2} \right\} \times \delta(x - L)$ is the vdW force acting on the AFM tip (A_1, A_2 are Hamaker constants, R is the tip radius, Z is the tip–surface separation distance). The first term indicates the repulsive force and the second one is the attractive vdW force. δ here is the Dirac delta function, which here indicates vdW force acting as a concentrated load on the AFM tip. Y and Ω are the excitation amplitude and frequency, respectively. Clearly, in the above model, the tip–surface interaction on the other parts of the AFM beam is ignored. As for the static or dynamic analysis of excitation frequency around the first AFM natural frequency, the vdW force on the AFM tip is dominant. In those scenarios it is valid to ignore the tip–surface interactions because this interaction dies out rapidly with the increase of the gap distance. However, when the AFM is driven with higher resonant frequency, the AFM vibration deflection becomes “wavy” (see Fig. 8.14), some parts of the beam becomes closer to the sample surface than the beam tip end, which will have larger AFM cantilever–sample surface interaction force. Therefore, the above model of ignoring the interaction force on other parts of the AFM beam may be problematic under those scenarios. The repulsive force in (8.11) will surpass the attractive one when the AFM is very close to the sample surface. In that sense, this repulsive force is due to the contact. However, as pointed by Rützel et al. [70] themselves that (8.11) should not be used to

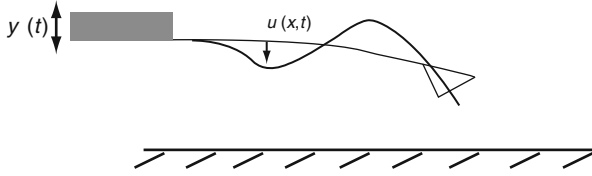


Fig. 8.14. The model of Rützel et al. [70] only considers the tip–surface interaction effect on the cantilever tip. As the excitation frequency approaches the natural frequencies other than the first natural frequency of the AFM cantilever structure, the beam response may look like that shown in the figure, in which the vdW effect on other part of the beam can be significant and cannot be ignored

model the tapping mode of intermittent contact. The contact force is a complicated one which depends on the contacting bodies' geometry and surface properties, different contact mechanics models are derived for different scenarios and (8.11) cannot accurately describe the contact force. Equation (8.11) represents a generic tip–sample interaction. With the inclusion of the repulsive force in (8.11), not only softening effect but also the hardening effect can be found in the numerical simulation [70]. Equation (8.11) is only suitable for the non-contact dynamics study of AFM and the dynamic study of the AFM intermittent contact is given in the next section. We also notice that the equation of motion above, that is, (8.11), does not account for the lumped/concentrated mass of AFM tip, which is shown by Cleveland et al. [26] to have a significant contribution to the system resonant frequencies. The concentrated tip mass reduces the AFM resonant frequencies, which has been utilized by Li et al. [77] as an effective way to achieve better AFM imaging in the intermittent contact. It is also worth pointing out that Argento and French [78] derive a different (concentrated) tip–surface interaction force F_{int} for a conical AFM tip, which is given as follows:

$$F_{\text{int}} = \frac{HR^2(1 - \sin \gamma)(R \sin \gamma - z_0 \sin \gamma - R - z_0)}{6z_0^2(R + z_0 - R \sin \gamma)^2} - \frac{H \tan \gamma [z_0 \sin \gamma + R \sin \gamma + R \cos(2\gamma)]}{6 \cos \gamma (z_0 + R - R \sin \gamma)^2} \quad (8.12)$$

Here H is the Hamaker constant for tip–surface interaction. z_0 , R are the tip–surface separation and the radius of the sphere at the AFM tip end, respectively. γ is the angle shown in Fig. 8.10. For the expressions of vdW force acting on AFM tip with pyramidal or other more complex shapes, the reader should refer to the works by Suresh and Walz [79], Touhari et al. [80] and Zanette et al. [81].

8.7 Dynamics of Intermittent Contact

The tapping mode is one of the most widely used high-resolution proximal probe techniques. The intermittent contact between the tip and the sample surface effectively eliminates the influence of lateral adhesion and friction, and thus, allows the topography imaging of rather soft and fragile samples [77]. The tip periodically tapping on the sample surface induces the periodic tip-sample force, which generates the motion with the high-frequency components. These high-frequency components are the harmonics of the excitation frequency. Those higher harmonics have been used to image the sample surface and show very significant improvement on the imaging [82]. The main advantage of using higher harmonics is the increased contrast and sensitivity to variations in properties of the sample. The strategy of intermittent contact modeling is to give two equations of describing the system with and without contact, respectively [83, 84]. If the AFM (tip) displacement is set as an unknown variable and the contact models discussed above are used, the governing equation for the AFM in contact is nonlinear. For simplicity, we can assume that when the AFM tip is in contact with the sample surface, it compresses a spring [20, 84] and the spring can be linearized. Even the two governing equations of AFM are both linear; the intermittent contact dynamics is still a highly nonlinear problem. The nonlinearity arises from the switching between the two governing equations, or say, the impact time is unknown. The rich nonlinear dynamics of the impact of one DOF mass-spring (in which the spring is a linear one) has been shown by Shaw and Holmes [83].

Due to the tip-sample impacts, the AFM cantilever motion can be either chaotic or periodic motion under external excitation and it has been experimentally observed [85]. The irregular, chaotic motion of the AFM cantilever is highly undesirable because of its feedback control problem (e.g., under frequency or force modulation) and inaccurate measurement [75]. There are abundant literatures in Thompson and Stewart's book [86] on the system response to the impact. To our best knowledge, Burnham et al. [85] are the first to use the coefficient of restitution method to study the AFM tip-sample impact response. The coefficient of restitution is less than unity, which thus accounts the energy loss during the impact. The AFM tip velocity after impact is the product of the coefficient of restitution and the velocity (just) before the impact. Under the different excitation amplitude, frequency and gap distance between the tip and sample, the tip may have periodic *subharmonic* resonance or chaotic motion. Before we go any further on this tip-sample response issue, we would like to spend some time explaining the concept of subharmonic resonance. In parametric excitation system (the excitation appears as a time-dependent variable in the governing equation), when the external excitation frequency Ω approaches the natural frequency ω_o (i.e., $\Omega \approx \omega_o$), the system resonance is called *primary resonance*. For a nonlinear system, when Ω shifts away from ω_o , the system may have the resonance called *secondary resonance*. If $\Omega \approx n\omega_o$ (n is an integer other than 1), the secondary resonance happens,

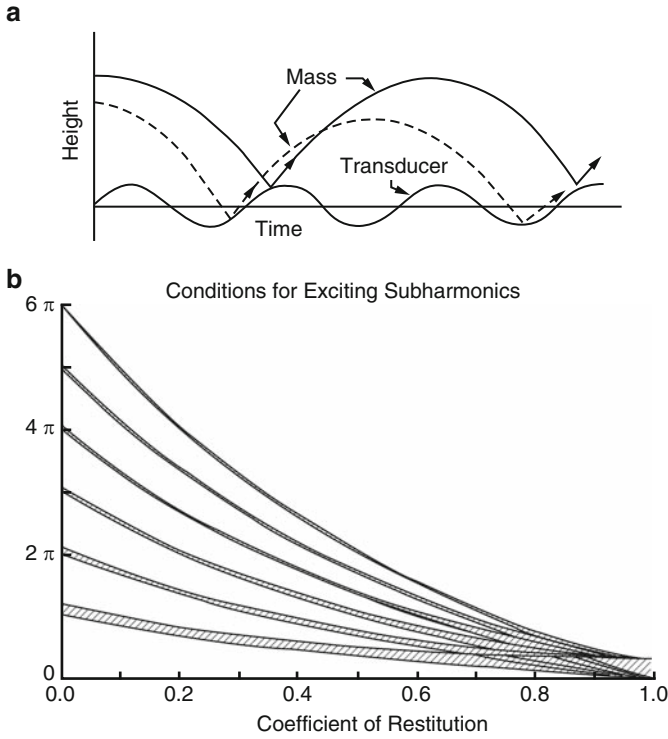


Fig. 8.15. (a) The heights of the mass and the transducer as a function of time. The mass here is the AFM cantilever effective tip mass and the transducer is mounted on the bottom of a sample. The sample is excited to hit/impact the AFM tip rather than the tip is excited. The broken (solid) line represents a nonperiodic (periodic) trajectory of the mass. From this schematic diagram, the subharmonic resonance (solid) is the resonance of period 2. (b) Regimes of stable periodic trajectories for the mass as a function of ε (coefficient of restitution) and the ratio of a/a_o . The shaded areas are where the conditions of (8.13) are satisfied (*Source*: Burham et al. [85])

it is called subharmonic resonance. If $\Omega \approx \frac{1}{n}\omega_o$, the secondary resonance happens, it is called *superharmonic* [72]. A subharmonic resonance of period n can be stable if the following equation is satisfied [85]:

$$\frac{1 - \varepsilon}{1 + \varepsilon} \pi n < \frac{a_o}{a} < \frac{1 - \varepsilon}{1 + \varepsilon} \left[1 + \left(\frac{1 + \varepsilon^2}{1 - \varepsilon^2} \frac{2}{\pi n} \right)^2 \right]^{1/2} \quad (8.13)$$

ε is the coefficient of restitution, a_o here is the maximum acceleration of the transducer and a is the free acceleration of AFM cantilever effective tip mass. Figure 8.15a shows a schematic subharmonic of period 2 (solid line). Figure 8.15b gives the stability zone (shaded area) of subharmonic of different periods. Berg and Briggs [87] and van de Water and Molenaar [88] further

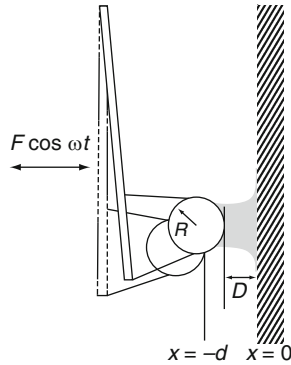


Fig. 8.16. Oscillating AFM. The cantilever with mass m is driven with a harmonic force $F \cos(\omega t)$. Its deflection is x , its velocity \dot{x} . It can collide with the substrate at $x = 0$. In equilibrium (drawn with dashed curves), the separation between the tip and the surface is d . The tip interacts with the substrate through a liquid bridge. In our model, the tip bounces elastically off the rigid surface when $D = 10^{-10}$ m. Impact energy loss is due to viscous damping of the liquid bridge [88]

extend this coefficient of restitution to study the dynamics of impact oscillators in the presence of a liquid-mediated interaction between the impacting tip and the surface (see Fig. 8.15). The model by Burnham et al. [85] and van de Water and Molenaar [88] treats the AFM system as one DOF system. In the model of van de Water and Molenaar [88], the AFM system transition to chaotic motion after two period-adding processes is shown (see Figs. 8.16 and 8.17).

Shaw and Holmes [83] did an excellent and classical analysis of one DOF piecewise linear oscillator. One end of mass attached to a linear spring and dashpot and the system is under a periodic excitation. When the mass moves a certain displacement, it will hit the other linear spring. According to their definition [83], if one of spring stiffness approaches infinite, the system becomes an impact oscillator. Shaw and Holmes [83] use the coefficient of restitution method for the analysis in the impact limit case. In their impact analysis, time is “frozen” when impact happens, that is, impact takes no time. So, when the AFM tip is in contact with soft sample material, the contact–separation process of AFM tip–sample can be a significant portion of the period of the external driving force, the coefficient of restitution method cannot be applied. van de Water and Molenaar [88] point out that the elastic collision energy loss cannot be solely responsible for the AFM tip–sample impact energy loss with liquid bridge. The viscous damping and acoustic radiation due to higher mode excitation of cantilever sometimes are the major reasons for the energy loss. The coefficient of restitution method fundamentally ignores the details of the AFM tip–sample interactions. For such scenario, the very general model provided by Shaw and Holmes [83] should be considered for the AFM tip–sample impact dynamics. Shaw [89] expanded the similar analysis above to a

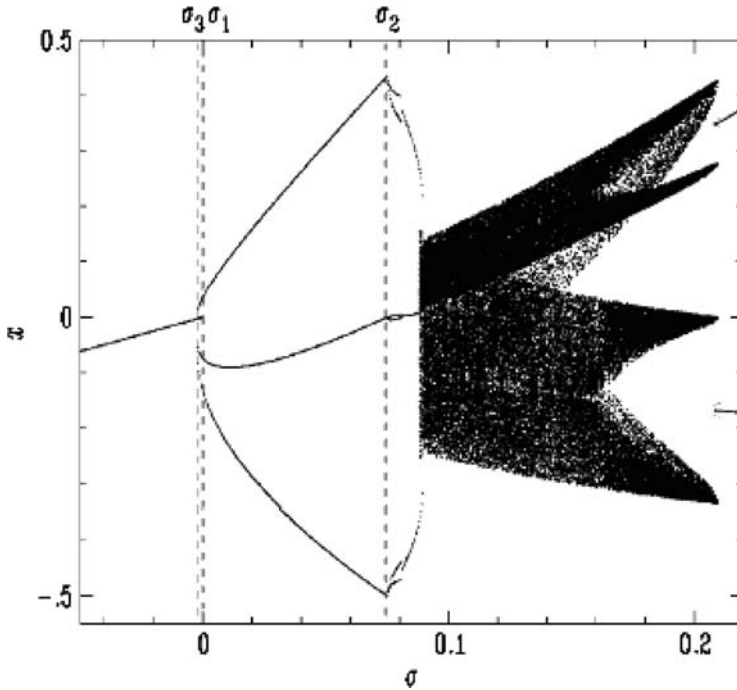


Fig. 8.17. One transition example of AFM motion from periodic to chaotic one is shown. σ is the scaled driving amplitude. When $\sigma < 0$, there is no impact happening. The AFM motion is period-1 motion. At $\sigma = \sigma_1 = 0$, the transition to period-3 orbit happens. The system undergoes a secondary bifurcation at $\sigma = \sigma_2$ to a period-3 orbit, which impacts twice in one period. For $\sigma \in [0.089, 0.209]$, the chaotic motion is seen [88]

cantilever driven by external force and there is a stop on its one side. The stop is modeled as the amplitude constraint. The cantilever displacement is discretized by modal expansion. There are two modal expansions for the cantilever. The constraint is incorporated in the modal expansion, when the cantilever tip hits the stop, the modal expansion of the cantilever displacement switches to the other form. But Shaw's modal expansion has only one term. Essentially, his model is still one DOF system. Even for this "simple" piecewise linear oscillator system (the nonlinearity is due to this piecewise linearity of springs, or say, the discontinuity of the stiffness), Shaw and Holmes [83] demonstrate that the system can have very complex responses such as subharmonic resonance, period doubling, chaotic motion. Nordmark [90] proposes the grazing impact to simplify the analysis of the impact dynamics. When the impact happens with zero velocity, it is called grazing impact. With the grazing impact scheme, the mathematics on the state-space (Poincaré mapping) can be dramatically reduced and the effect of derivative singularity of Poincaré mapping on periodic orbit can be analyzed. For the Poincaré mapping, the

difficult differential governing equations do not need to be solved and are transformed into algebraic difference equations. By studying the mapping, one can obtain the information on the system dynamic state. For the grazing impact, very complicated system response can still happen [91]. And using the grazing impact model for general case of AFM tip-sample impact can also be problematic [88] for obvious reasons.

In reality, AFM can neither bounce back instantly as suggest by the coefficient of restitution method nor impact the sample surface with zero velocity as indicated by the grazing impact. In reality, the AFM tip impacts with the sample surface with finite speed and the contact time can be a significant fraction of the forcing period. Zhang and Zhao [84] present a model to account the above two effects (nonzero impact speed and significant contact time). The model is that an AFM impacts with a spring and a damper. The spring to account the sample surface elasticity and the damper is to account the energy loss during the impact. This simple model shows that during the intermittent contact, multiple modes of the AFM cantilever are excited, which suggests that the one DOF model is not suitable for the intermittent contact dynamics. The problem/difficult of the above model [84] is to find the proper spring stiffness and damping coefficient for the contacting sample.

8.8 Summary

We present a comprehensive review of the AFM modeling analyses on its statics, tip-sample contact mechanics, non-contact and contact dynamics, etc. The focus is on the physical assumptions explicitly or implicitly included in those models. Modeling the AFM correctly by considering all the aspects of AFM structure, tip-sample interactions, residual stress, etc. is extremely difficult and certain simplifications are thus introduced to keep the model workable. Different researchers use different physical assumptions to avoid the modeling difficulty for different specific AFM operating scenarios. Therefore, the applicability regimes of those models should be clear for those who use them. By analyzing those physical assumptions concerning the modeling AFM, we try to offer our comments on the issues of the applicability regimes and reliability problems of those models. We also present the analysis difficulties on modeling AFM. The dynamic impact response of AFM system demonstrates very complicated and various motions. The subharmonic and chaotic motions are highly undesirable and harmful to the feedback control of AFM. But they are unavoidable in many AFM applications. Few researches have been done in this area. However, there are abundant research literatures concerning this impact response for macroscopic system. How to utilize and extend those research results of macroscopic system to the AFM microstructure by incorporating the effects such as capillary, adhesion and vdW force will be a challenge yet helpful to the AFM development.

Acknowledgement

We are thankful to Dr. Weixue Tian of Caterpillar Champaign Simulation Center and Dr. Zhenyu Yang of Beijing University of Aeronautics and Astronautics for proof-reading the manuscript. This research is supported by the National Natural Science Foundation of China (NSFC) (Grant No. 10721202) and the LNM initial funding for young investigators.

References

1. W.A. Hofer, A.S. Foster, A.L. Shluger, Theories of scanning probe microscopes at the atomic scale. *Rev. Mod. Phys.* **75**, 1287–1331 (2003).
2. J.E. Sader, Surface stress induced deflections of cantilever plate with application to the atomic force microscope: rectangular plates. *J. Appl. Phys.* **89**, 2911–2921 (2001).
3. J.E. Sader, Surface stress induced deflections of cantilever plate with application to the atomic force microscope: V-shaped plates. *J. Appl. Phys.* **91**, 9354–9361 (2002).
4. H. Dai, J.H. Hafner, A.G. Rinzler, D.T. Colbert, R.E. Smalley, Nanotubes as nanoprobe in scanning probe microscopy. *Nature* **384**, 147–150 (1996).
5. A. Hall, W.G. Matthews, R. Superfine, M.R. Falvo, S. Washburn, Simple and efficient method for carbon nanotube attachment to scanning probes and other substrates. *Appl. Phys. Lett.* **82**, 2506–2508 (2003).
6. S.I. Lee, S.W. Howell, A. Raman, R. Reifengerger, C.V. Nguyen, M. Meyyappan, Nonlinear tapping dynamics of multi-walled carbon nanotube tipped atomic force microcantilevers. *Nanotechnology* **15**, 416–421 (2004).
7. K. Moloni, M.R. Buss, R.P. Andres, Tapping mode scanning force microscopy in water using a carbon nanotube probe. *Ultramicroscopy* **80**, 237–246 (1999).
8. E.S. Snow, P.M. Campbell, J.P. Novak, Single-wall carbon nanotube atomic microscope probes. *Appl. Phys. Lett.* **80**, 2002–2004. 2002a
9. E.S. Snow, P.M. Campbell, J.P. Novak, Atomic force microscopy using single-wall C nanotube probes. *J. Vac. Sci. Tech. B* **20**, 822 (2002b).
10. R.M.D. Stevens, N.A. Frederick, B.L. Smith, D.E. Morse, G.D. Stucky, P.K. Hansma, Carbon nanotubes as probes for atomic force microscopy. *Nanotechnology* **11**, 1–5 (2000).
11. T. Miyatani, M. Fujihira, Calibration of surface stress measurement with atomic force microscopy. *J. Appl. Phys.* **81**, 7099–7115 (1997a).
12. Y. Song, B. Bhushan, Coupling of cantilever lateral bending and torsion in resonance and lateral excitation modes of atomic force microscopy. *J. Appl. Phys.* **99**, 094911 (2006a).
13. T. Miyatani, M. Horii, A. Rosa, M. Fujihira, O. Marti, Mapping of electrical double-layer force between tip and sample surfaces in water with pulsed-force-mode atomic force microscopy. *Appl. Phys Lett.* **71**, 2632–2634 (1997b).
14. A. Rosa-Zeiser, E. Weilandt, S. Hild, O. Marti, The simultaneous measurement of elastic, electrostatic and adhesive properties by scanning force microscopy: pulsed-force mode operation. *Meas. Sci. Technol.* **8**, 1333–1338 (1997).
15. P.J. de Pablo, J. Colchero, J. Gómez-Herrero, A.M. Baró, Jumping mode scanning force microscopy. *Appl. Phys. Lett.* **73**, 3300–3302 (1998).

16. Q. Zhong, D. Innis, K. Kjoller, V. Elings, Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy. *Surf. Sci. Lett.* **290**, L688–L692 (1993).
17. T. Kasai, B. Bhushan, L. Huang, C. Su, Topography and phase imaging using the torsional resonance mode., *Nanotechnology* **15**, 731–742 (2004).
18. J.E. Sader, C.P. Green, In plane deformation of cantilever plates with applications to lateral force microscopy. *Rev. Sci. Instrum.* **75**, 878–883 (2004).
19. Y. Song, B. Bhushan, Finite-element vibration analysis of tapping-mode atomic force microscopy. *Ultramicroscopy* **107**, 1095–1104 (2007).
20. Y. Song, B. Bhushan, Quantitative extraction of in-plane surface properties using torsional resonance mode of atomic microscopy. *J. Appl. Phys.* **97**, 083533 (2005).
21. Y. Song, B. Bhushan, Simulation of dynamic modes of atomic force microscopy using a 3D finite element model. *Ultramicroscopy* **106**, 847–873 (2006b).
22. T.R. Albrecht, S. Akamine, T.E. Carver, C.F. Quate, Microfabrication of cantilever styli for the atomic force microscope. *J. Vac. Sci. Technol. A* **8**, 1368–1374 (1990).
23. R.J. Roark, W.C. Young, *Formulas for stress and strain* (McGraw-Hill, New York, 1975).
24. J.E. Sader, I. Larson, P. Mulvaney, L.R. White, Method for calibration of atomic force microscope cantilevers. *Rev. Sci. Instrum.* **66**, 3789–3798 (1995a).
25. W. Fang, J.A. Wickert, Post buckling of micromachined beams. *J. Micromech. Microeng.* **4**, 116–122 (1994).
26. J.P. Cleveland, S. Manne, D. Bocek, P.K. Hansma, A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. *Rev. Sci. Instrum.* **64**, 403–405 (1993).
27. J. Chon, P. Mulvaney, J.E. Sader, Experimental validation of theoretical models for the frequency response of atomic force microscope cantilever beams immersed in fluids. *J. Appl. Phys.* **87**, 3978–3988 (2000).
28. C.T. Gibson, G.S. Watson, S. Myhra, Determination of the spring constants of probes for force microscopy/spectroscopy. *Nanotechnology* **7**, 259–262 (1996).
29. J.E. Sader, Frequency response of cantilever beams immersed in viscous fluids with application to the atomic force microscope. *J. Appl. Phys.* **84**, 64–76 (1998).
30. J.E. Sader, J.W.M. Chon, P. Mulvaney, Calibration of rectangular atomic force microscope cantilever. *Rev. Sci. Instrum.* **70**, 3907–3969 (1999).
31. J.E. Sader, Parallel beam approximation for V-shaped atomic force microscope cantilevers. *Rev. Sci. Instrum.* **66**, 4583–4587 (1995b).
32. H.J. Butt, P. Siedle, K. Seifert, K. Fendler, T. Seeger, E. Bamberg, A.L. Weisenhorn, K. Goldie, A. Engel, Scan speed limit in atomic force microscopy. *J. Microscopy* **169**, 75–84 (1993).
33. D.F. Ogletree, R.W. Carpick, M. Salmerson, Calibration of frictional forces in atomic force microscopy. *Rev. Sci. Instrum* **67**, 3298–3306 (1996).
34. O. Pietrement, J. Beaudoin, M. Troyon, A new calibration method of the lateral constant stiffness and lateral force using modulated lateral force microscopy. *Tribol. Lett.* **7**, 213–220 (1999).
35. R.G. Cain, S. Biggs, N.W. Page, Force calibration in lateral force microscopy. *J. Colloid Interface Sci.* **227**, 55–65 (2000).
36. A. Feiler, P. Attard, I. Lardon, Calibration of torsional spring constant and lateral photodiode response of frictional force microscopes. *Rev. Sci. Instrum.* **71**, 2746–2750 (2000).

37. R.G. Cain, M.G. Reitsma, S. Biggs, N.W. Page, Quantitative comparison of three calibration techniques for the lateral force microscope. *Rev. Sci. Instrum.* **72**, 3304–3312 (2001).
38. S. Ecke, R. Raiteri, E. Bonaccorso, C. Reiner, H.J. Deiseroth, H.J. Butt, Measuring normal and friction force acting on individual fine particle. *Rev. Sci. Instrum* **72**, 4164–4170 (2001).
39. C.P. Green, J.E. Sader, Torsional frequency response of cantilever beams immersed in viscous fluids with application to the atomic force microscope. *J. Appl. Phys.* **92**, 6262–6274 (2002).
40. J.E. Sader, Susceptibility of atomic force microscope cantilevers to lateral forces. *Rev. Sci. Instrum.* **74**, 2439–2443 (2003).
41. M. Varenberg, I. Etsion, G. Halperin, An improved wedge calibration method for lateral force in atomic force microscopy. *Rev. Sci. Instrum.* **74**, 3362–3367 (2003).
42. P.V. Bayly, K.D. Murphy, Coupling between dissimilar modes in an asymmetrically forced string. *J. Acoust. Soc. Am* **103**, 3362–3369 (1998).
43. Y. Ahn, H. Guckel, J.D. Zook, Capacitive microbeam resonator design. *J. Micromech. Microeng.* **11**, 70–80 (2001).
44. E.M. Abdel-Rahman, M.I. Younis, A.H. Nayfeh, Characterization of the mechanical behavior of an electrically actuated microbeam. *J. Micromech. Microeng.* **12**, 759–766 (2002).
45. E.K. Chan, K. Garikipati, R.W. Dutton, Characterization of contact electromechanics through capacitance-voltage measurements and simulations. *J. Micromech. Microeng.* **8**, 208–217 (1999).
46. S. Belaidi, P. Girard, G. Leveque, Electrostatic forces acting on the tip in atomic force microscopy Modelization and comparison with analytic expressions. *J. Appl. Phys* **81**, 1023–1030 (1997).
47. R. Raiteri, H.J. Butt, Measuring electrochemically induced surface stress with an atomic force microscope. *J. Phys. Chem.* **99**, 15728–32 (1996).
48. R. Raiteri, H.J. Butt, M. Grattarola, Changes in surface stress measured with an atomic force microscope. *Scanning Microscopy* **12**, 243–251 (1998).
49. Y. Zhang, Q. Ren, Y. Zhao, Modelling analysis of surface stress on a rectangular cantilever beam. *J.Phys. D: Appl. Phys.* **37**, 2140–2145 (2004).
50. Y. Zhang, Y. Zhao, Applicability range of Stoney’s formula and modified formulas for a film/substrate bilayer. *J. Appl. Phys.* **99**, 053513 (2006a).
51. Y. Zhang, Extended Stoney’s formula for a film-substrate bilayer with the effect of interfacial slip. *J. Appl. Mech.* **75**(011008–1) 9 (2008a).
52. W. Fang, J.A. Wickert, Determining mean and gradient residual stresses in thin films using micromachined cantilevers. *J. Micromech. Microeng.* **6**, 301–309 (1996).
53. Y. Zhang, Y. Zhao, An effective method of determining the residual stress gradients in a micro-cantilever. *Microsyst. Techno.* **12**, 357–364 (2006b).
54. Y. Zhang, Deflections and curvatures of film-substrate structure with the presence of gradient stress in MEMS application. *J. Micromech. Microeng.* **17**, 753–762 (2007).
55. K.L. Johnson, K. Kendall, A.D. Roberts, Surface energy and the contact of elastic solids. *Proc. Roy. Soc. London A* **324**, 301–313 (1971).
56. B.V. Derjaguin, V.M. Muller, Y.P. Toporov, Effect of contact deformations on the adhesion of particles. *J. Colloid Interface Sci.* **53**, 314–326 (1975).
57. D. Maugis, Adhesion of spheres: the JKR-DMT transition using a Dugdale model. *J. Colloid Interface Sci.* **150**, 243–269 (1992).

58. V.M. Muller, V.S. Yushenko, B.V. Derjaguin, On the influence of molecular forces on the deformation of an elastic sphere and its sticking to a rigid plane. *J. Colloid Interface Sci* **77**, 91–101 (1980).
59. X. Shi, Y.P. Zhao, Comparison of various adhesion contact theories and the influence of dimensionless load parameter. *J. Adhesion Sci. Technol.* **18**, 55–68 (2004).
60. C. Morrow, M. Lovell, X. Ning, A JKR-DMT transition solution for adhesive rough surface contact. *J. Phys. D: Appl. Phys.* **36**, 534–540 (2003).
61. Y. Zhang, Transitions between different contact models. *J. Adhesion Sci. Technol.* **22**, 699–715 (2008b).
62. Y.P. Zhao, L.S. Wang, T.X. Yu, Mechanics of adhesion in MEMS- a review. *J. Adhesion Sci. Technol.* **17**, 519–546 (2003).
63. F. Yang, X. Zhang, J.C.M. Li, Adhesive contact between a rigid sticky sphere and an elastic half space. *Langmuir* **17**, 716–719 (2001a).
64. K.L. Johnson, *Contact Mechanics* (Cambridge University Press, 1985).
65. F. Yang, Load-displacement relation in adhesion measurement. *J. Phys. D: Appl. Phys.* **36**, 2417–2420 (2003).
66. M. Giri, D.B. Bousfield, W.N. Unertl, Dynamic contact on viscoelastic film: work of adhesion. *Langmuir* **17**, 2973–1981 (2001).
67. F. Yang, J.C.M. Li, Adhesion of a rigid punch to an incompressible elastic film. *Langmuir* **17**, 6524–6529 (2001b).
68. H. She, M.K. Chaudhury, Estimation of adhesion hysteresis using rolling contact mechanics. *Langmuir* **16**, 622–625 (2000).
69. L. Meirovitch, *Analytical Methods in Vibrations* (Macmillan, New York, 1967).
70. S. Rützel, S.I. Lee, A. Raman, Nonlinear dynamics of atomic-force-microscope probes driven in Lennard-Jones potentials. *Proc. R. Soc. Lond. A* 01PA313/1–24 **459**, 1925–1948 (2003).
71. N. Sasaki, M. Tsukada, The relation between resonance curves and tip-surface interaction in noncontact atomic force microscopy. *Jpn. J. Appl. Phys.* **38**, L533–535 (1998).
72. A.H. Nayfeh, D.T. Mook, *Nonlinear Oscillation* (Wiley, 1979).
73. H. Hölscher, W. Gutschmann B Allers, U.D. Schwars, H. Fuchs, Wiesendanger, Measurement of conservative and dissipative tip-sample interaction forces with a dynamic force microscope using the frequency modulation technique. *Phys. Rev. B* **75**, 1287–1331 (2003).
74. M. Hoummady, E. Farnault, Enhanced sensitivity to force gradients by using higher flexural modes of the atomic force microscope cantilever. *Appl. Phys. A* **66**, S361–364 (1998).
75. M. Ashhab, M.V. Salapaka, M. Dahleh, I. Mezi, Dynamical analysis and control of microcantilevers. *Automatica* **35**, 1663–1670 (1999).
76. M.Y. Cuberes, G.A.D. Briggs, O. Kolosov, Nonlinear detection of ultrasonic vibration of AFM cantilevers in and out of contact with the sample. *Nanotechnology* **12**, 53–59 (2001).
77. H. Li, Y. Chen, L. Dai, Concentrated-mass cantilever enhances multiple harmonics in tapping-mode atomic force microscopy. *Appl. Phys. Lett.* **92**, 151903 (2008).
78. C. Argento, R.H. French, Parametric tip model and force-distance relation for Hamaker constant determination from atomic-force microscopy. *J. Appl. Phys.* **80**, 6081–6090 (1996).

79. L. Suresh, Y.J. Walz, Effect of surface roughness on the interaction energy between a colloidal sphere and a flat plate. *J. Colloid Interface Sci.* **193**, 199–213 (1996).
80. F. Touhari, G. Cohen-Solal, X. Bouju, C.h. Girard, M. Devel, Scanning force microscopy simulations of well-characterized nanostructures on dielectric and semiconducting substrates. *Appl. Surf. Sci.* **125**, 351–359 (1998).
81. S.I. Zanette, A.O. Caride, V.B. Nunes, G.L. Klimchitskaya, F.L. Freire Jr., R. Prioli, Theoretical and experimental investigation of the force-distance relation for an atomic force microscope with a pyramidal tip. *Surf. Sci.* **453**, 75–82 (2000).
82. J. Preiner, J. Tang, V. Pastushenko, P. Hinterdorfer, Higher harmonic atomic force microscopy: imaging of biological membrane in liquid. *Phys. Rev. Lett.* **99**, 046102 (2007).
83. S.W. Shaw, P.J. Holmes, A periodically forced piecewise linear oscillator. *J. Sound Vibr.* **90**, 129–155 (1983).
84. Y. Zhang, Y. Zhao, Nonlinear dynamics of atomic force microscopy with intermittent contact. *Chaos, Solitons and Fractals* **34**, 1021–1024 (2007).
85. N.A. Burnham, A.J. Kulik, G. Gremaud, G.A.D. Briggs, Nanosubharmonics: the dynamics of small nonlinear contacts. *Phys. Rev. Lett.* **74**, 5092–5095 (1995).
86. J.M.T. Thompson, H.B. Stewart, *Chapter 15: Chaotic motions of an impacting system Nonlinear Dynamics and Chaos*, 2nd edn. (Wiley, 2002).
87. J. Berg, G.A.D. Briggs, Nonlinear dynamics of intermittent-contact mode atomic force microscopy. *Phys. Rev. B* **55**, 14899–14908 (1997).
88. W. van de Water, J. Molenaar, Dynamics of vibrating microscopy. *Nanotechnology* **11**, 192–199 (2000).
89. S.W. Shaw, Forced vibrations of a beam with one-sided amplitude constraint: theory and experiment. *J. Sound Vibr.* **99**, 199–212 (1985).
90. A.B. Nordmark, Non-periodic motion caused by grazing incidence in an impact oscillator. *J. Sound Vibr.* **145**, 279–297 (1991).
91. K.D. Murphy, T.M. Morrison, Grazing instabilities and post-bifurcation behavior in an impacting string. *J. Acoust. Soc. Am* **111**, 884–892 (2002).
92. G.Y. Chen, T. Warmack, T. Thundat, D.P. Allison, A. Huang, Resonance response of scanning force microscopy cantilevers. *Rev. Sci. Instrum.* **65**, 2531–2537 (1994).
93. J.N. Israelachvili, *Intermolecular and Surface Forces with Applications to Colloidal and Biological Systems*, (Academic, 1985).
94. O. Kolosov, K. Yamanaka, Nonlinear detection of ultrasonic vibrations in an atomic force microscope. *Japan. J. Appl. Phys.* **32**, L1095–1098 (1993).
95. Y. Liu, T. Wu, D.E. Evans, Lateral force microscopy study on the shear properties of self-assembled monolayers of dialkylammonium surfactant on mica. *Langmuir* **10**, 2241–2245 (1994).
96. J.E. Sader, L. White, Theoretical analysis of the static deflection of plates for atomic force microscope applications. *J. Appl. Phys* **74**, 1–9 (1993).
97. L.N. Virgin, *Introduction to Experimental Nonlinear Dynamics* (Cambridge University Press, Cambridge, UK, 2000).

Experimental Methods for the Calibration of Lateral Forces in Atomic Force Microscopy

Martin Munz

Summary. As a variation of atomic/scanning force microscopy (AFM/SFM), lateral force microscopy (LFM) relies on the torsional deformation of the AFM cantilever upon lateral forces acting between tip and sample surface. LFM enables imaging of frictional properties of the surfaces of materials, thin films or monolayers at a high spatial resolution. Furthermore, LFM is increasingly used to study the effect of shear loading on nanostructures or nanoparticulates. Albeit a large variety of applications have been demonstrated and the measurement mode is implemented in most commercially available AFM instruments, LFM seems to suffer from the lack of reliable and established calibration methods for lateral forces. However, general acceptance of LFM requires quantification coupled with a solid understanding of the sources of uncertainty. This chapter reviews the available experimental calibration methods. In addition to a description of these methods, a table including information on the key characteristics is provided as well as an overview of the basic equations.

Key words: lateral force microscopy (LFM), friction force microscopy (FFM), atomic force microscopy (AFM), scanning force microscopy (SFM), scanning probe microscopy (SPM), torsional spring constant/stiffness, lateral spring constant/stiffness, sensitivity, optical crosstalk, mechanical crosstalk, friction loop, friction law, contact stiffness, force-distance curve (FDC), reference beam, suspension stiffness, torsional resonance, position-sensitive detector (PSD), microcantilever, micro-electromechanical systems (MEMS), calibration, uncertainty, nanometrology.

Abbreviations

AFM	Atomic force microscopy
CL	Cantilever
D-LFC	Diamagnetic lateral force calibrator
DMT	Derjaguin–Muller–Toporov

FDC	Force-distance curve
FEA	Finite element analysis
FIB	Focused ion beam
JKR	Johnson–Kendall–Roberts
LEN	Lateral electrical nanobalance
LFM	Lateral force microscopy
MEMS	Micro-electromechanical systems
MLFS	Micro-lateral force sensor
NEMS	Nanoelectromechanical systems
PG	Pyrolytic graphite
PSD	Position-sensitive detector
SBP	Scanning-by-probe
SBS	Scanning-by-sample
SNR	Signal-to-noise ratio

9.1 Introduction

Since its invention in 1987 [1], lateral force microscopy (LFM), frequently referred to as friction force microscopy (FFM) as well, has found a steadily increasing number of applications [2–4]. Among others, it has been employed for mapping spatial variations in the surface properties of heterogeneous polymers [5, 6], thin films [7], or surfaces patterned using lithography techniques [8, 9].

Further major drivers for the advancement of nanoscale friction force measurements are emerging or rapidly developing technologies with requirements for nanomechanical characterisation, such as micro- and nanoelectromechanical systems (MEMS/NEMS) [7, 10, 11] or biomedical engineering and biotechnology [12, 13]. Concerning the fabrication process of MEMS structures, Kim et al. [14] employed LFM to evaluate the forces required for the collapse of photoresist patterns. Such patterns tend to fail in the course of the rinsing required for the development process. Hence, in addition to mere imaging of lateral forces, LFM can also be utilised for the nanomechanical testing of thin films or small structures. Owing to the lateral forces acting between AFM tip and sample surface, LFM can be used for combined mechanical loading and lateral force measurement. Wright and Armstrong [12] studied the removal of microbial cells from medically relevant surfaces. Considering that in many cases microbial cells are exposed to hydrodynamic shear fields, the application of lateral forces can be used to measure the critical shear strength of such cell–surface interfaces.

In a similar manner, the shear strength of nanoparticles adhering to a support surface has been studied [15], and Ling et al. [16] have employed LFM to investigate the friction between two single microspheres. As compared to macroscopic techniques averaging over a large number of nanoscale entities,

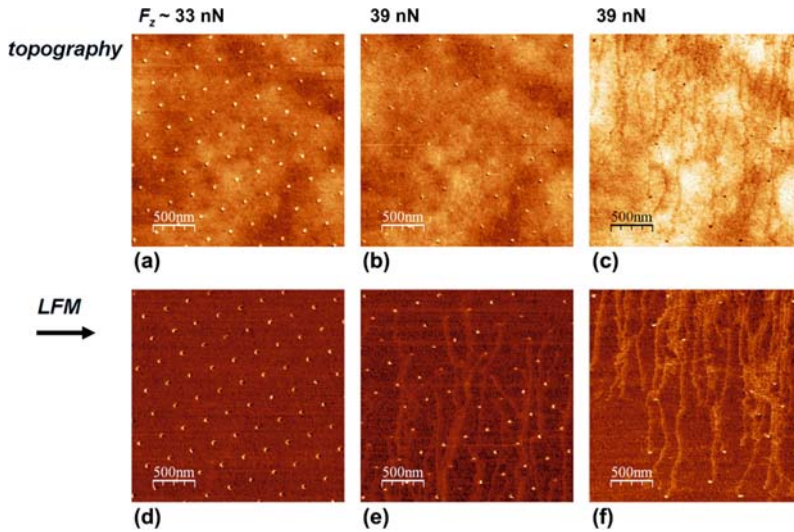


Fig. 9.1. Excerpt from a series of LFM scans on an array of W nanoparticles. (a)–(c) Topography images, and (d)–(f) LFM images, measured upon scanning from *left to right*. Using electron beam deposition (EBD), the nanoparticles were grown on *top* of a gold surface. The images (b) and (e) were measured upon scanning from *top to bottom*, whereas the images (c) and (f) were measured upon scanning from *bottom to top*. The LFM images (e) and (f) show tracks that resulted from loose particles moved along by the AFM tip. Scan length 15.0 μm . AFM system type MFP-3D by Asylum Research

AFM allows for force measurements on single particles, a prerequisite for studies of the relationship between particle shape and the critical force required for their detachment.

An example of nanoparticle detachment experiments is given in Fig. 9.1. The sample consisted of an array of tungsten nanoparticles grown on top of a gold surface [15]. The normal force applied by the AFM tip was step-wise increased. As can be seen from the Figs. 9.1b and c, disintegration and detachment of the nanoparticles occurred around a normal force of approximately 39 nN. The corresponding LFM images (Figs. 9.1e, f) show tracks that originate from positions pertaining to the array of nanoparticles. They can be attributed to loose nanoparticles moved along by the tip. Notably, these tracks also indicate the scan direction along the axis of slow scanning. Figures 9.1b and e were recorded upon scanning from top to bottom, whereas Figs. 9.1c and f were recorded upon scanning from bottom to top.

Although the normal load applied when the detachment of nanoparticles occurs can be indicative of the adhesion strength of the interface with the substrate surface [15], the lateral force applied to the nanoparticles would be more direct a measure of the adhesion strength. Accurate evaluation of the LFM data requires calibration of the lateral force corresponding to the measured

cantilever (CL) twist. Hence this application demonstrates the importance of lateral force calibration for quantitative measurements of the nanoparticle–substrate adhesion strength. A sufficient adhesion strength helps ensure that the nanoparticles do not fall off the surface or cannot be easily swept away.

Furthermore, it should be noted that the dependence of frictional forces on the chemical species present can be utilised for differentiating chemical groups by measuring the force interactions with a chemically functionalised tip [17–19]. Owing to this capability, LFM can be considered as a form of chemical force microscopy (CFM). Frequently, CFM is identified with the measurement of force–distance curves (FDCs). However, the lateral forces measured upon scanning with a chemically functionalised tip were shown to be closely related to the adhesion force values extracted from FDCs [8, 17, 20].

Despite the early and continued demonstration of the use of LFM for a wide range of applications, a lack of established calibration methods almost certainly has impeded its widespread acceptance. In particular, the somewhat reduced pace of development of LFM manifests itself when comparing it with rapidly advancing techniques that rely on the analysis of FDCs, such as molecular pulling [13]. Another need for the development of accurate and reliable calibration techniques is to understand and quantify the fundamental mechanisms governing sliding friction. The existing lack of understanding means that friction modelling frequently involves parameters such as surface energy or interfacial shear strength. Essentially, they are used as fit parameters, and a major means of assessing the resulting data is to cross-check if they are consistent with the values derived from several calibration routines [21].

An AFM CL acts as the core element of force sensing. The force interactions between the tip attached to the CL and the surface under investigation result in deformations of the CL. Using a position-sensitive detector (PSD), these deformations are converted into measurement signals. A bidirectional PSD allows both normal and lateral forces to be detected via corresponding deflections of the CL, i.e. via the CL bending (flexural deformation) and the CL twist (torsional deformation), respectively. In the most common version of the PSD, a laser beam bounces off the back of the CL and the angular deflection of the reflected beam is detected using a segmented photodiode. Essentially, quantitative measurement of lateral forces by means of AFM requires two separate calibrations [22], namely the calibration of the lateral PSD response to convert the measured signal into the twist angle of the CL, and the calibration of the torsional spring constant of the CL that converts the twist angle into a torque. This is analogous to the calibration of normal forces, where both the related optical lever sensitivity and the CL normal spring constant (stiffness) need to be calibrated [23].

Over the last 15 years, a variety of methods for the calibration of lateral forces have been demonstrated. Since the year 2000, an increased level of related activities can be observed. The recent rise of the subject may indicate the widely accepted relevance of LFM as well as the need for an established calibration technique. Indeed a considerable variety of methods

for the calibration of lateral forces in AFM has been suggested. Considering the way a torsional CL deformation is induced, five major groups of calibration methods can be identified. Essentially, it can be induced by scanning across a surface and utilising the frictional forces [16, 24–26], by direct application of a force acting at a position off the long axis of the CL [13, 22, 27–32], by loading the tip-substrate contact in a lateral direction [21, 33–35], by loading a compliant structure of known stiffness [36–38], or by exciting the torsional CL resonance [39, 40]. A detailed table of the particular methods is provided in Sect. 9.8 (Table 9.3).

One obvious way of classification of the available methods is the involvement of a sliding motion, i.e. whether or not the calibration method relies on a scanning motion between tip and sample [34]. Among the scanning calibration techniques are the wedge calibration method [24, 25], and the parallel scan technique where the scan movement is parallel to the CL long axis [41, 42]. In particular, the wedge method is frequently used as an experimental reference technique [30, 34, 36].

Considering that these methods are based on empirical relationships between friction force and normal load, Asay et al. [34] classified them as indirect. Thus, the lateral forces causing the CL twist are not directly known or measured. For instance, techniques [24–26] based on Amontons’ law [43] involving the friction coefficient, μ , assume that its value is constant over time; however, in the case of sharp AFM tips, it is not unlikely to change in the course of the experiment. In particular, the tip can be affected by wear, and an increase in tip radius entails increases in contact area as well as in total adhesion force.

Another point of criticism is that the friction law can be a matter of the tip shape and properties. A friction force scaling in a linear manner with the load, as given by Amontons’ law, is known to hold for multi-asperity contacts only, whereas contacts with a single-asperity are more likely to follow a power law, as given by the Johnson–Kendall–Roberts (JKR) or the Derjaguin–Muller–Toporov (DMT) model [44–47].

A classification of the calibration method, in terms of direct or indirect, has been applied by various authors, however, with slight differences in meaning. In line with the elucidation given earlier, a calibration method can be referred to as direct if the force applied to the CL is known from basic principles or experiments, e.g. by measurement of the deformation of a calibrated reference spring. However, other authors [26, 36] also used the term ‘direct’ for one-step techniques, i.e. calibration methods that yield a lateral force calibration factor in a single step, without the requirement of a separate measurement for the determination of the PSD sensitivity [26, 35]. In these terms, the wedge method is a direct technique, since it delivers an overall lateral calibration constant and does not require separate measurements for the determination of torsional CL spring constant and lateral PSD sensitivity.

This chapter is organised as follows: the basic definitions of the calibration constants related to LFM as well as some relationships between these

constants are given in Sect. 9.2. In Sect. 9.3, an overview of the available methods for the calibration of the lateral sensitivity of the PSD is given. Sections 9.4–9.7 provide an overview of the methods that aim at the calibration of lateral forces. Methods relying on a scanning movement between tip and substrate are described in Sect. 9.4. An outline of the configurations employing a force balance between the tip and a rigid structure is presented in Sect. 9.5, and configurations involving compliant structures are considered in Sect. 9.6. Afterwards, in Sect. 9.7 methods are presented that utilise torsional CL resonances. Finally in Sect. 9.8, a discussion of various methods is given. Largely, it is in the form of a table that provides an overview of the methods and their key characteristics. Furthermore, it indicates the particular calibration constants quantified by each method and the reported uncertainties. For the purpose of clarity, such details are not typically provided in the main text. Rather, the main text aims at an outline of the various approaches.

9.2 Basic Definitions and Relationships

9.2.1 The Calibration Constants Involved in a Lateral Force Measurement

The quantitative analysis of LFM data requires the knowledge of the factor linking the measured signal either with the force or with the deformation. In general, lateral force calibration involves two steps [22]:

1. The calibration of the lateral PSD response to convert the measured signal to the twist angle of the CL
2. The calibration of the torsional spring constant of the CL that converts the twist angle to a torque.

The related calibration factors are the torsional sensitivity of the PSD, S_ϕ , and the torsional spring constant, k_ϕ . Assuming small deflections of the CL, the relationships between the acting forces, the deformations of the CL, and the resulting measurement signals are of a linear form. Hence, the output voltage ΔV_ϕ of the lateral direction PSD can be written as

$$\Delta V_\phi = S_\phi \Delta\phi, \quad (9.1)$$

where $\Delta\phi$ is the angle of torsional deformation (twist) of the CL, and S_ϕ is the torsional angle sensitivity of the PSD (in V/rad). It depends on the geometry of the laser beam path and frequently on the total signal on the PSD also. Hence, S_ϕ needs to be calibrated every time the position of the laser beam or of the PSD changes.

The linear relationship between the torsional CL deformation and the acting torque, T , involves the torsional CL spring constant, k_ϕ :

$$T = k_\phi \Delta\phi. \quad (9.2)$$

Correspondingly, k_ϕ is measured in Nm/rad. As a measure for the amount of deformation resulting from a certain torsion moment or force, the spring constant is a stiffness by nature. In principle, both the terms spring constant and stiffness can be used interchangeably. For the purpose of clarity and to align with the standardised vocabulary for surface chemical analysis [48], however, it is advisable to refer to the CL stiffness as spring constant.

Furthermore, it is worthwhile noting that as a quantity characterising the CL mechanical properties, k_ϕ is more fundamental than S_ϕ . By definition, it is independent of the characteristics of the readout system for the CL deflection. Knowledge of the overall conversion factor α (Table 9.1) is sufficient to convert the readout values into force units. However, α changes with every repositioning of the laser focus or of the segmented photodiode. Thus, quantitative comparison of various CLs or comparison with alternative calibration techniques requires knowledge of the torsional spring constant, k_ϕ . Notably, the overall conversion factor, α , is the representative of one-step calibration techniques that do not involve a separate measurement for the determination of the PSD sensitivity (Sect. 9.1).

An overview of the defining equations is given in Table 9.1, and a schematic representation of the mutual relationships of the calibration constants is shown in Fig. 9.2. For the purpose of distinction from the other equations, the defining equations are referred to by the numbers (D1)–(D9).

These definitions establish a loop that can be represented by a triangle (Fig. 9.2). Once two calibration constants are known, the third one can be calculated. Combination of the equations given in Table 9.1 yields the following relationships between the calibration constants:

$$\alpha_z = \frac{k_z}{S_z} \tag{9.3a}$$

$$\alpha_\phi = \frac{k_\phi}{S_\phi}. \tag{9.3b}$$

Table 9.1. Equations that define the calibration constants related to normal and torsional deformations

	PSD sensitivity	CL spring constant	Overall conversion factor
Torsional	$\Delta V_\phi = S_\phi \Delta\phi$ (D1)	$T = k_\phi \Delta\phi$ (D2)	$T = \alpha_\phi \Delta V_\phi$ (D3)
(in terms of $\Delta\phi$)			
lateral (in terms of Δx)	$\Delta V_\phi = S_x \Delta x_\phi$ (D4)	$F_x = k_x \Delta x_\phi$ (D5)	$F_x = \alpha_x \Delta V_\phi$ (D6)
normal	$\Delta V_z = S_z \Delta z$ (D7)	$F_z = k_z \Delta z$ (D8)	$F_z = \alpha_z \Delta V_z$ (D9)

The vertical displacement and the twist angle are denoted by Δz and $\Delta\phi$, respectively, and the related normal force and torque are denoted by F_z and T , respectively. The PSD sensitivities, S_z and S_ϕ , are measures for the signal changes upon deformation of the CL. The CL spring constants, k_z and k_ϕ , are measures for the forces related to the deformations. The conversion factors, α_z and α_ϕ , are measures for the changes in the measurement signal upon torque acting on the CL

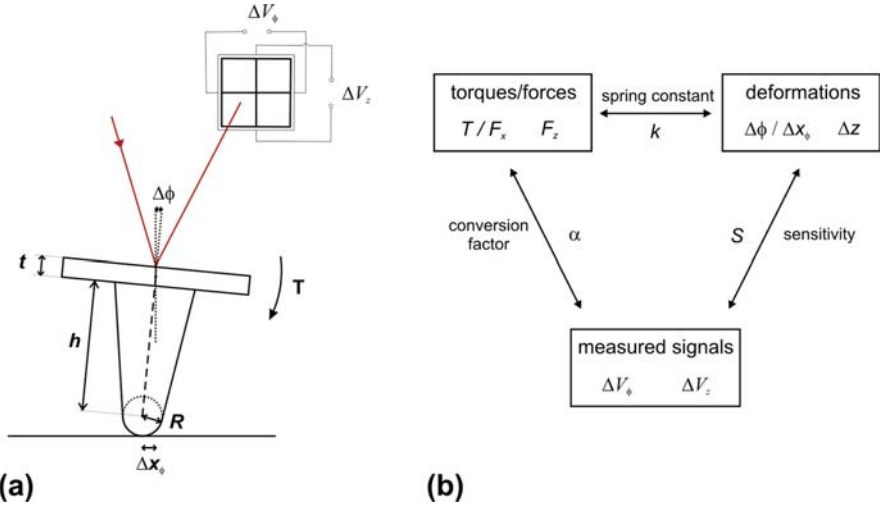


Fig. 9.2. (a) Schematic representation of the core-sensing elements of an AFM instrument, consisting of a CL with integrated tip, laser beam, and segmented photodiode. In the case of a lateral force acting at the apex of the tip, the length of the torsional arm is given by $H = R + h + t/2$, where R is the radius of curvature of the tip apex, h is the tip height, and t is the thickness of the CL beam. (b) Schematic representation of the calibration triangle, illustrating the relationships between the applied forces, the resulting deformations, and the measured signals. Such relationships can be applied to both the flexural and the torsional CL deformations

A torque, T , is applied to the CL by the lateral force, F_x , acting via a lever arm of length H , between the CL neutral axis and the contact point. The torque causes a CL twist, and for small values of the twist angle, $\Delta\phi$, the related displacement at the tip-sample interface, Δx_ϕ , can be approximated by

$$\Delta x_\phi = H \Delta\phi. \tag{9.4}$$

In the case of a symmetric tip position and a symmetric CL beam, it acts as a torsional arm of length H , which is given by the sum of the tip radius, R , the tip height, h , and half of the CL thickness, t , i.e. $H = R + h + t/2$ (Fig. 9.2a).

In addition to the torsional CL spring constant, k_ϕ , also a lateral CL spring constant, k_x , can be defined:

$$F_x = k_x \Delta x_\phi. \tag{9.5}$$

The related calibration constants are given in the second row of Table 9.1.

9.2.2 Basic Relationships Involving the Calibration Constants

Further to these basic equations, a couple of related equations and comments shall be given:

1. In vector representation, the torque, \mathbf{T} , acting on the CL is given by the vector product of the distance vector, \mathbf{q} , and the lateral force, \mathbf{F}_{lat} :

$$\mathbf{T} = \mathbf{q} \times \mathbf{F}_{\text{lat}}. \quad (9.6)$$

Hereby, \mathbf{q} is the distance vector between the tip apex and the CL long axis. Assuming a perfectly symmetric CL, the tip apex is situated exactly beneath the CL long axis, and the CL shear centre is situated on its long axis. In this case, the amount of the torque can be written as $T = |\mathbf{q} \times \mathbf{F}_{\text{lat}}| = qF_x$. As the force acting perpendicularly to the tip, the lateral force, \mathbf{F}_{lat} , is horizontal if \mathbf{q} is vertical. Then it can be identified with F_x , the force acting along the x -axis. As discussed in Sect. 9.5.5, this does not hold in the case of a positional offset between tip and CL shear centre.

For the case of a perfectly symmetric CL, \mathbf{q} is vertical, and (9.6) can be written in terms of the total tip height, H :

$$T = HF_x. \quad (9.7)$$

2. The total tip height, H , represents the lever arm along which the lateral force, F_x , generates a torque and causes a CL twist. By comparing the equations (D1) and (D4) and inserting (9.4), an alternative equation for the lateral PSD sensitivity can be derived:

$$S_x = S_\phi / H. \quad (9.8)$$

3. Furthermore, inserting the equations (9.4) and (D1) into (D5) yields an expression for the lateral force, F_x , in terms of the voltage ΔV_ϕ :

$$F_x = \frac{k_x H}{S_\phi} \Delta V_\phi = \frac{k_x}{S_x} \Delta V_\phi. \quad (9.9)$$

4. The combination of (9.4) with the expressions for the torsional and the lateral spring constants ((D2) and (D5)) as well as with (9.6) results in the following relationship:

$$k_\phi = H^2 k_x. \quad (9.10)$$

Using this equation in combination with (9.3b), the lateral CL spring constant, k_x , can be written in terms of the torsional conversion factor, α_ϕ :

$$k_x = \frac{\alpha_\phi S_\phi}{H^2}. \quad (9.11)$$

5. The lateral conversion factor, α_x , is related to the torsional conversion factor, α_ϕ , by the following relationship:

$$\alpha_x = \frac{\alpha_\phi}{H}. \quad (9.12)$$

Equations (9.8), (9.10), and (9.12) provide direct links between the relationships based on the torsional angle, $\Delta\phi$, and the relationships based on the lateral displacement, Δx_ϕ .

6. For the convenient comparison of the normal and the lateral conversion factors, both the lateral and the normal PSD sensitivities can be defined in terms of the angular deflections. Denoting these angles with $\Delta\phi$ and $\Delta\eta$, respectively, the PSD sensitivity, \hat{S}_z , related to the angular deflection, $\Delta\eta$, resulting from the vertical tip displacement, Δz , is defined by:

$$\Delta V_z = \hat{S}_z \Delta\eta. \quad (9.13)$$

This alternative PSD sensitivity, \hat{S}_z , is measured in V/rad. Considering the relationship between Δz and $\Delta\eta$, as resulting for a beam of effective length L :

$$\Delta z = \frac{2}{3} L \Delta\eta, \quad (9.14)$$

the normal force, F_z , can be written as:

$$F_z = k_z \Delta z = \frac{2}{3} k_z L \Delta\eta \equiv \hat{k}_z \Delta\eta, \quad (9.15)$$

with the alternative spring constant, \hat{k}_z , in N/rad. Similarly, an alternative formulation for the normal PSD sensitivity can be given by:

$$S_z = \frac{\Delta V_z}{\Delta z} = \frac{3}{2} \frac{1}{L} \frac{\Delta V_z}{\Delta\eta} \equiv \frac{3}{2} \frac{1}{L} \hat{S}_z. \quad (9.16)$$

7. Assuming validity of Amontons' law for the given tip-sample system, the lateral force signal, ΔV_ϕ , and the CL twist angle, $\Delta\phi$, scale in a linear manner with the applied load, F_z [49]. Thus, when writing $\Delta V_\phi = m F_z$ and $\Delta\phi = \tilde{m} F_z$ with the constant factors m and \tilde{m} , the defining equation for the torsional PSD sensitivity, (D1), transforms into a particularly simple expression:

$$S_\phi = \frac{\Delta V_\phi}{\Delta\phi} = \frac{m}{\tilde{m}}. \quad (9.17)$$

9.2.3 The Lateral and the Normal Spring Constant of a Rectangular CL

For a homogeneous CL of rectangular shape, the normal as well as the lateral spring constant can be calculated from its dimensions and the moduli of its material. The lateral spring constant, k_x , is given by:

$$k_x = \frac{G w t^3}{3 L H^2}, \quad (9.18)$$

where G is the shear modulus of the CL material. w , t , and L are the CL width, thickness, and the distance from the fixed end of the CL to the position near its

free end where the tip is situated. For an isotropic and homogeneous material of Poisson's ratio ν , the shear modulus, G , and the Young's modulus, E , are coupled via the equation $G = E/(2(1 + \nu))$ [50].

Similarly to k_x , the normal CL spring constant, k_z , can be calculated from:

$$k_z = \frac{Ewt^3}{4L^3}. \quad (9.19)$$

Herewith, E denotes the Young's modulus of the CL material.

Analytical equations are more difficult to derive for the case of V-shaped CLs. Neumeister and Ducker [51] used approximations such as a triangular plate or two prismatic beams.

Importantly, the pronounced dependence of k_z and k_x on the CL thickness, t , means that any uncertainty in the measurement of t is bound to entail a large uncertainty in the CL spring constants. Similarly, spatial variations in the CL thickness can lead to a large difference between the calculated and the real CL spring constant. For instance, spatial variations in the CL thickness or in its shear modulus can result from imperfections related to irregularities in the processing steps involved in the CL fabrication [52].

A combination of (9.18) and (9.19) yields the ratio of the lateral to the normal spring constant:

$$\frac{k_x}{k_z} = \frac{4G}{3E} \left(\frac{L}{H} \right)^2. \quad (9.20)$$

The ratio k_x/k_z is independent of the CL dimensions w and t , since both spring constants scale with wt^3 . With G/E in the order of 1 and $L/H \gg 1$, the value of the ratio k_x/k_z is $\gg 1$. Its value increases with the length L . Furthermore, it decreases with increasing values of the tip height, H . The scaling of k_x/k_z with H^{-2} (9.20) accounts for the fact that the tip acts as a moment arm. For a higher tip, a certain lateral force causes a larger torque and hence a larger torsional deformation of the CL. Thus the lateral CL stiffness, k_x , decreases with increasing values of H . Since k_z is independent of H , the ratio k_x/k_z shows the same scaling as k_x .

Notably, (9.20) can be used to calculate the lateral CL spring constant, k_x , if the normal spring constant, k_z , is known [39,53,54]. However, propagation of the uncertainties in k_z as well as in the dimensions, H and L , and in the elastic properties, E and G , can lead to a considerable uncertainty in k_x . Moreover, the shear modulus, G , is frequently expressed in terms of the Young's modulus, E , by employing the equation $G = E/(2(1 + \nu))$ [50]; however, this approach is valid only for CLs made of isotropic materials. Indeed, in the case of single crystal Si CLs, this approach was found to result in large uncertainties, as compared to the case of CLs made of Si_3N_4 [39].

Due to the scaling of k_x and k_z with t^3 (9.19) and (9.20), any method involving thickness measurements is inherently error-sensitive [55]. For applications beyond the quick estimation of the CL spring constant, mitigation

strategies should be applied, such as the replacement of the CL thickness with a parameter related to its effective thickness [56]. Considering that the fundamental flexural resonance frequency is directly related to the CL spring constant, it is evident that it can be utilised as a measure for its thickness [39, 56]. From the Euler–Bernoulli equations for a rectangular and homogeneous beam, an equation for the CL thickness can be derived [35, 57, 58]:

$$t = \frac{\omega_n}{K_n^2} \sqrt{12 \frac{\rho}{E}}. \quad (9.21)$$

Herewith, ω_n is the n th flexural resonance frequency, K_n is the related wavenumber, and ρ is the mass density. For the fundamental mode ($n = 1$), the related value of the wavenumber is given by $K_1 L_0 = 1.875$, with the CL length, L_0 .

It should be noted that the equations given earlier assume a rectangular CL geometry and perfectly homogeneous materials properties. However, if the CL geometry becomes more complex and variations in its local materials properties and its coatings are significant, then the simple beam mechanics approach leads to large uncertainties [54, 59].

Furthermore, it is worthwhile mentioning that also the torsional resonance can be utilised for lateral force calibration (Sect. 9.7). Similar to the flexural resonances of the CL, also torsional resonance frequencies depend on the CL spring constant, mass distribution, and dimensions. A comprehensive review of the CL dynamic properties and the related dynamic AFM operation modes has been given by Song and Bhushan [60].

9.2.4 The Case of In-Plane Deformations

The particulars of the CL geometry are also of relevance for the amount of CL deformation parallel to its surface plane. Throughout the existing calibration methods, it is assumed that the lateral force acting on the tip leads purely to a torsional CL deformation. In principle, however, it can also lead to an in-plane deformation. This deformation mode can be neglected only within certain ranges of the CL geometry parameters, such as its width, thickness, or tip height. Considering the spring constants for torsional and in-plane deformation, the torsional deformation is dominating if the following condition is satisfied [61]:

$$\lambda = \frac{k_\phi}{k_{\text{in}}} \ll 1. \quad (9.22)$$

Herewith, k_{in} denotes the spring constant for in-plane deformation of the CL.

For rectangular CLs with $L \gg w$, k_{in} is given by [61]:

$$k_{\text{in}} = \frac{1}{4} Et \left(\frac{w}{L} \right)^3. \quad (9.23)$$

Notably, this expression refers to a lateral loading perpendicular to the CL long axis. In combination with the expression for the lateral spring constant (9.18), the ratio λ is given by:

$$\lambda = \frac{4G}{3E} \left(\frac{tL}{Hw} \right)^2. \quad (9.24)$$

Clearly, the condition $\lambda \ll 1$ (9.22) is satisfied if $t \ll H$ and $L \ll w$, since the ratio G/E is in the order of 1. This simple consideration shows that for CLs with $L \gg w$, (9.22) does not necessarily hold and that the assumption of negligible in-plane deformation is invalid [61]. For instance, with $H = 10 \mu\text{m}$, $t = 2 \mu\text{m}$, $L = 400 \mu\text{m}$, and $w = 40 \mu\text{m}$, the ratio λ has the value 2.13, if the shear modulus is replaced with $G = E/(2(1 + \nu))$ [50] and if $\nu = 0.25$. For a CL of length $L = 200 \mu\text{m}$, however, the value of λ is 0.53, that is, < 1 .

Since a lateral movement of the CL surface has no effect on the angle of incidence at the position of the laser spot, the laser deflection technique is not sensitive for in-plane bending of the CL. Owing to the fact that the in-plane bending is negligible for typical CLs with $w \gg t$, however, the torsional CL deformation dominates and it is justified to rely on the laser deflection technique. Nevertheless, this assumption should be validated for the particular CL under consideration [61].

9.3 Calibration of the Lateral Sensitivity of the PSD

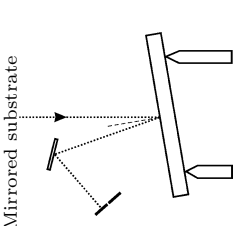
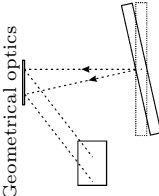
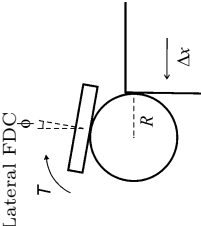
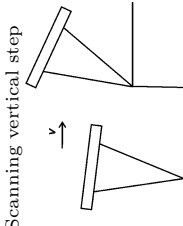
9.3.1 Available Methods

Various methods for the measurement of the lateral deflection sensitivity have been suggested. In the simplest case, the PSD is assumed to be rotationally symmetric, that is, to be equally sensitive for normal and for torsional angular deflections. In terms of the PSD sensitivities defined in Table 9.1, this approach can be expressed as $S_x = S_z$. However, this is not likely to be the case. Possible reasons for different values of the normal and the lateral PSD sensitivities are asymmetries in the shape of the reflected laser spot [35] and different electronic amplification factors for the normal and the lateral PSD signals. Hence, dedicated methods for the calibration of the lateral PSD sensitivity are required. An overview of the available methods is given in Table 9.2.

Mirrored Substrate Method

The lateral sensitivity of the PSD can be measured by mimicking the torsional deformation of the CL and recording the change in the LFM signal. For instance, the angle between PSD and reflecting surface can be altered by

Table 9.2. Overview of methods for calibration of the PSD sensitivity

Group of methods	Approach	Major advantage	Main source of uncertainty	Refs.
Mirrored substrate	 <p>The laser beam is reflected off a mirrored substrate and the PSD output measured as a function of the substrate tilt angle</p>	Non-contact technique (i.e. no risk of tip damage)	Different conditions of the mirrored substrate (distance from the mirror, shape of the reflected laser beam cross-section)	[22] [29]
Geometrical optics	 <p>Geometrical analysis of the optical beam path and measuring the signal changes upon displacement of the PSD</p>	Non-contact technique (i.e. no risk of tip damage)	Non-linearities in the detector system (incl. mechanical backlash in the PSD positioner)	[21] [62] [63]
Lateral FDC	 <p>Pressing a colloidal probe (test probe) against a vertical and rigid step, recording the changes in the LFM signal as a function of the lateral displacement (lateral FDC)</p>	A standard calibration grating can be used; no risk of damage to the tip of the target CL since a test probe of similar mech. and optical properties is used	Length of the torsional moment arm, both for the test and the target probe	[30]
Scanning vertical step	 <p>Scanning an integrated tip across the vertical step of a calibration grating</p>	Requires no particular experimental measures, only a standard calibration grating and a sharp tip	Assumes that the global tip shape applies to the tip apex; relies on the readout of the peak voltage in the lateral force signal	[49]

using a mirrored substrate and by tilting it by means of a stepper motor [64]. In a similar approach, the AFM head can be tilted [29]. By recording the associated changes in the lateral force signal, ΔV_ϕ , and using (D1), the PSD sensitivity, S_ϕ , can be calculated.

A highly linear relationship between the PSD response and the tilt angle was observed by Rutland's group [13, 29]. The laser spot position on the PSD did not show a significant effect on the PSD sensitivity [13]. On the other hand, Cannara et al. [30] found a parabolic dependence of the lateral sensitivity on the lateral laser spot position. Such findings are reminiscent of the fact that the PSD response depends on the particulars of the instrumental design. Thus, it should be borne in mind that the instrument to be used for LFM measurements should be thoroughly tested [13]. As pointed out by Asay et al. [34], in many cases the assumption of a negligible optical crosstalk does not hold, thus leading to a poor reproducibility of measurements performed in different laboratories.

Geometrical Optics Method

As an alternative to tilting the AFM head or a mirrored substrate, the PSD response can be measured as a function of its displacement, that is, by moving it along its two major axes [62]. The PSD sensitivity can be calculated by means of a suitable model for changes in the beam path as the CL bends and twists [21]. Using such an approach based on geometrical optics, Liu et al. [62] converted the PSD displacement into an equivalent lateral tip displacement. The method requires knowledge of the tip height and of the lateral CL spring constant, k_x . The equation derived by Liu et al. [62] relates the lateral force, F_x , to the friction loop width, W . It involves the sensitivities of the PSD for vertical and horizontal movements [21, 62].

As reported by Cain et al. [21], the sensitivity for horizontal PSD movements can depend on the level of the bending signal, V_z , that is, on the setpoint value of the topography feedback. In such a case, the scaling of raw friction data, V_ϕ , is a function of the load applied during the friction measurement. A similar dependence was also found for friction versus load data measured on a polished SrTiO₃ (305) surface.

In many cases, the measurements need to be undertaken with the sample surface covered by a liquid, thus requiring the usage of a fluid cell. In particular, this is more the rule than the exception when it comes to experiments involving biological or biochemical specimens. With a refractive index, n , of the liquid different from the one of air ($n \approx 1$), it affects the path of the laser beam. When comparing the results for air and water, Pettersson et al. [13] observed a significant change in the sensitivity, S_ϕ . Hence, the PSD calibration needs to be performed under the same conditions as the real measurement. More recently, Tocha et al. [63] presented a geometrical analysis accounting for the different refractive indices of air and the liquid contained in the fluid

cell. They derived a correction factor that relates the PSD response for measurements in air and in the liquid to be used. This correction factor is given by the ratio of the refractive indices of the two media [63].

Alternatively to the approach taken by Liu et al. [62], torsional CL deflections can also be generated via an offset in the tip position. In the case of an AFM measurement, such a positional offset from the CL long axis causes an apparent frictional force even without any sliding movement (Sect. 9.5.5). In the course of the calibration of the PSD response for CL deflection by going into contact with a hard and stiff substrate, that is, by a purely normal loading, an offset in the tip position results in a torque and a corresponding torsional CL deformation. However, for typical CLs the vertical tip displacement related to the torsional deformation of the CL, $\Delta z_\phi = \Delta z - \Delta z_{\text{norm}}$, is comparatively small, owing to a large ratio of torsional and normal CL spring constants, $k_\phi/k_z \gg 1$. Bogdanovic et al. [29] reported that the relative amount of Δz_ϕ was less than 10%. Nevertheless, the occurrence of such a mechanical crosstalk (Sect.9.5.5) underscores the necessity of placing the colloidal probe as close as possible to the long axis of the CL.

Lateral FDC Method

Approaches relying on geometrical optics or a mirrored substrate circumvent the usage of an AFM CL. However, under certain conditions the calibration can also be achieved using an AFM probe. Given a well-defined contact geometry and a rigid structure, a lateral displacement, Δx , can be assumed to be fully converted into a torsional deflection of the CL beam. The approach chosen by Cannara et al. [30] takes benefit of the sphere-like shape of a colloidal probe. By approaching it to a vertical step, they recorded lateral FDCs. According to (D4), the required sensitivity, S_x , is given by the slope of the linear compliance regime of such lateral FDCs (see Sect. 9.6.1).

In greater experimental detail, a colloidal glass sphere of approximately 70 μm diameter was attached to a Si_3N_4 test CL with 40 μm width. A lateral force-displacement measurement was achieved by approaching a vertical step on the surface of a freshly cleaved GaAs wafer [30]. Owing to its cleavage behaviour, a plane at a perfect 90° angle to the (100) surface can be generated, the $(\bar{1}10)$ plane of the GaAs crystal.

In order to allow for calibration of standard CLs with an integrated tip or of colloidal probes with a small diameter, they suggested a procedure utilising the colloidal probe as a test probe. Under the assumption of an unchanged condition of the PSD, the sensitivity can be calculated for a CL of arbitrary length provided that the sensitivity was measured for the test probe. However, the width and reflectivity of the target CL should not deviate significantly from the ones of the test probe. A major advantage of the test probe method is that the risk of damage to the tip of the target CL is eliminated.

It should be mentioned that in addition to the torsional response, also lateral in-plane bending may occur (Sect. 9.2.4). Hence both the in-plane bending

of the test and the target CL should be negligible. Under the assumption that the calibration is independent of the laser spot position if it is situated between the tip and the free end of the CL, Cannara et al. [30] provided an equation to correct for in-plane bending of the target probe.

Furthermore, calibration of the target CL requires accounting for differences in the torsional moment arm length as well as in the total signal on the PSD.

Scanning Across a Vertical Step

In the search for a simple calibration technique without any special experimental requirements, Choi et al. [49] suggested a method involving only a standard CL and a grating featuring vertical steps. Similar to the approach of Cannara et al. [30], the method is based on the changes in lateral force upon contact with a vertical wall. However, the method of Choi et al. [49] employs a sharp tip rather than a colloidal probe and considers the lateral forces at two distinct points of contact, namely a point on the flat area and the point at the upper edge of the step. At the latter, the lateral force signal is assumed to show a maximum.

An expression for the CL twist angle can be derived from the balance of the torsional moments. This expression is quadratic in the CL tilt angle measured on the flat area [49]. Provided knowledge of the torsional CL spring constant, k_ϕ , the quadratic equation delivers a value for the CL twist angle, and combination with (D1) yields the torsional PSD sensitivity.

In their experimental approach, Choi et al. [49] used rectangular CLs and calculated the k_ϕ values via the linear-elastic continuum-mechanical model, that is employing the equations (9.18) and (9.19), with the CL thickness measured indirectly via the torsional resonance frequency of the CL.

It should be noted that this method assumes a conically symmetric global tip shape only, as expressed in terms of a tip-opening angle, ψ . In reality, however, the local tip shape can deviate significantly from the shape of the global tip (tip shank). Typically, the local tip shape is accounted for by describing the tip apex in terms of a sphere of radius R . The approach involving exclusively the global tip shape is more likely to deliver useful results in case of very sharp tips, that is, for very small values of R .

9.3.2 Optical Crosstalk

A major complication of LFM measurements can result from a rotational misalignment of the PSD [41, 65]. Such a misalignment causes a crosstalk between the PSD signals related to the normal and the torsional deflections of the CL. It leads to systematic errors in lateral force as well as in height measurements (Fig. 9.3). In particular, such a crosstalk can be dominating in the case of nanoscale structures featuring significant variations in frictional

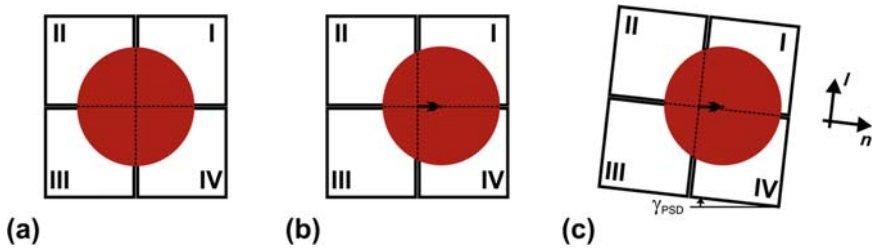


Fig. 9.3. Schematic representation of the optical crosstalk effect. (a) The ideal case of a perfectly symmetric and centred laser spot. The segmented photodiode comprises four segments, I–IV. (b) The laser spot is shifted along the x -axis, thus indicating a torsional deformation of the CL. (c) Result of a misalignment of the segmented photodiode. The misalignment of the PSD is given by the angle γ_{PSD} . In the coordinate system associated with the segmented photodiode, (l, n) , the vector representing the lateral shift of the laser spot has a non-zero component in vertical direction

properties [65, 66], that is, the superposition of significant variations in the lateral force signal onto an almost constant normal force signal.

It should be pointed out that the crosstalk related to a misalignment of the PSD (optical crosstalk) needs to be differentiated from the mechanical one, resulting from a positional offset of the CL shear centre (Sect. 9.5.5). Furthermore, it should be borne in mind that crosstalk can also result from signal mixing effects in the electronic parts of the AFM instrument [67].

Crosstalk between the lateral and the normal responses is likely to occur, and for many AFM instruments it is difficult to eliminate. For instance, if the PSD is not oriented strictly perpendicular to the incident laser beam, the projection of the beam onto the PSD surface is distorted [41, 67]. In this way, the spatial distribution of the light intensity on the PSD surface does not exactly reflect the deformation state of the CL.

Experiments by Cannara et al. [30] have shown a decrease in the lateral deflection sensitivity, S_ϕ , with increasing offset of the PSD position. Approximately, the observed dependence was of a quadratic form. This finding was attributed to the Gaussian distribution of the laser spot. That is, the PSD output may also be affected by the intensity distribution of the reflected laser spot, depending on its extent as compared to the area of the PSD. In order to reduce such non-linearities, it is advisable to avoid large offsets between the laser spot and the centre of the PSD.

Prunici et al. [67] analysed the amount of optical crosstalk in terms of a misalignment factor deduced from the changes in friction loop with sample height. They observed strong changes in the misalignment factor with changes in the position of the reflection point on the CL. In particular, for narrow CLs strong variations were found due to the amount of laser light not reflected by the CL.

In the case of a scanning-by-probe (SBP) scheme, an optical crosstalk can occur even for a free CL, that is, when it is in its undeformed condition [68]. On the contrary to a scanning-by-sample (SBS) scheme, the scanning and the measurement functions are not separated. If the axes of the scanning system are not properly aligned with the axes of the PSD, the reflected lased spot does not move along the PSD axes even if the scanner motion is purely along one of its inherent axes. Varenberg et al. [68] devised an electronic circuit to compensate for the linear crosstalk characteristic observed when approaching the CL to the sample surface.

More recently, the complete detection system consisting of CL and the optical beam path has been modelled to enable a simulation of the LFM system and to perform parametric studies [69]. The model involves a finite element analysis (FEA) of the CL deformation and matrix algebra to describe the path of the reflected laser beam. The access to parametric simulations allows for analysis of the effects of the system variables. In particular, the effect of misalignments in the optical beam path on the measurement signals can be studied in a systematic manner. Hence, crosstalk effects can be identified more easily and the consideration of crosstalk coefficients is important for quantification.

9.4 Methods Relying on a Scanning Motion

9.4.1 The Wedge Method

When plotting the lateral force versus the lateral position, x , for both the forward and the backward scan directions, a loop results. It is generally referred to as the friction loop. Essentially, it is defined by two parameters, namely its width, W , and its centre line, Δ . Both these parameters are related to the torsional moment acting on the tip. In the case of a positive slope, the forward and the backward scans are related to uphill and downhill motions, respectively. The uphill motion causes an additional contribution to the torsional moment acting on the CL. This effect has been referred to as the ratchet mechanism [70]. With the local slope angle θ , the related contribution scales with $\tan \theta$.

By denoting the related torsional moments as T^u and T^d , respectively, the friction loop parameters W and Δ can be written as follows:

$$2W = T^u - T^d \quad (9.25a)$$

$$2\Delta = T^u + T^d. \quad (9.25b)$$

For a perfectly symmetric CL with a tip exactly positioned on its long axis, the lateral forces acting during the uphill and the downhill motions differ in direction only but are identical in amount, and the difference in related torques is constant. It should be noted that potentially the frictional forces

depend on the direction of scanning even in the case of a perfectly flat surface; however, this case seems to be rare and is widely neglected.

Thus, when evaluating lateral force data, the friction loop width, W , is typically considered rather than the lateral force signal, V_ϕ , since the lateral force contribution related to topography is eliminated as well as any offset in V_ϕ .

In terms of the measured PSD signal, the moments acting on the tip can be written as follows:

$$T^{(\tilde{\kappa})} = \alpha_\phi \Delta V_\phi^{(\tilde{\kappa})}, \quad (9.26)$$

with $\tilde{\kappa} = u$ and $\tilde{\kappa} = d$ for an uphill and a downhill scan, respectively.

By expressing the moments in terms of the applied load, F_L , and the adhesion force, F_A , the following equation for the friction loop width, W , can be derived [25]:

$$\frac{W}{H} = \mu \frac{F_L + F_A \cos \theta}{\cos^2 \theta - \mu^2 \sin^2 \theta}. \quad (9.27)$$

Similarly, for the centre line value, Δ , of the friction loop, the following equation holds:

$$\frac{\Delta}{H} = \frac{\sin \theta \{F_L \cos \theta + \mu^2 (F_A + F_L \cos \theta)\}}{\cos^2 \theta - \mu^2 \sin^2 \theta}. \quad (9.28)$$

From the ratio of (9.27) to (9.28), an equation can be derived that is quadratic in μ :

$$\sin \theta (F_A + F_L \cos \theta) \mu^2 - \frac{\Delta}{W} (F_L + F_A \cos \theta) \mu + F_L \sin \theta \cos \theta = 0. \quad (9.29)$$

The solutions of (9.29) depend on the quantities Δ and W , resulting from the lateral force measurement, as well as on the forces F_L and F_A . In particular, this means that calculation of μ requires calibration of the normal CL spring constant, k_z . Thus, strictly speaking the wedge method is not truly a single-step method, although it provides an overall torsional conversion factor.

The calibration grating used by Varenberg et al. [25] features grooves with a trench-like structure. Slopes of an angle of approximately $\pm 54.7^\circ$ connect the flat bottom areas of the trenches with the flat ridge areas separating adjacent trenches. These slope angles are given by the Si crystal lattice. When scanning across such grooves, excellent feedback performance is required to ensure a reasonably constant load. Furthermore, in the case of colloidal probes, the limited width of the grooves can be an issue. Although the method is suitable for colloidal probes, for particle diameters larger than $2R \approx 3.9 \mu\text{m}$ it cannot reach the planar bottom of the grating [71].

It should be mentioned that the slope angle of 54.7° is comparatively large and thus demanding for the topography feedback. This slope angle results

from the structure of the Si crystal lattice, but it is not necessarily an optimal angle for LFM measurements. Moreover, for the given step height of $1.80\ \mu\text{m}$, the lateral extent of the slope is only $1.27\ \mu\text{m}$. For smaller angles, the slope is more extended and the corresponding plateau in the lateral force profile encompasses a larger number of data points, thus enhancing the statistics of the measurement. In addition, steep slopes are more likely to result in departures from the set load and in variations in the tip-sample contact area. In turn, such variations will affect the lateral force signal.

To slant the edges of the calibration grid, focused ion beam (FIB) milling can be applied. In this way, a section of the right-sided edge of a trench was milled. The resulting slope angle was approximately 18.5° (with a global tilt angle of -2.5° , a slope angle of approximately 16.0° was measured). With a lateral extent of approximately $2.8\ \mu\text{m}$ (Fig. 9.4a), the slanted area was more than double as wide as the initial slope. Furthermore, it is interesting to note that the edge on the opposite side of the trench was undercut.

Whereas the flat areas of the TGF11 calibration grid are pertaining to (100) planes of the Si crystal lattice, the sloped area was created by means

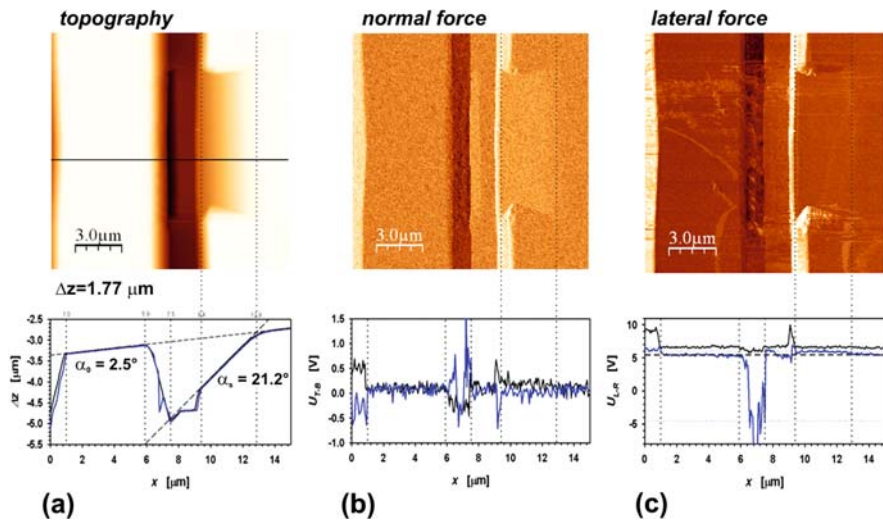


Fig. 9.4. LFM measurement on a modified TGF11 calibration grid. The grid is used for the wedge method. The plots in the lower row show cross-sections of the images of (a) topography, (b) normal force, and (c) lateral force, respectively. The position of the cross-section is indicated in the topography image. The images were measured upon scanning in forward direction, i.e. from *left to right*. The corresponding cross-sections are given in black colour, whereas the profiles given in blue colour resulted from the scan in backward direction, i.e. from *right to left*. A CL of spring constant $k_z \sim 2.3\ \text{N m}^{-1}$ was used (type FMR). Scan length $15.0\ \mu\text{m}$, maximum height difference $\Delta z = 1.77\ \mu\text{m}$, scan rate $0.1\ \text{lines/s}$. AFM system type CPII by Veeco Instruments

of FIB and is not related to a particular plane of the Si lattice. The surface properties of the micromachined slope are likely to be affected by the FIB milling process, for instance by implantation of Ga^+ ions. This can also affect the LFM contrasts. The results of an LFM measurement on such a 20° slope are given in Fig. 9.4c. In the normal and lateral force images, the slanted region corresponds to a plateau, that is, more or less constant forces are measured. Clearly, on the plateaus the force values are different from the ones on the flat areas. Since the lateral extent of the slanted slope increases with the depth of milling, wider plateaus can be observed in the force images. Hence for a given scan rate, the height change per second is smaller and the topography feedback is more likely to keep the normal force within a narrow band around the set value. That is, the tip-sample contact is in a stable condition. Moreover, a larger number of data points are recorded, thus providing a larger dataset.

The adhesion force, F_A , was measured by recording FDCs on the flat area of a ridge. Five different points were selected, at least five curves per point were recorded, and the pull-off forces were read out. The normal stiffness of the CL used for this measurement was $k_z \sim 2.3 \text{ N/m}$. A slope angle of 21° was measured. For the calculations of the conversion factor, α , the difference angle between the ‘flat’ and the sloped regions was taken into account. The values of this difference angle were $\Delta\theta = 18.5^\circ$ and 18.7° , respectively. The measured adhesion force was $F_A \sim 1.32 \times 10^2 \text{ nN}$ and for the conversion factor, α , a value of approximately $0.14 \mu\text{N/V}$ was calculated.

It should be noted that a similar approach has been demonstrated by Tocha et al. [26]. Using FIB, they milled wide notches with tilt angles between 20° and 35° into a Si(100) wafer surface. By fabricating wide slopes, they ensured compatibility with large tips, such as colloidal probes. Furthermore, they extended the equations describing the wedge method to the general case of two non-zero slopes. In this way, they can account for the likely case of a non-zero slope of the surface that is supposed to be flat. The uncertainty analysis performed by Tocha et al. [26] delivered a range of 5–15% for the overall conversion factor. The uncertainty was found to depend strongly on the instrument stability, and the lower end of the uncertainty range was corresponding to the condition of a well-equilibrated AFM system.

As mentioned before, the wedge method can be considered a direct technique, since it delivers an overall lateral calibration constant and does not require separate measurements for the determination of the torsional CL spring constant and lateral PSD sensitivity. On the contrary, however, Asay et al. [34] classified the wedge method as indirect since it involves a friction law, and Li et al. [36] classified it as a semi-direct method, considering the facts that it is a single-step technique but that some geometric variables need to be evaluated separately. On the bottom side, as a result of its single-step nature, the wedge method does not yield a separate value of the torsional spring constant. The overall torsional conversion factor value resulting from the wedge method also depends on the particulars of the PSD system and its settings. Furthermore, it requires knowledge of the normal CL spring constant.

In this sense, it is not genuinely a one-step technique. Varenberg's version of the wedge method relies on the measurement of FDCs and their evaluation in terms of the pull-off force [25]. Similarly, the initial approach of Ogletree et al. [24] requires normal force calibration since it relies on the analysis of friction forces for a range of loads.

The effective adhesion force, F_A , acting under the conditions of sliding friction is not necessarily identical with the pull-off force extracted from a normal FDC. This issue was pointed out by Butt et al. [71]. They identified the approach of determining F_A via FDCs [25] as a major source of uncertainty. Indeed, Varenberg et al. [25] reported a load-dependence of both the calibration coefficient, α_ϕ , and the friction coefficient, μ . On the contrary to the approach taken by Varenberg et al. [25], the original wedge method of Ogletree et al. [24] relied on measuring friction loops for a range of loads and evaluating their load dependence.

To account for the load dependence of the wedge method in the formulation by Varenberg et al. [25], modified data evaluation schemes have been developed. Wang and Zhao [72] attributed the load-dependence to the finite contact stiffness and expressed the overall conversion factor as the product of a factor given by the CL properties and a contact factor. The latter is given by the ratio of the CL stiffness to the stiffness of the tip-substrate contact.

Similarly to Wang and Zhao, Butt et al. [71] employed a mathematical separation to address the load-dependence of the wedge method. With the aim to separate the influence of F_A on the lateral deflection, they developed an ansatz that is linear in F_L and folds the F_A -dependence into the offset. The offset constants are denoted as I_W and I_Δ . Hence the key equations given by Varenberg et al. [25] can be written in the following form:

$$W_0^0(M, \theta) = S_W F_L + I_W \quad (9.30a)$$

and

$$\Delta_0^0(M, \theta) = S_\Delta F_L + I_\Delta, \quad (9.30b)$$

where

$$S_W = \frac{R(1 + \cos \theta) + t/2}{\beta} \cdot \frac{\mu}{\cos^2 \theta - \mu^2 \sin^2 \theta} \quad (9.30c)$$

$$I_W = \frac{R(1 + \cos \theta) + t/2}{\beta} \cdot \frac{\mu \cos \theta}{\cos^2 \theta - \mu^2 \sin^2 \theta} F_A \quad (9.30d)$$

$$S_\Delta = \frac{R(1 + \cos \theta) + t/2}{\beta} \cdot \frac{\mu^2 \sin \theta \cos \theta + \sin \theta \cos \theta}{\cos^2 \theta - \mu^2 \sin^2 \theta} - \frac{R \sin \theta}{\beta} \quad (9.30e)$$

$$I_\Delta = \frac{R(1 + \cos \theta) + t/2}{\beta} \cdot \frac{\mu^2 \sin \theta}{\cos^2 \theta - \mu^2 \sin^2 \theta} F_A. \quad (9.30f)$$

The values of the slope S_W result from linear plots of W_0^0 values measured for various loads, F_L . Then the friction coefficient, μ , and the overall torsional

conversion factor, $\alpha_\phi = \beta/H$, can be calculated using the equations (9.30c) and (9.30e). Notably, both these equations do not depend on F_A .

In addition to the effective adhesion force, F_A , Butt et al. [71] identified several error sources related to the CL deflection and its detection. The effects to be considered are:

1. A non-zero lateral CL deflection even if no torque is applied
2. A non-constant normal CL deflection due to poor feedback performance or due to pronounced topography features
3. Changes in the background deflection signal with CL height due to optical interference effects.

Accordingly, the torsional deflection signal can be written as a sum of contributions, including contributions from the torsional moment T , from the coupling to the static component of the normal signal, from the coupling to the dynamic component of the normal signal, and from the noise floor associated with optical interference effects. In particular, coupling to the dynamic component of the normal signal can be due to a limited response of the topography feedback resulting in changes in the lateral deflection signal. Hence it is likely to scale with the scan speed. In a similar manner, the various contributions to the friction loop offset, Δ_0 , and to the friction loop half-width, W_0 , can be summed up. Only the terms Δ_0^0 and W_0^0 of order zero are connected to the torsion moment, T . Since the first and the third terms are independent of the scan direction, they do not contribute to W_0 but to Δ_0 . However, their contribution to Δ_0 can be evaluated by scanning on a flat segment of the calibration grid and subtracting the related offset from the offset measured on a sloped facet.

Another point of concern is the effect of optical or mechanical crosstalk on the wedge method. The resulting values of the calibration factors may vary significantly with the degree of crosstalk, and parameters such as the friction coefficient could show a non-linear effect on the signals. Reportedly, the friction loops measured on two flat surfaces of different height can show a mutual offset, reflecting a coupling between the topography and the lateral force signal [68]. At present, however, no dedicated experimental techniques are available for the quantification of crosstalk effects. A valuable approach towards such quantification could be a rigorous modelling of the LFM detection system. Michal et al. [69] suggested parametric studies based on a system model.

Furthermore, it should be noted that the employment of the friction loop width, W , for the calibration implies that the method is not suitable for tip-sample systems that exhibit directional effects [36]. Although it appears negligible for many systems, a dependence of the absolute value of the friction force on the scan direction cannot be generally ruled out. In the case of a directional effect, an asymmetric friction loop would be measured even for a perfectly aligned PSD and a fully symmetric CL.

9.4.2 Methods Involving the Normal Spring Constant

When scanning parallel to the CL long axis rather than perpendicular to it, the frictional forces acting between tip and sample result in flexural CL deformations. In the ideal case of a perfectly flat and smooth surface, the flexural deformation can be exclusively attributed to the frictional forces. If the normal spring constant, k_z , and the related PSD sensitivity, S_z , are known, the frictional force can be quantified. Hence, by first measuring the frictional forces in a parallel scan and then recording the lateral force signal in the usual perpendicular scan, the lateral force measurement can be calibrated [41, 42]. This approach relies on the assumption that the frictional forces measured in parallel scan mode are equal to the ones measured in perpendicular scan mode.

In the general case of a non-flat surface, the flexural CL deformation includes a contribution from surface height variations, and this contribution needs to be separated from the frictional one. The separation can be achieved by accounting for the fact that the sign of the frictional forces changes with the scan direction, whereas the height variations due to the surface topography are independent of it. With the topography feedback being engaged, the changes in total flexural deformation with scan direction will translate into a difference in the height profiles measured in forward and in backward scan direction. That is, when scanning parallel to the CL long axis, the height profiles of the forward and backward scans will differ due to the frictional forces. As shown by Ruan and Bhushan [42], the frictional force can be calculated from the height difference between the two scans, the normal CL spring constant, the CL length, and the tip height. This method is frequently referred to as parallel scan method.

An underlying assumption is that the hysteresis in the topography signal can be fully attributed to the frictional contribution. However, piezoelectric scanners typically show a non-zero hysteresis due to non-linear behaviour. Although most AFMs feature a built-in linearisation, it is not always fully efficient. Since the parameters characterising the non-linearity are likely to change over time, in general numerical linearisation schemes (software linearisation) are less effective than active ones that involve a real-time measurement of the actual displacements generated by the piezoelectric scanner (hardware linearisation).

Another route for lateral force calibration is provided by (9.20). In addition to the Young's modulus, E , the shear modulus, G , the effective CL length, L , and the tip height, H , its employment requires knowledge of the normal CL spring constant, k_z . Hence, in an indirect way, this approach takes benefit from the methods available for the calibration of the normal spring constant. However, it should be borne in mind that large uncertainties in k_z will propagate into the uncertainty of k_x (Sect. 9.2.3).

Provided validity of Amontons' friction law, a linear relationship between frictional and normal forces can be expressed in terms of the friction coefficient, μ . Its experimental determination involves variation of the normal force and measurement of the lateral force signal. For instance, in the approach

demonstrated by Bilas et al. [53], normal force variation along a scan line was generated by exploiting the bow of a tube scanner. In their SBP system, the laser beam was not moved along with the CL. In combination with the action of the topography feedback, the corresponding relative motion between the laser spot and the CL resulted in normal force variations.

9.5 Methods Relying on a Force Balance upon Contact with a Rigid Structure

9.5.1 Normal Loading upon Contact with a Sloped Substrate

A torsional deformation of the CL can be generated in various ways. In addition to the typical case of a frictional force resulting from sliding motion, a lateral force can also be applied to the tip via the contact with a sloped surface, that is, without application of a scanning motion. Asay et al. [34] utilised such a static force balance to determine the lateral force at any applied load. Two major advantages of such non-scanning calibration techniques are that no frictional law is involved (the validity of which may depend on the particulars of the tip-sample interaction), and that the risk of tip degradation by wear is minimised.

The characteristic quantities can be extracted from the changes in lateral force upon approach of a ridged calibration grid featuring facets of the slope angles 0 and $+\theta$.

Torsional deformation results from any force acting perpendicular to a moment arm. As can be seen from Fig. 9.5, for an applied load, F_L , and a slope angle, θ , the friction force, F_f , is

$$F_f = F_L \sin \theta. \quad (9.31)$$

The horizontal force, F_x , is given by the sum of the horizontal components of the normal and the friction force:

$$F_x = (F_N - F_A) \sin \theta = F_f \cos \theta \quad (9.32)$$

Using this expression for F_x , S_x can be calculated from the defining equation (D4):

$$S_x = \frac{\Delta V_\phi}{\Delta x_\phi} = k_x \frac{\Delta V_\phi}{F_x} = k_x \frac{\Delta V_\phi}{(F_N - F_A) \sin \theta - F_L \sin \theta \cos \theta} \quad (9.33)$$

This equation holds for the ideal case of a perfectly symmetric CL. Importantly, any torsional deformation of the CL involves a rotation of the tip axis and a corresponding slip in the tip-substrate contact. Thus, a no-slip condition leads to erroneous results. In a more recent publication, Asay et al. clarified this point and provided a revised analysis.

A positional offset of the AFM tip causes an asymmetry with respect to the curve measured on the flat area (facet 2), that is, in a plot of the

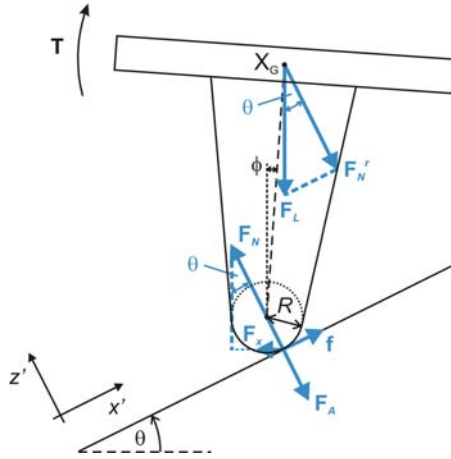


Fig. 9.5. Vector analysis of the forces acting on the AFM tip when loading a positive slope. The point X_G denotes the intersection with the CL long axis

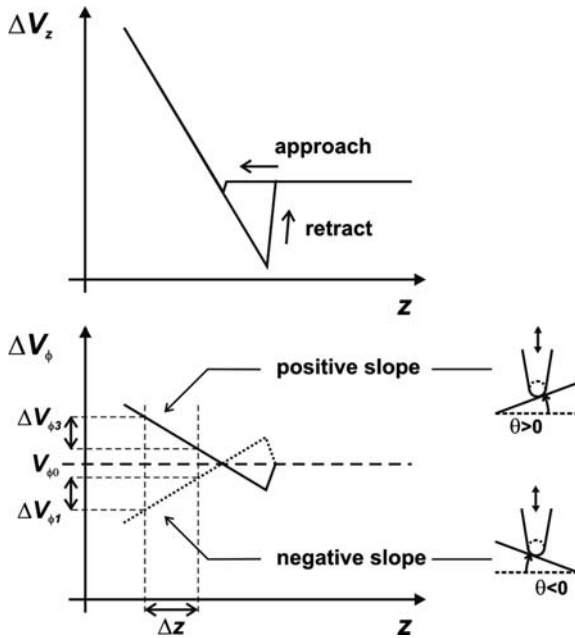


Fig. 9.6. Schematic signal-distance curves for the case of a FDC measured on a sloped surface. The changes in the normal signal, ΔV_z , with the vertical displacement are shown in the upper plot. The corresponding changes in the lateral force signal, ΔV_ϕ , are shown in the lower plot. In the regime of the FDC contact branch, ΔV_ϕ scales more or less linearly with ΔV_z , and its sign is given by the sign of the slope angle, θ

voltage ΔV_ϕ versus the z -position the voltage differences are not identical, $|\Delta V_{\phi 1}| \neq |\Delta V_{\phi 3}|$ (Fig. 9.6). Since the resulting lateral signal shifts in the same manner for positive and for negative slopes, the optical crosstalk can be described in terms of an offset voltage, $V_{\phi 0}$, in the lateral PSD signal. Hence, the voltage representing the degree of the optical crosstalk, $V_{\phi 0}$, is given by the arithmetic mean of $|\Delta V_{\phi 1}|$ and $|\Delta V_{\phi 3}|$ [34].

Furthermore, by using a commercially available calibration grid featuring trenches with sloped edges and by following the procedure given by Varenberg et al. [25], Asay et al. [34] compared their results from their non-scanning technique with the results from the wedge method (Sect. 9.4.1). Reportedly, they found reasonable agreement both for the lateral PSD sensitivity, S_ϕ , and for the friction coefficient, μ , measured on the horizontal surface.

Potential experimental complications are a slip-like movement of the tip while recording an FDC and a change in contact point between the tip and facet surface. A stick-slip transition should be recognisable from a step-like non-linearity in the FDC. Since nanoscale slipping would imply slight variations in contact point and moment arm, it is advisable to check the FDCs for the presence of stick-slip effects.

9.5.2 Normal Loading with the Contact Point off the CL Long Axis

A force applied at a lever arm distance, a , from the CL symmetry axis causes a CL with a horizontal lever attached to it to deflect flexurally as well as torsionally (Fig. 9.7a). A CL with a significant offset between its centre axis and the probe position may lend itself to a lateral force calibration experiment [32]; however, a general calibration method should be applicable to any CL.

Provided a CL similar to the target CL (i.e. the CL to be used for quantitative LFM measurements), it can be facilitated for lateral force calibration

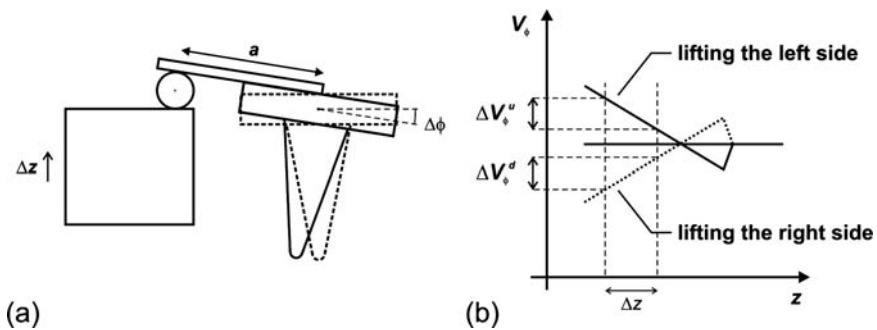


Fig. 9.7. (a) Schematic of the calibration method involving a horizontal bar attached to the CL beam. Using a piezoelectric transducer, it is lifted at a distance, a , from the CL long axis and a torque is applied to the CL. The distance, a , defines the torsional arm length. (b) Plot of the lateral signal, V_ϕ , versus the vertical displacement, z . Lifting the right side of the CL leads to voltage changes, ΔV_ϕ , of the opposite sign

by attaching a horizontal lever to the upper side of the CL beam. Hence a normal force can be applied to the free end of the horizontal CL, resulting in a torque [22, 28]. A schematic representation of the approach is shown in Fig. 9.7. The schematic (Fig. 9.7a) depicts the situation of a lever arm of length $a > w/2$, where w denotes the width of the CL.

Using the equations (D7) and (D8), the applied force ΔF_z can be written in terms of the related voltage change of the PSD output, ΔV_z :

$$F_z = \frac{k_z}{S_z} \Delta V_z, \quad (9.34)$$

where S_z is the normal force sensitivity. The torque resulting from the force F_z applied at a distance a from the CL long axis is given by

$$T = F_z a. \quad (9.35)$$

In combination with (D3) and (D9), this equation delivers an expression for the ratio of the two voltages readout from the PSD:

$$\frac{\Delta V_\phi}{\Delta V_z} = \frac{\alpha_z}{\alpha_\phi} a. \quad (9.36)$$

Similar to the approach of Feiler et al. [22], Reitsma [28] applied a lever technique by using a CL with a crossbar attached to it, that is, a bar across both the left and right edge of the CL. This prototype is referred to as ‘hammerhead’ CL [28]. The Si crossbeam was glued to the free end and perpendicular to the long axis of the CL.

A torque can be applied via a tungsten sphere glued to a ramp chip attached to a piezoelectric actuator. When lifting the ramp chip, the sphere presses against the left or the right end of the crossbeam. The distances, a_L and a_R , of the contact points between sphere and hammerhead from the CL central axis were approximately $73 \mu\text{m}$ [28]. When plotting their data according to (9.36), clearly Reitsma [28] observed a linear relationship, with the curve passing through the origin. The latter finding indicates a proper alignment between the CL and the detection system. Furthermore, the distance of the sphere-lever contact from the CL long axis was identified as a main source of uncertainty [28].

Reitsma [28] found that the precision was increased by a factor of 5 as compared to the single-point method of Feiler et al. [22]. It is worthwhile mentioning that micro-fabricated CLs of a shape similar to the ‘hammerhead’ CL used by Reitsma [28] are now commercially available [73]. Although these CLs may have been developed for different purposes [74], it is conceivable that in the future similar CLs could be dedicated to lateral force calibration.

A torsional deformation of the CL can also be generated by an asymmetric mass distribution. Toikka et al. [27] attached a fibre to the CL, with its axis perpendicular to the CL long axis. On the contrary to a crossbar used by Feiler et al. [22] and Reitsma [28], the fibre was mounted asymmetrically to cause

an imbalance in the gravitational pull. The resulting CL twist was quantified by approaching the assembly to a flat substrate and measuring the distance required to level the fibre.

In general, a torque can be generated by any loading point off the CL long axis. On the contrary to the case $a > w/2$ where an external loading point is generated by the attachment of a crossbar or by a special CL design, a loading point with $a < w/2$ is provided by a tip or probe at a position off the CL long axis. Using a colloidal particle as a probe, such a configuration has been employed by Quintanilla and Goddard [32]. As the main sources of uncertainty, they identified the particle offset distance and the alignment of the CL with the plane given by the laser beam path. Although such a positional offset of the probe can be beneficial for the purpose of calibration, the coupling between normal and lateral forces impairs the orthogonality of the related signals necessary for a facile interpretation of LFM measurements (Sect. 9.5.5).

9.5.3 Lateral Loading of a Horizontal Surface

When bringing the AFM tip in contact with a stiff sample surface and loading the contact in lateral direction, the contact region is sheared and the CL shows a torsional deformation. In the sticking regime, that is, before the onset of sliding motion, these two deformations add up to the lateral displacement applied to the tip-sample contact. This condition of static friction corresponds to the initial slope of a friction loop. In the limit of a very stiff tip-sample contact, the lateral displacement is completely converted into a torsional CL deformation. Hence such a measurement can be exploited to determine the lateral deflection sensitivity [33].

Considering that in the case of sharp tips also a significant tip deformation can occur [75], it is recommendable to make such measurements using a large colloidal probe and a stiff substrate. That is, a configuration should be chosen that ensures that the torsional CL deformation dominates. Since the lateral stiffness of the tip-sample contact increases with the tip radius, R , the condition of a stiff tip-sample contact is met with such a probe as well, that is, the total lateral deformation is dominated by the torsional deformation of the CL. In addition to the calibration of friction forces, the technique allows for the extraction of both the lateral contact stiffness, k_{xc} , and the contact shear strength, τ , from the friction trace.

For a quantitative analysis of the initial slope of the friction loop, the deflections of two springs need to be considered. With the springs arranged in series, the one represents the lateral CL spring constant, k_x , and the other represents the lateral contact stiffness, k_{xc} [76]. As indicated in Fig. 9.8, the lateral scanner displacement, Δx , can be decomposed into the lateral deformation of the contact, Δx_c , and the lateral movement at the interface leading to twisting of the CL, Δx_ϕ :

$$\Delta x = \Delta x_\phi + \Delta x_c. \quad (9.37)$$

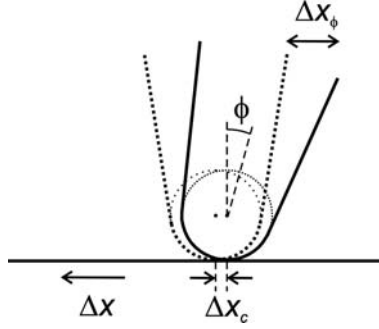


Fig. 9.8. Schematic representation of the compliance method. When operating in the sticking regime, a lateral displacement of the sample, Δx , causes a lateral deformation of the tip-substrate contact, Δx_c , and a lateral deformation of the CL, Δx_ϕ

The related forces are given by:

$$F_x = k_{x,\text{tot}}\Delta x = k_x\Delta x_\phi = k_{x_c}\Delta x_c. \quad (9.38)$$

The PSD signal, ΔV_ϕ , resulting from a twist angle $\Delta\phi$ of the CL is given by (D1). Accordingly, the slope, σ , of the stick region of the friction loop is given by:

$$\sigma = \frac{\Delta V_\phi}{\Delta x} = S_\phi \frac{\Delta\phi}{\Delta x}. \quad (9.39)$$

Combined with (9.4), this relationship can be written as follows:

$$H\sigma = S_\phi \frac{\Delta x_\phi}{\Delta x}. \quad (9.40)$$

Using (9.38), the ratio $\Delta x_\phi/\Delta x$ can be expressed in terms of the related spring constants:

$$H\sigma = S_\phi \frac{k_{x,\text{tot}}}{k_x}. \quad (9.41)$$

Provided a constant sensitivity S_ϕ and a constant lateral CL spring constant, k_x , $H\sigma$ varies with the total stiffness, $k_{x,\text{tot}} = (1/k_x + 1/k_{x_c})^{-1}$. As pointed out by Cain et al. [33], in the limit $k_x \ll k_{x_c}$ is $k_{x,\text{tot}} \approx k_x$ and the left side of (9.41) is equal to the scaling factor S_ϕ .

Using the equations (9.10), (9.37), and (9.38), (9.41) can be rewritten as follows:

$$k_\phi = H^2 k_{x,\text{tot}} \left(1 + \frac{\Delta x_c}{\Delta x_\phi} \right). \quad (9.42)$$

With increasing sphere size, the ratio $\Delta x_c/\Delta x_\phi$ is $\ll 1$ and $H^2 k_{x,\text{tot}}$ approaches k_ϕ .

The experimental results given by Cain et al. [33], however, departed from the asymptotic behaviour of $H\sigma$ expected for the limit of high loads. In fact, the product $H\sigma$ was found to decrease. Since it is unlikely that the lateral CL spring constant or the lateral contact stiffness decreases with increasing load, other effects need to be considered, such as the finite stiffness of the adhesive used to fix the colloidal probe. For larger values of the sphere diameter, $H^2k_{x,\text{tot}}$ was found to approach k_ϕ more rapidly with increasing load [33]. The torsional spring constant, k_ϕ , was calculated using FEA [33].

The approach of analysing the static friction regime for the purpose of lateral force calibration was referred to as static friction method [30,33]. Alternatively, it can be called lateral compliance method, reflecting the fact that it relies on an analysis of the deformations resulting from a lateral displacement.

9.5.4 Lateral Loading of a Vertical Surface

Using (9.10) and (9.18), the lateral spring constant, k'_x , can be calculated for a rectangular CL with the tip at a distance $L < L_0$ from the CL base:

$$k'_x = \frac{k_\phi L_0}{H^2 L}, \quad (9.43)$$

where L_0 is the full length of the CL. Once k'_x is known, the lateral conversion factor, α_x , can be calculated:

$$\alpha_x = \frac{F_x}{\Delta V_\phi} = k'_x \frac{\Delta x_\phi}{\Delta V_\phi} = k'_x \frac{1}{S_x}. \quad (9.44)$$

Equation (9.44) results from the defining relationships given in Table 9.1, namely the equations (D4), (D5), and (D6).

If the target probe is also the test probe, the lateral in-plane bending (Sect. 9.2.4) can be ignored. The uncertainty analysis performed by Cannara et al. [30] shows that ignoring the in-plane bending (Sect. 9.2.4) can lead to a significant increase in the error margin. Further major sources of uncertainty are the tip height, H_{target} , of the target CL, and the CL width, w . Another assumption is that the spatial distribution of reflected laser intensity on the PSD is very similar for both CLs. Comparison of the test probe technique with the wedge method delivered similar results, but a slightly lower uncertainty in the case of the wedge method where the calculated uncertainty in the lateral conversion factor, α_x , was 4.2% [30].

A major methodical advantage of the test probe technique is that the tip of the target CL is not in a mechanical contact [30]. In particular, the approach of using a test probe can be valuable for functionalised or coated tips that could take damage in the course of the calibration procedure.

9.5.5 Mechanical Crosstalk

Inspection of many commercially available AFM CLs reveals an offset in the tip position with respect to the CL symmetry axis. Such an offset results in

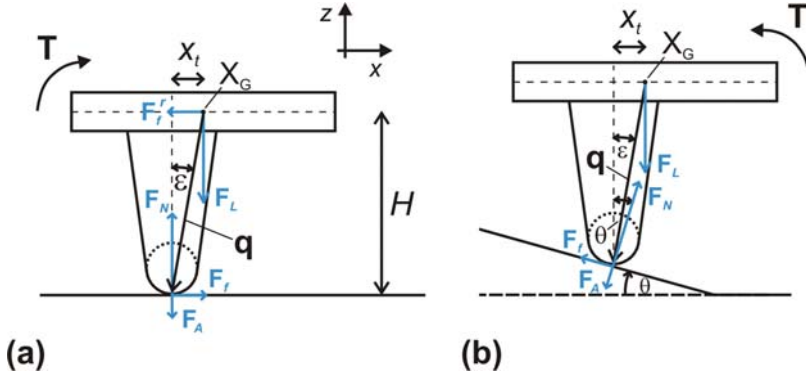


Fig. 9.9. Consideration of the case of an offset, x_t , in the tip position. (a) Introduction of the moment arm, \mathbf{q} , related to the tip. The point X_G denotes the intersection with the CL long axis. (b) Vector analysis of the forces acting when scanning across a slope

a torque acting on the CL, the amount of which depends on the normal force (Fig. 9.9). This effect can be employed for lateral force calibration; however, it should be borne in mind that any LFM measurement can be affected by such a crosstalk between the normal and the lateral forces. Notably, this purely mechanical crosstalk effect can occur on top of the optical crosstalk effect resulting from a misalignment of the PSD (Sect. 9.3).

Considering the Effect of an Offset in the Tip Position

On the contrary to a PSD misalignment (Sect. 9.3.2), an off-centred tip makes the lateral force responses on sloped surfaces non-symmetric about the horizontal segment. As shown in Fig. 9.9a, a positional offset with respect to the CL long axis can be accounted for by considering the angle, ε , of the moment arm of length q . The moment arm is defined by the vector, \mathbf{q} , between the tip end and the CL centre axis, and its angle, ε , with respect to the line perpendicular to the CL beam. Forces perpendicular to the moment arm cause CL twisting and the corresponding torque, \mathbf{T} , is given by:

$$\mathbf{T} = \mathbf{q} \times \mathbf{F} = \begin{pmatrix} -H \tan \varepsilon \\ 0 \\ -H \end{pmatrix} \times \begin{pmatrix} -F_f \cos \theta + (F_N - F_A) \sin \theta \\ 0 \\ F_f \sin \theta + (F_N - F_A) \cos \theta \end{pmatrix}. \quad (9.45)$$

The force causing the CL to twist can be calculated from the projection of the lateral force F_{lat} onto the axis orthogonal to the moment arm \mathbf{q} (Fig. 9.9b). Here, this force is denoted as F_{lat}^{oq} :

$$F_{lat}^{oq} = F_f \cdot \cos(|\theta| - \varepsilon) - (F_N - F_A) \cdot \sin(|\theta| - \varepsilon). \quad (9.46)$$

It should be noted that F_{lat}^{oq} depends on the distance angle $|\theta| - \varepsilon$.

Considering the Effect of an Offset in the Position of the Shear Centre

In general, any deviation of the real CL geometry from the ideal one as well as any heterogeneity in its materials' properties can lead to an offset between the nominal CL symmetry axis and the CL shear centre. The shear centre can be defined as the point at which any resulting concentrated force will not induce a torsional twist on the CL [36]. As a consequence, a comprehensive characterisation of the CL torsional properties includes a crosstalk constant, that is, a measure for the change in lateral force upon a change in the normal load applied to the tip.

As depicted in Fig. 9.10, such an offset between the geometrical centre, X_G , of the local cross-section of the CL on the projection plane and the CL shear centre, X_C , can be accounted for by the misfit angle, γ_m (Fig. 9.10). It contributes to the initial tilt angle, ϕ_0 , in the undeformed configuration, that is, $\phi_0 = \theta - \gamma_m$, where θ denotes the slope angle of the surface feature.

The coordinate systems involved are the friction force system, (x', z') , associated with the tip-sample contact, the mechanical sensing system, (x, z) , associated with the undeformed CL, and the optical sensing system, (n, l) , associated with the PSD (Fig. 9.10). The relationships between the forces and the deformation quantities can be derived by describing the forces acting in the tip-substrate contact in the (x', z') -system and the CL deformations in the (x, z) -system. Application of the orthogonal transformation matrix yields an expression for the friction force, $f^{(\kappa)}$, in terms of the normal and the lateral output voltages [36]:

$$f^{(\kappa)} = \text{sgn}(\kappa) \left\{ \alpha_{xx} V_\phi^{(\kappa)} + \alpha_{xz} V_z^{(\kappa)} \right\}. \quad (9.47)$$

Herewith, α_z denotes the normal conversion factor, α_{xx} denotes the 2nd rank lateral conversion factor, and α_{xz} denotes the crosstalk conversion factor. The prefactor $\text{sgn}(\kappa)$ is equal to +1 or -1 for scans in the forward or in the reverse direction, respectively. The factors α_{xx} and α_{xz} of (9.47) are related to intrinsic structural quantities of the CL, namely the misfit angle, γ_m , and the reduced parameter $\xi = R/H$:

$$\alpha_{xx} = \alpha_x \frac{1}{(1 - \xi) \cos \gamma_m + \xi \cos \theta} \quad (9.48)$$

$$\alpha_{xz} = \alpha_z \frac{(1 - \xi) \sin \gamma_m + \xi \sin \theta}{(1 - \xi) \cos \gamma_m + \xi \cos \theta}. \quad (9.49)$$

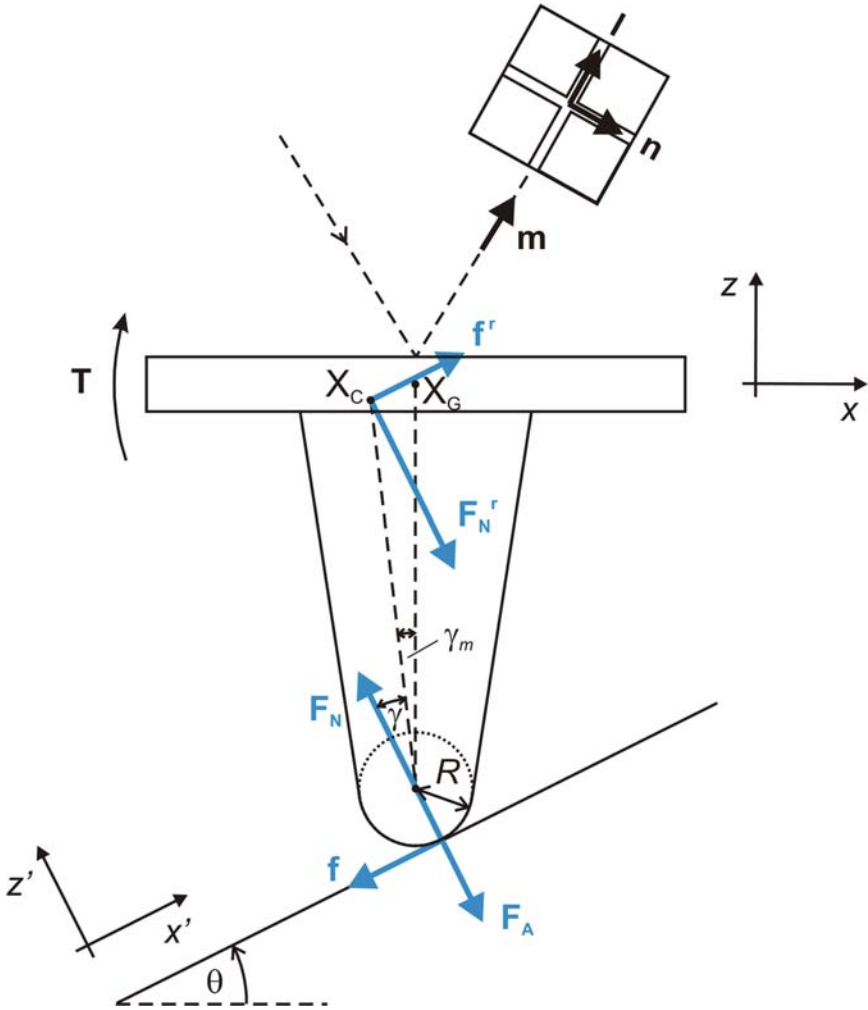


Fig. 9.10. Similar to Fig. 9.5, however, for the case of an offset between the symmetry centre and the shear centre of the CL. Considering the plane of the tip cross-section, the two centres are represented by the points X_G and X_C , respectively. The offset is accounted for via an offset angle, γ_m

In the case of $\theta = \gamma_m = 0$, the crosstalk lateral conversion factor is zero, $\alpha_{xz} = 0$, and the 2nd rank lateral conversion factor is equal to α_x , that is, $\alpha_{xx} = \alpha_x$. That is, no mechanical crosstalk between the normal and the lateral force occurs and α_{xx} reduces to the lateral conversion factor of (D6), respectively.

Eliminating the Mechanical Crosstalk Effect by Novel Design Concepts

Considering the susceptibility of conventional CLs for mechanical crosstalk, novel concepts of micromechanical probes have been suggested. Essentially, such studies aim at a complete decoupling of the deformations resulting from normal and lateral forces.

Fukuzawa et al. [77] devised such a probe by combining a double CL with a torsion beam. The torsion beam is horizontal and supports the double CL beam. The amount of its torsional deformation is a measure for the normal force acting on the tip. The lateral force, however, is measured via the lateral bending of the double CL beam. Owing to its stiffness against torsional deformation, it is deflected parallel to a lateral force. Yet a standard PSD setup can be used, that is, a laser beam deflection system. By providing a groove-like structure etched into the probe surface, the lateral movement of the groove-like structure causes spatial variations in the laser beam spot bounced off the CL, thus resulting in variations in the lateral force signal.

It is interesting to note that on the contrary to schemes aiming at the elimination of the mechanical crosstalk effect, for special measurement modes also, its augmentation can be of interest. Recently, a new CL design has been devised where the tip is deliberately offset from the CL long axis [74]. When operating in intermittent contact mode, the resulting torsional oscillations enrich the CL frequency spectrum and help reconstruct the tip-sample interactions. As compared to the higher harmonics of the fundamental CL resonance, the torsional resonance peaks were found to show a good signal-to-noise ratio (SNR).

9.6 Methods Relying on a Force Balance Upon Contact with a Compliant Structure

9.6.1 The Case of a Vertical Reference Beam

As discussed in Sect. 9.3.1, the lateral PSD sensitivity can be determined by pressing a colloidal probe against a vertical and rigid wall. If this wall is replaced with a compliant beam of known stiffness and vertical alignment, also the torsional CL spring constant can be measured. This approach was demonstrated by Ecke et al. [31].

As depicted in Fig. 9.11a, a beam-shaped reference lever was pushed against the equator of the probe. Such an experiment yields a lateral FDC (Fig. 9.11b), and the lateral sensitivity, S_x^{ref} , of the setup is given by the slope of the compliant region of the FDC. Repeating this experiment, using a rigid edge rather than a reference lever yields a lateral force curve with a steeper compliant region. The respective sensitivity can be denoted as S_x^0 .

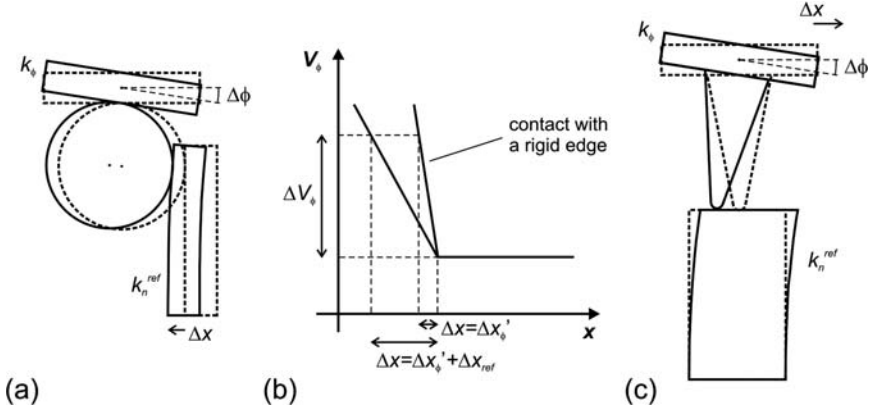


Fig. 9.11. Schematic representation of the vertical beam method. (a) Sketch of the configuration of colloidal probe and vertical beam of stiffness, k_n^{ref} . (b) The related lateral signal-displacement curves both for the case of a compliant beam and of a rigid structure. (c) Sketch of a similar approach, applying the lateral force via the contact between the tip apex and a beam, i.e. operating in the sticking regime of a lateral contact. If the reference beam is made of a piezoresistive material, its deformation is converted into a voltage

Similar to the calibration technique for normal CL spring constant [41,78], the lateral CL spring constant can be calculated from the condition of force balance between the two springs:

$$F_x = k'_x \Delta x'_\phi = k_n^{\text{ref}} \Delta x^{\text{ref}}. \quad (9.50)$$

The lateral displacements $\Delta x'_\phi$ and Δx^{ref} sum up to the total lateral displacement generated by the scanner:

$$\Delta x = \Delta x'_\phi + \Delta x_{\text{ref}}. \quad (9.51)$$

Considering (6.10), the lateral CL spring constant, k'_x , for the case of loading at the position of the probe equator is related to the lateral CL spring constant, k_x , for the standard case of loading at the lowest point of the probe:

$$\frac{k_x}{k'_x} = \frac{H'^2}{H^2}, \quad (9.52)$$

where $H' = D/2 + t/2$ with the diameter, D , of the colloidal probe.

Taking into account that in the case of a rigid counter-body the total displacement is accommodated solely by the CL deformation, the lateral sensitivity, S_x^{ref} , of the measurement involving both CLs can be written in the following form:

$$S_x^{\text{ref}} = \frac{\Delta V_\phi}{\Delta x} = \left(\frac{\Delta x'_\phi + \Delta x_{\text{ref}}}{\Delta V_\phi} \right)^{-1} = \left(\frac{1}{S_x^0} + \frac{\Delta x_{\text{ref}}}{\Delta V_\phi} \right)^{-1}, \quad (9.53)$$

where S_x^0 is the lateral sensitivity in the case of a rigid vertical edge. Combining this with the equation representing the force equilibrium (9.50), the ratio k'_x/k_n^{ref} can be expressed in terms of the sensitivities:

$$\frac{k'_x}{k_n^{\text{ref}}} = \frac{S_x^0 - S_x^{\text{ref}}}{S_x^{\text{ref}}}. \quad (9.54)$$

By accounting for the shorter torsional arm length in the case of the calibration measurement, (9.52) leads to an expression for the lateral CL spring constant, k_x :

$$k_x = k_n^{\text{ref}} \frac{S_x^0}{S_x^{\text{ref}} - S_x^0} \left(\frac{D}{2} + \frac{t}{2} \right)^2. \quad (9.55)$$

The lateral sensitivity, S_x , results from $S_x/S_x^0 = (S_\phi/H) / (S_\phi^0/H') = H'/H$:

$$S_x = S_x^0 \frac{D}{D + \frac{t}{2}}. \quad (9.56)$$

By considering a CL of rectangular shape and the related beam mechanics, the lateral force can be calculated using the equations (9.7), (D2), (9.10), and (9.18):

$$F_x = \frac{T}{H} = \frac{k_\phi \Delta\phi}{H} = \frac{H^2 k_x \Delta\phi}{H} = H \frac{Gwt^3}{3LH^2} \Delta\phi = \frac{1}{3} \frac{Gwt^3}{LH} \Delta\phi. \quad (9.57)$$

Using this expression as well as the equations (D6), (D4), and (9.4), the lateral conversion factor, α_x , can be expressed in terms of the lateral PSD sensitivity, S_x , and the lateral CL spring constant, k_x :

$$\alpha_x = \frac{F_x}{\Delta V_\phi} = \frac{F_x}{S_x \Delta x_\phi} = \frac{1}{S_x} \frac{F_x}{H \Delta\phi} = \frac{1}{S_x} \frac{Gwt^3}{3LH^2} = \frac{k_x}{S_x}. \quad (9.58)$$

This equation is consistent with the corresponding expression for the torsional conversion factor (9.3b).

In a simplified approach, the PSD can be assumed to be rotationally symmetric, that is, to be equally sensitive for normal and torsional deflections. Then, S_x can be replaced with S_z . However, as mentioned in Sect. 9.3, the PSD response is likely to be asymmetric, for example, due to asymmetries in the laser spot shape [35].

Xie et al. [35] employed a commercially available piezoresistive CL as a portable microforce calibration standard. As a reference lever with an integrated piezoresistive sensor, this calibration standard allows for a voltage output proportional to its deflection.

The normal spring constant, k_n^{ref} , of the reference lever was calibrated using Cleveland's mass loading method [79]. The resulting value of k_n^{ref} was

given as $18.21 \pm 0.47 \text{ N/m}$ [35]. For the calibration of the piezoresistive force sensor sensitivity [35], the piezoresistive CL was mounted horizontally on a movable platform. It was loaded normal to its longitudinal axis via a glass microsphere attached to a nanopositioning stage. The voltage output, V_p , of the Wheatstone bridge connected to the piezoresistive CL was recorded as a function of the microsphere displacement. The resulting value of the piezoresistive force sensor sensitivity, $C_p = dF_z/dV_p$, was $10.36 \pm 0.27 \mu\text{N/V}$ [35]. The effect of torsional loading on the output voltage, V_p , was found to be negligible. Thus, it was concluded that points on the side edges as well as on the top edge of the piezoresistive CL can be used as loading locations [35].

For the calibration of AFM CLs, the piezoresistive force sensor was either mounted vertically or horizontally on the AFM stage along its longitudinal axis. Then the tip of the CL to be calibrated was brought into contact with the top edge of the piezoresistive force sensor. The respective measurement configurations are referred to as top-loading (Fig. 9.11c) and side-loading (*not shown*), respectively. In both configurations, the AFM tip was moved laterally after the loading location had been determined by scanning in contact mode. Upon lateral movement, the voltage outputs both of the piezoresistive sensor, ΔV_p , and of the lateral PSD signal, ΔV_ϕ , were recorded.

- For the top-loading configuration, force equilibrium between the lateral force exerted by the CL, F_x , as given by (D5) and the restoring force exerted by the reference beam, $k_n^{\text{ref}} \Delta x^{\text{ref}}$, yields the following equation:

$$k_x \Delta x_\phi = F_x = k_n^{\text{ref}} \Delta x^{\text{ref}} = C_p V_p. \quad (9.59)$$

As resulting from (9.37) and (9.38), the total lateral stiffness, $k_{x,\text{tot}}$, of the mechanical system consisting of CL and tip-sensor contact is given by

$$k_{x,\text{tot}} = \left(\frac{1}{k_x} + \frac{1}{k_{xc}} \right)^{-1}. \quad (9.60)$$

Herewith, k_x and k_{xc} are the lateral stiffness of the AFM CL and the tip-substrate contact, respectively.

Using (9.59), the lateral conversion factor, α_x , can be expressed in terms of the voltage output, V_p , of the piezoresistive force sensor:

$$\alpha_x = \frac{F_x}{\Delta V_\phi} = C_p \frac{\Delta V_p}{\Delta V_\phi}. \quad (9.61)$$

- For the side-loading configuration, however, the loading direction is perpendicular to the reference beam, at a certain distance from its clamping end. In order to eliminate the effects of finite lateral contact stiffness, Xie et al. [35] applied the lateral force via the tip shank rather than via the tip apex. Furthermore, from the condition that the same moment and voltage are generated for both configurations, a correction formula for the conversion factor measured in the side-loading configuration was deduced.

Xie et al. [35] gave an overall error of 12.4% in the lateral conversion factor, α . Notably, it was found to depend largely on the uncertainty in C_p . For the case of a piezoresistive force sensor calibrated using an absolute force standard, they expect an error of <6%. The uncertainty of their method was compared with the one of a beam mechanics method where the thickness was determined via the CL fundamental resonance frequency. With the uncertainties $\delta L_{\text{eff}} = 2\%$, $\delta w = 2\%$, $\delta h = 10\%$, $\delta t = 10\%$, and $\delta S_z = 10\%$, they arrived at an overall error of 33% in α_0 .

9.6.2 The Case of a Horizontal Reference Beam

A horizontal fibre of known dimensions and modulus can be utilised to apply a lateral force to the tip. If laid down on a flat substrate and attached to it at one of its two ends, the free end of the fibre bends laterally when it encounters the scanning AFM tip.

For small deflections of the horizontal fibre, $\Delta x_{\text{hf}} \ll L_{\text{hf}}$, the restoring force exerted by the in-plane fibre, F_{hf} , is proportional to the displacement at the position of the tip, Δx_{hf} :

$$F_{\text{hf}} = \frac{3\pi}{4} \frac{E_{\text{hf}} r_{\text{hf}}^4}{L_{\text{hf}}^3} \Delta x_{\text{hf}}, \quad (9.62)$$

with the fibre bending stiffness

$$k_{\text{hf}} = \frac{3\pi}{4} \frac{E_{\text{hf}} r_{\text{hf}}^4}{L_{\text{hf}}^3}. \quad (9.63)$$

Herewith, E_{hf} , r_{hf} , and L_{hf} are the Young's modulus, the radius, and the free length of the fibre, respectively [59]. The fibre bending stiffness, k_{hf} , scales with $1/L_{\text{hf}}^3$. When plotting the PSD output against the restoring force, F_{hf} , the lateral force conversion factor results from the slope of the curve.

In the experimental realisation of Liu et al. [59], a glass fibre was used. The loading was achieved by scanning across the fibre. The loading was assumed quasi-static, and the lateral force conversion factor was derived from the balance of the restoring forces.

The conversion factor value was found to increase slightly with the free length of the glass fibre, L_{hf} . Notably, for smaller values of L_{hf} , the assumption of small deflection ($\Delta x_{\text{hf}} \ll L_{\text{hf}}$) may not hold. Furthermore, for increasing deflections, Δx_{hf} , the contact point between tip and glass fibre is likely to move a small distance along the fibre axis, thus resulting in an effective increase in L_{hf} , that is, in a slightly reduced spring constant. That is, for a larger free length, L_{hf} , a more accurate lateral force conversion factor should be obtained. With uncertainties of 3% in both r_{hf} and L_{hf} , and of 5% in Δx_{hf} , Liu et al. [59] calculated the overall uncertainty of the technique to be 26%. However, the effective free length of the fibre may differ from the value of L_{hf} extracted from the images, if the glue used to attach the fibre to the substrate is not rigid but accommodates a fraction of the deformation. In these terms, effectively

the value of L_{hf} is likely to be slightly larger, resulting in a lower value of the restoring force, F_{hf} .

Another point of concern is the offset between the tip apex and the contact point with the fibre. Even if the lower side of the fibre is in genuine contact with the substrate, a fibre of micron-size radius touches the tip shank rather than the very end of the tip. Hence, the torsional arm length differs from the one acting in the case of an apex–surface contact. If the torsional arm length is shorter than the full tip height, H , then the same lateral force generates a smaller torque, and the CL appears stiffer for torsional deformations. This effect is not unlikely to make a major contribution to the total uncertainty budget, in addition to the uncertainties considered by Liu et al. [59]. Provided accurate knowledge of the glass fibre diameter as well as the tip shape, however, it should be possible to account for this effect.

Furthermore, sliding motion of the fibre across the substrate surface may involve a frictional force contribution to the total lateral force acting on the tip. The assumed force balance neglects such a frictional contribution.

Owing to its cylindrical geometry, a perfect fibre is isotropic for all radial directions. Thus the same bending stiffness applies to both lateral and normal loading conditions. Once the bending stiffness of the fibre is known, it can be utilised to calibrate the CL for normal as well as for lateral loading. Such a measurement configuration can be achieved with a fibre that is clamped at one end and that is not suspended by a substrate, that is, if the other end of the fibre is free to move in all radial directions. By loading the free end of the fibre at various distances from the clamped end, Morel et al. [80] determined the CL stiffness and the PSD sensitivity. Using a calibrated weight, the fibre bending stiffness was measured by loading the fibre at various distances from the clamped end.

9.6.3 The Case of a Mechanically Suspended Platform

Provided a calibrated reference spring and a rigid tip-sample contact, the lateral CL spring constant can be calculated without much effort. Force equilibrium between the spring representing the lateral CL spring constant, k_x , and the reference spring of stiffness, k_s , yields the following equation:

$$k_x = k_s \frac{\Delta x_s}{\Delta x_\phi}. \quad (9.64)$$

Herewith, Δx_s is the lateral displacement of a platform suspended by the reference spring. The restoring force exerted by this spring is denoted as F_s . A total displacement Δx generated by an actuator causes a lateral displacement Δx_s of the platform, depending on the stiffness of its suspension arms, k_s .

With the aim of providing a user-friendly calibration standard, purpose-made MEMS devices were prototyped and demonstrated [37,38]. Core element of the demonstrated MEMS-devices is a platform suspended by four parallel

arms. Similar to an AFM CL, the stiffness of the suspension system can be adjusted via the dimensions of the arms. The suspension arms of the platform should have a large stiffness for the deformation in normal direction but a much lower stiffness for the deformation in the direction perpendicular to their length axis. When scanning across the platform surface, the AFM-tip applies a lateral force and displaces the platform laterally. The demonstrated suspension systems [37, 38] consist of two pairs of parallel arms. The suspension system can be represented by a reference spring of stiffness k_s , and the restoring force generated by the suspension system is denoted by F_s .

Similar to the lateral electrical nanobalance (LEN) developed by Cumpson et al. [37], Ando et al. [38] demonstrated a calibration device based on MEMS technology, dubbed microlateral force sensor (MLFS). On the contrary to the LEN device, the MLFS employs a feedback circuit to keep the lateral position of the platform constant. Upon application of a voltage to the comb drive, the suspension system deforms elastically and the platform moves along the x -axis. The electrodes of a detector attached to the platform are used to measure a tunnelling current. Depending on the magnitude of the tunnelling current, the voltage applied to the comb drive is adjusted to keep the tunnelling current constant.

When the AFM tip is in mechanical contact with the suspended platform, the latter is laterally displaced by a distance Δx_s , as a function of the lateral force exerted by the CL. The position of the platform is restored by means of a feedback circuit. It applies a voltage to the comb drive that scales with the tunnelling current measured. Approximately, the applied voltage, ΔV_d , increases linearly with the displacement Δx_s :

$$\Delta V_d = g_{cd} \Delta x_s. \quad (9.65)$$

Herewith, the gain factor of the feedback connected with the comb drive is denoted as g_{cd} .

The lateral force applied by the tip, F_x , can be measured by recording the changes in the driving voltage, V_d . The driving force, F_d , generated when applying the voltage V_d to the comb drive is given by the dimensions of the interdigitating fingers and the permittivity of the medium between the comb electrodes:

$$F_d = 2\varepsilon_0 \frac{N_{cd} T_{cd}}{D_{cd}} V_d^2, \quad (9.66)$$

where ε_0 is the permittivity of air (8.85×10^{-12} F/m), D_{cd} is the width of the gap between the comb fingers, T_{cd} is the thickness of the comb fingers, and N_{cd} is the number of fingers constituting the comb drive.

By considering the dimensions of the suspension system, Ando et al. [38] employed FEA to calculate its lateral stiffness, k_s . The resulting value of k_s was 26 ± 3 N/m, with the main error source given by the error in the thickness of the leaf springs. The relationship between the driving voltage, V_d , and the

displacement of the sensing table, Δx_s , was measured using a confocal laser-scanning microscope. The displacement Δx_s was found to scale with $V_d^{1.8}$ and a sensitivity of 1.1 nN/mV was calculated for a bias voltage of $V_d = 90$ V.

However, it should be noted that deviations of the fabricated structures from the layout are likely to occur. Typically, they result from variations in the etch steps involved in the fabrication process of such MEMS structures. Such variations can cause a large uncertainty in the suspension stiffness. Frequently occurring etch effects are edge biases and non-orthogonal sidewall angles, and they can result in significant changes in the mechanical properties, such as the fundamental resonance frequency [52]. With the aim of providing a valid calibration as well as traceability to SI units, Cumpson et al. [37] devised a procedure for the calibration of their suspension system, involving electrical current measurement, laser Doppler vibrometry, and white light interferometry. Owing to such a set of complimentary measurements, this approach does not depend on knowledge of the elastic properties or dimensions of the suspension arms.

When scanning on the surface of the suspended platform, a lateral force acts, depending on the surface properties of both tip and platform. In the case of a low friction coefficient, μ_{tp} , of the tip-platform contact, the signals resulting from the torsional CL deformation and the platform displacement can be as small as the noise level. Hence the calibration measurement will suffer from a large uncertainty. To ensure a sufficient SNR, it would be necessary to select a useful LEN device depending on the spring constant of the CL as well as on the largely unknown friction coefficient μ_{tp} . Hence to make the selection solely dependent on the CL spring constant, it is helpful to provide a suitable topographical structure leading to significant changes in the measured signals.

Using FIB, Ando et al. [38] created an array of asperities on the platform surface. When scanning across such asperities, variations in the lateral forces occur. The array pitch size was 250 nm. The normal spring constant, k_z , of the CL used for the demonstration of the technique was approximately 2 N/m. It should be noted that the AFM measurement was performed under vacuum conditions (3×10^{-5} Pa) to ensure a stable tunnelling current.

Rather than providing an array of topographic structures, a single but more extended structure can also be utilised. In their design of the LEN device, Cumpson et al. [37] included a slit-shaped groove, situated in the centre of the platform. When scanning across such a groove, the lateral force acting between tip and platform increases, resulting both in a significant change in the CL torsional deformation and in a lateral displacement of the platform. However in the case of such slit-shaped grooves, the global tip shape comes into play when the tip touches the groove edges with its shank rather than with its apex, thus leading to an ill-defined tip-sample contact. To ensure a stable and continuous contact between tip apex and platform, more shallow grooves seem to be advisable. A trench-shaped groove could be a more viable approach, similar to the elementary unit of the grating structure used in Varenberg's version of the wedge method [25, 26].

9.6.4 The Case of a Magnetically Suspended Platform

By employing a compliant load cell, Li et al. [36] demonstrated a direct method for the calibration of the lateral conversion factor of an AFM CL-tip system. In a similar manner to the MEMS devices, it relies on a suspended platform; however, the suspension is achieved via diamagnetic forces rather than via suspension arms. When providing an inhomogeneous magnetic field exhibiting a minimum, a diamagnetic object is levitated, forced to the point of lowest magnetic field, and gets trapped at the position of the field minimum. The gradient of the diamagnetic levitation force represents a spring that can be used as the reference spring, k_s , of the calibration measurement. Accordingly, the suspension system was dubbed diamagnetic lateral force calibrator (D-LFC) [36].

Li et al. [36] relied on a static friction measurement, that is, the lateral force is kept below the force required for the stick–slip transition (Sect. 9.5.3). The total lateral displacement, Δx , generated by the scanner results in lateral displacements of the suspended platform, Δx_s , and of the tip apex, Δx_t , respectively:

$$\Delta x = \Delta x_s + \Delta x_t. \quad (9.67)$$

In turn, the tip displacement, Δx_t , can be split into a fraction due to the compliance of the tip-sample contact, Δx_c , and a fraction due to the torsional deformation of the CL, Δx_ϕ (Sect. 9.5.3). Owing to the low lateral stiffness, k_s , of the magnetic suspension system, however, the component Δx_t is negligible, $\Delta x_t \ll \Delta x_s$, and the lateral force acting on the tip can be written as

$$f' = k_s \Delta x_s \approx k_s \Delta x. \quad (9.68)$$

When writing (9.47) in a difference form, $\Delta f = \alpha_{xx} \Delta V_\phi + \alpha_{xz} \Delta V_z$, the inverse slope, $\partial f / \partial V_x$, of the resulting plot of V_ϕ versus $f \approx k_s \Delta x_s$ represents the lateral conversion factor, α_{xx} . The measurement can be made for various values of V_z by changing the setpoint value of the force feedback. Similarly to the determination of α_{xx} , the crosstalk lateral conversion factor, α_{xz} , is given by the inverse slope, $\alpha_{xz} = \partial f / \partial V_z$. Alternatively to a two-step procedure, the lateral conversion factors can be determined in a single step by a 3D fit procedure for the full dataset $f(V_x, V_z)$.

It is worthwhile underscoring that apart from the normal conversion factor, α_z , and the lateral conversion factor, α_{xx} , the method provides a conversion factor, α_{xz} , related to crosstalk. As elucidated in the Sects. 9.3.2 and 9.5.5, respectively, a crosstalk between normal and lateral forces can be induced optically or mechanically.

Interestingly, the availability of both lateral conversion factors allows for the measurement of the absolute value of the friction force. Hence it enables studies of directional effects. Notably, this is not the case for methods that rely on the measurement of the friction loop width, W (Sect. 4.1).

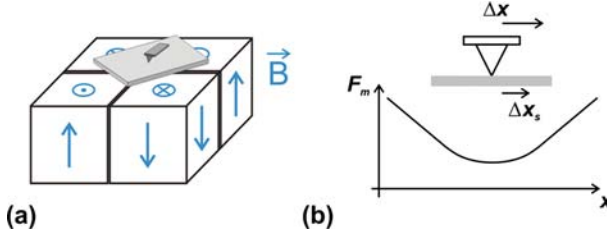


Fig. 9.12. Schematic representations of (a) the array of four permanent magnets used to create a magnetic spring, and (b) a PG sheet levitated and trapped in the region of high magnetic field gradients. When operating in the sticking regime of the tip-substrate contact, a small lateral displacement, Δx , of the CL causes a lateral displacement of the levitated PG sheet, Δx_s , as well as a lateral deformation, Δx_t , of the system consisting of CL, tip and tip-substrate contact. If the lateral stiffness of the levitation system is very small, the component Δx_t is negligible.

The levitation system of Li et al. [36] consists of a set of permanent magnets and the levitated (suspended) platform (Fig. 9.12a). They utilised the highly anisotropic magnetic behaviour of pyrolytic graphite (PG). PG has a layered structure and is highly diamagnetic. Its magnetic susceptibility in the direction perpendicular to the basal plane is several times higher than that in the direction parallel to the plane [36, 81]. Provided a vertical magnetic field, a PG sheet can be suspended with its basal plane aligned horizontally. Lateral stability can be achieved by providing a lateral modulation of the magnetic field, for example, by arranging a number of permanent magnets with alternating magnetisation. A system of four permanent magnets allows for stability along both lateral directions, with the minimum of the magnetic field at its centre (Fig. 9.12b). For the normal spring constant of their system, Li et al. [36] obtained a value of approximately 10 pN/nm. The lateral stiffness, however, is much smaller.

Key parameters for the diamagnetic levitation effect are the magnetic flux density, B , and its gradient, ∇B , the susceptibility of the diamagnetic material, χ , as well as its mass density, ρ . The magnetic force, F_m , acting on a particle of volume, V , and susceptibility, χ , is given by:

$$F_m = \frac{1}{2\mu_0} \chi V \nabla B^2. \quad (9.69)$$

For diamagnetic materials, the susceptibility is negative, that is, $\chi < 0$. Balance between the diamagnetic and the gravitational force yields the following condition [82]:

$$B \nabla_z B = \mu_0 g \frac{\rho}{\chi}, \quad (9.70)$$

where g denotes the gravitational acceleration. That is, materials with a large ρ/χ -ratio require a large value of the product $B \nabla_z B$.

9.7 Methods Relying on Torsional Resonances of the CL

In addition to the common measurement of lateral forces in LFM (i.e. the lateral forces induced by the scanning motion of the tip across the sample surface), more recently developed mapping techniques employ the torsional resonance of the CL. In the torsional resonance mode, a special configuration of piezoelectric actuators is used to drive the CL into torsional resonance [60, 83]. The related amplitude and phase signals are monitored and their changes upon tip-sample interaction were shown to enable distinct materials contrasts [83].

In another mode relying on the CL torsional resonance, a hammerhead-shaped CL is operated in intermittent contact mode [74]. The tip of the asymmetric CL is placed off the CL long axis by a distance larger than the CL half-width. As a result of this asymmetry, a torsional moment is generated (Sect. 9.5.5), and vibrations at the fundamental torsional resonance can be excited. While the amplitude of the CL flexural deformation is used for operation of the topography feedback, the amplitude and phase signals related to the torsional resonance provide additional information on the sample properties, such as the elastic modulus [74, 84].

Due to the development of such techniques relying on the CL torsional resonance, it seems to be of growing interest. However, it also can be exploited for the calibration of the CL torsional spring constant. In many ways, a direct analogy exists with the use of the fundamental flexural resonance for the calibration of normal forces. Commonly used methods for the calibration of the normal CL spring constant are the added mass method of Cleveland et al. [79] and the unloaded resonance technique of Sader et al. [85]. The former relies on the change in resonance frequency with known masses attached to the free end of the CL. The latter relies on the resonance frequency and quality factor of the fundamental flexural resonance as well as on the plan view dimensions of the CL. In a similar manner, both these methods can be applied to determine the torsional CL spring constant [39].

The fundamental torsional resonance frequency, ω_ϕ , depends on the torsional spring constant, k_ϕ , and on the mass moment of inertia, J :

$$\omega_\phi^2 = \frac{k_\phi}{J}. \quad (9.71)$$

In addition to the mass moment of inertia, J_i , inherent to the CL, the moment J also comprises the contribution, J_a , due to a mass attached to the CL:

$$J = J_i + J_a. \quad (9.72)$$

For a sphere-shaped mass, M_s , attached to the free end of the CL, the total added mass moment of inertia, J_a , is given by [39]:

$$J_a = \frac{7}{5} M_s r_s^2. \quad (9.73)$$

Herewith, M_s and r_s denote the mass and the radius of the attached sphere. Rearrangement of (9.71) shows that a plot of J_a versus ω_ϕ^{-2} results in a linear curve, the slope of which is given by the torsional spring konstant, k_ϕ . That is, the calibration procedure requires a series of torsional resonance frequency measurements, using different spherical masses attached to the free end of the CL.

Notably, the mass should be placed at the free end of the CL and aligned with the CL long axis. Offsets from the free end of the CL and from its long axis need to be accounted for:

- Off-end loading can be covered by employing (9.10) and (9.18). That is, the torsional spring constant for the case of loading at a distance $\Delta L = L_0 - L$ from the CL free end is given by:

$$\frac{k_\phi}{k_{\phi 0}} = \frac{L_0}{L}, \quad (9.74)$$

where L_0 is the full CL length and $k_{\phi 0}$ the related torsional spring constant.

- Off-axis loading can be accounted for via the mass moment of inertia, J_a , due to the added mass. With the distance, e , between the sphere centre of mass and the CL long axis (excentricity), J_a can be expressed as follows [39]:

$$J_a = \left(\frac{7}{5} + \frac{e^2}{r_s^2} \right) M_s r_s^2. \quad (9.75)$$

Hence, the mass moment of inertia of a mass applied off-axis is larger than that of an identical mass placed on the long axis of the CL. Consideration of (9.71) and (9.72) shows that the increase in J_a will cause ω_ϕ to decrease.

Moreover, off-axis positioning of the mass can also result in a coupling of the flexural and torsional modes of vibration. A similar asymmetry in mass distribution occurs for the case of a positional offset of the tip from the CL long axis (Sect. 9.5.5). By imaging CLs upon vibration in their resonance modes, it has been shown that asymmetries in the CL geometry lead to significant mode coupling [86].

Various techniques for the deliberate excitation of torsional resonances can be employed:

- In the torsional resonance mode, a pair of piezoelectric actuators is used [60, 83]. They are integrated into the CL holder in a way that the left- and the right-hand side of the CL can be displaced separately, in a direction normal to the CL plane. When they are operated 180° out of phase, a torsional motion is generated.
- Since the usage of piezoelectric actuators involves a potential hazard of electrical shorting when operating in liquids, Mullin and Hobbs [87] devised a different setup. They replaced the piezoelectric actuators with slabs of a magnetostrictive material. In combination with a modulated magnetic

field, one slab expands while the other one contracts. The modulated magnetic field can be generated by superimposing the static magnetic field of a permanent magnet with the alternating magnetic field of a solenoid connected to an ac voltage source [87].

- Another way of exciting torsional resonances has been demonstrated by Jeon et al. [40]. Using a triangular-shaped CL, they induced a torque by exploiting the Lorentz force. Via an electrically conductive coating on the backside of the CL, an electric current was conducted across the two legs of the CL. In combination with an in-plane static magnetic field aligned perpendicular to the CL long axis, a Lorentz force is induced that acts upwards on one side of the CL and downwards on the other side of the CL. The magnitude of the resulting torque can be calculated from the magnetic field strength, the electric current, and the area enclosed by the two legs of the CL [40].

9.8 Discussion

Considering the large number of methods related to force calibration in LFM, it is advisable to organise the comparison of the methods in the form of a table. As can be seen from Table 9.3, the applied criteria are as follows:

- The resulting calibration constant, that is, if the method yields separate values of the PSD sensitivity and of the torsional spring constant or if it yields the overall conversion factor
- The suitability of the method for colloidal probes, that is, if it is applicable to probes incorporating micron-size beads rather than a standard AFM tip
- The risk of tip damage, that is, if the method requires a mechanical contact with the tip and if this contact is likely to cause some damage
- The degree of user-friendliness, that is, the level of skills required
- The methodological requirements or complications inherent to the method
- The uncertainties as reported by the authors and the identified main source of uncertainty.

Typically, measurement of an overall (combined) torsional calibration factor means that neither k_ϕ nor the sensitivity of the optical detection system were calibrated separately. On the contrary, if separate calibrations are achieved then the overall factor can be calculated but is not primarily measured.

For the evaluation of the risk of tip damage as well as of the degree of user friendliness, a scoring range of “L”-“M”-“H” (low-medium-high) was applied. For instance, a score of “L” indicates a low risk of tip damage, and it was attributed to the methods where the tip is not involved in the calibration procedure. A scoring value of “M” was applied if a static contact with the tip is involved and a scoring value of “H” if a scanning motion is involved.

Table 9.3. Overview of the demonstrated methods for the calibration of lateral forces in LFM

Method (number, name, reference)	Approach	Parameters calibrated	Suitable for colloidal probes	Risk of tip damage	User friendliness	Methodological requirements or complications	Reported uncertainties; main source of uncertainty	compared with method
E1	<i>Optical geometry</i> Liu et al. [62] Sect. 9.3.1	S_ϕ k_ϕ α_ϕ	✓	L	M	Requires knowledge of the optical beam path, incl. the PSD dimensions, distance from CL to PSD, and probe moment arm	$\delta k_\phi = 10\%$; non-linearities in the detector system and its associated mechanical positioning	
E2	<i>Optical geometry method for liquid cells</i> Tocha et al. [63] Sect. 9.3.1	✓	✓	L	M	Requires knowledge of the optical beam path and of the refractive index of the liquid	Unknown; non-linearities in the detector system and its associated mechanical positioning	optical geometry method E1

(Continued)

Table 9.3. (Continued)

Method (number, name, reference)	Approach	Parameters calibrated	Suitable for colloidal probes	Risk of tip damage	User friendliness	Methodological requirements or complications	Reported uncertain- ties; main source of uncertainty	compared with method
E3 <i>Pivot and mirror tilting</i> Bogdanovic et al. [29] Sect. 9.3.1	Pushing against a protuber- ance offset from the CL long axis, PSD sens. by changing the angle between AFM-head and a mirrored substrate	S_ϕ k_ϕ α_ϕ ✓ ✓	colloidal probes	L (low) M (med) H (high)	L (low) M (med) H (high)	Demonstrated for a tipless target CL only, requires well-defined tilting of the AFM head	$\delta k_\phi = 15\%$; off-axis distance a , different conditions of the mirrored substrate	dimensional
E4 <i>Pivot and torsional Sader</i> Petterson et al. [13] Sect. 9.3.1	Torsional Sader + pivot as of method E3, reverse calculation yields the PSD sens.	✓ ✓	✓	H	M	Requires meas. of the torsional resonance frequency	$\delta k_\phi =$ 16.8%; off-axis distance a	torsional Sader + lever method, also torsional added mass method E25, mirror tilting

E5	<p><i>Wedge method by Carpick (facets of a crystal surface)</i> Ogletree et al. [24] Sect. 9.4.1</p>	<p>Scanning across well-defined slopes of a crystal surface (SrTiO₃) and analysing the friction loops, the meas. need to be made for a range of loads</p>	✓	H	M	<p>The sharp edges between two facets can lead to tip damage, restriction to a CL with a sharp tip, assumes Amontons' law, considers the adhesion as a force offset</p>	$\delta \left(\frac{\alpha_x}{\alpha_z} \right) = 12\%$; piezo response	dimensional
E6	<p><i>Wedge method by Varenberg (micro-fabricated trenches)</i> Varenberg et al. [25] Sect. 9.4.1</p>	<p>Scanning across well-defined slopes of a grid (TGF11) and analysing the friction loops, requires meas. of the pull-off force but no variation in load</p>	✓	H	M	<p>Grid commercially available, scanning steep slopes is demanding for the topogr. feedback, calibr. factor was found to be load-dependent, n.a. to large colloidal probes</p>	$\delta\alpha_x = 11\%$ integrated probe $\delta\alpha_x = 8\%$ colloidal probe; adhesion force via pull-off meas, results in load-dependence of calibr. Factor	

(Continued)

Table 9.3. (Continued)

Method (number, name, reference)	Approach	Parameters calibrated	Suitable for colloidal probes	Risk of tip damage	User friendliness	Methodological requirements or complications	Reported uncertainties; main source of uncertainty	compared with method
E7	<p><i>Wedge method by Varenberg, using shallow slopes and the friction loops</i></p> <p>Tocha et al. [26]</p> <p>Sect. 9.4.1</p>	<p>S_ϕ</p> <p>k_ϕ</p> <p>α_ϕ</p>	✓	<p>H</p>	<p>M</p>	<p>Fabrication of the slopes using FIB, prone to optical interference on the sloped areas</p>	<p>$\delta\alpha_x = 5-15\%$;</p> <p>non-zero global sample tilt</p>	<p>calc. of the CL spring constant (analytical & via FEA) + opt. geometry method E1</p>
E8	<p><i>Wedge method by Varenberg, case of colloidal probes</i></p> <p>Ling et al. [16]</p> <p>Sect. 9.4.1</p>	✓	✓	<p>H</p>	<p>M</p>	<p>Grid commercially available, scanning steep slopes is demanding for the topogr. feedback, n.a. to large colloidal probes</p>	<p>$\delta\alpha_\phi = 4\%$;</p> <p>ill-defined contact (contact area, load) due to large slope angles</p>	

E9	<p><i>Parallel scanning</i> Ruan et al. [42] Sect. 9.4.2</p>	<p>Employing the calibrated normal forces by scanning in direction parallel to the CL long axis</p>	✓	✓	H	<p>Requires calibration of the normal force signal; assumes negligible hysteresis of the scanner; assumes a symmetric tip</p>	<p>Unknown; piezo response, tip asymmetry, normal spring constant</p>	<p>macro-scale friction meas.</p>
E10	<p><i>Piezotube bow</i> Bilas et al. [53] Sect. 9.4.2</p>	<p>Employing the calibrated normal forces by calc. of lateral forces via relationship between k_z and k_x</p>	✓	✓	H	<p>Generates variation in normal force along scan line via bow of a tube scanner</p>	<p>Unknown; piezo response</p>	
E11	<p><i>Normal FDCs on wedge sample</i> Asay et al. [34] Sect. 9.5.1</p>	<p>Vector analysis of forces acting when measuring FDCs on slopes</p>	✓	✓	M	<p>Grid commercially available</p>	<p>$\delta\alpha_x = 15.2\%$; slope dependence of adhesion force</p>	<p>wedge method (dev. $\sim 13\%$)</p>

(Continued)

Table 9.3. (Continued)

Method (number, name, reference)	Approach	Parameters calibrated	Suitable for colloidal probes	Risk of tip damage	User friendliness	Methodological requirements or complications	Reported uncertainties; main source of uncertainty	compared with method
E12 <i>Horizontal lever</i> Feiler et al. [22] Sect. 9.5.2	Generating a torque via a lever attached to the target CL	S_ϕ k_ϕ α_ϕ	✓	L M (med) H (high)	L (low) M (med) H (high)	Requires attachment of a stiff lever to the CL, assumes that the CL of interest as well as the PSD settings are equivalent to the ones used for calibration	$\delta\alpha_\phi = 10\%$; lever arm length	dimensional via analytical continuum mechanics method [51], tilting a mirrored substrate
E13 <i>Hammer-head CL</i> Reitsma [28] Sect. 9.5.2	Lever method by lifting the outer ends of cross-beam attached to the CL	✓	✓	L	L	Requires attachment of a cross-beam to the CL, the additional mass reduces its resonance frequency	$\delta\alpha_\phi = 0.8\%$; lever arm length	

E14	<i>Horizontal lever</i> Toikka et al. [27] Sect. 9.5.2	Generating a torque via the gravitational pull of a lever attached in an asymmetric manner to the CL	✓	✓	M	L	Requires attachment of a stiff lever to the CL, significant deflection for soft CLs only	Unknown; distance needed to level the lever	dimensional via analytical continuum mechanics method [51]
E15	<i>Off-axis colloidal probe</i> Quintanilla et al. [32] Sect. 9.5.2	Torque via a colloidal particle mounted off-axis + normal added mass method [79]	✓	✓	L	M	Requires a test probe of similar width and reflectivity, fitted with a colloidal particle	$\delta S_x = 11.9\%$ $\delta k_\phi = 16.6\%$; particle offset distance	lateral compliance method E16 [33], k_ϕ/k_z ratio
E16	<i>Lateral compliance method</i> Cain et al. [21, 33] Sect. 9.3.1 Sect. 9.5.3	Calc. of k_ϕ using FEA + static friction meas. and relying on high contact stiffness by using colloidal probes	✓	✓	M	M	Assumption of stiff probe-substrate contact requires colloidal probe	Unknown; (lateral contact stiffness)	ref. [21] provides comparison with the methods of E1, E5 and ref. [33]

(Continued)

Table 9.3. (Continued)

Method (number, name, reference)	Approach	Parameters calibrated	Suitable for colloidal probes	Risk of tip damage	User friendliness	Methodological requirements or complications	Reported uncertainties; main source of uncertainty	compared with method	
		S_ϕ	k_ϕ	α_ϕ					
E17 <i>Lateral FDC using a test probe</i> Cannara et al. [30] Sect. 9.3.1 Sect. 9.5.4	Torsional Sader + lateral FDC against vertical wall	✓	✓		L	M	Requires a test probe of similar width and reflectivity, assumes same laser intensity distribution for test CL and for target CL	$\delta S_x^{\text{target}} = 0.5\%$ $\delta k_x^{\text{target}} = 6.8\%$ $\delta \alpha_x^{\text{target}} = 7.5\%$; length of target CL 4.2% torsional moment arm	wedge method $\delta \alpha_x^{\text{wedge}} = 4.2\%$
E18 <i>Vertical lever</i> Ecke et al. [31] Sect. 9.6.1	Pushing a colloidal probe laterally against a rigid vertical surface and a vertical lever	✓	✓		M	M	Requires a colloidal probe and a test lever of known normal stiffness	Unknown; piezo response	dimensional, horizontal lever method E12

E19	<i>Vertical lever with integrated piezo-resistive sensor</i> Xie et al. [35] Sect. 9.6.1	Bending a reference CL via lateral loading of the contact between its edge and the tip, its bending is measured using an integrated piezores.	✓	✓	M	M	Requires a calibration, reference lever with integrated piezores. sensor and related electronics, contact between tip and the edge-sided face of the reference lever	$\delta\alpha_x = 12.4\%$; piezoresistive force sensor sensitivity, vertical offset between contact point and tip apex	modified dimensional (t via f_1), $\delta\alpha_x^{\text{dim}} = 33\%$
E20	<i>Bending a fibre clamped at its one end</i> Morel et al. [80] Sect. 9.6.2	Loading a glass fibre at various distances from its clamped end	✓	✓	✓	M	Contact point between tip and fibre ill-defined	$\delta k_x = 9\%$; $\delta S_x = 3\%$; tip-fibre contact point defining the torsional arm length	friction force meas. on various materials
E21	<i>Bending a fibre supported by a substrate</i> Liu et al. [59] Sect. 9.6.2	Scanning across the free end of a glass fibre clamped at the other end and supported by a substrate	✓	✓	H	M	Severe risk of damage to the tip, contact point between tip and fibre ill-defined, neglects frictional forces fibre-substrate	$\delta\alpha_\phi = 26\%$; tip-fibre contact point defining the torsional arm length	Dimensional $\delta\alpha_c^{\text{dim}} = 56\%$

(Continued)

Table 9.3. (Continued)

Method (number, name, reference)	Approach	Parameters calibrated	Suitable for colloidal probes	Risk of tip damage	User friendliness	Methodological requirements or complications	Reported uncertain- ties; main source of uncertainty	compared with method
		S_ϕ k_ϕ α_ϕ						
E22 <i>Lateral electrical nanobalance (LEN)</i> Cumpson et al. [37]	Scanning on a suspended platform of a MEMS device	✓	✓	H	M	Design, fabrication, and calibration of the MEMS devices are time-consuming	$\delta\alpha_x = 7\%$; stiffness of the suspension system	
E23 <i>Micro-lateral force sensor (MLFS)</i> Ando et al. [38]	Scanning on a suspended platform of a MEMS device with feedback via comb drive	✓	✓	H	M	Design, fabrication, and calibration of the MEMS devices are time-consuming	unknown; stiffness of the suspension system	plausibility test using an expression for tangential force from method E5
E24 <i>Diamagn. lateral force calibrator (D-LFC)</i> Li et al. [36]	Lateral loading (static friction) of a platform suspended by magnetic forces	✓	✓	M	M	The magnetic levitation system can be too large for many AFM systems, n.a to CLs with magn. coating	(16.5%) $\delta\alpha_{xx} = 0.5\%$ $\delta\alpha_{xz} = 2.4\%$; lateral displacement of PG sheet	k_z

E25	<i>Torsional added mass</i> Green et al. [39] Sect. 9.7	Attaching small masses to the CL and analysing the changes in its fundamental torsional resonance frequency	✓	✓	L	M	Requires accurate attachment of small spheres to the CL and the meas. of its fundamental torsional resonance frequency	torisional Sader, normal added mass method [79] + equ. for k_ϕ/k_z -ratio
E26	<i>Torsional Sader</i> Green et al. [39] Sect. 9.7	Exploiting the fundamental torsional resonance	✓	✓	L	H	Requires meas. of the torsional resonance peak and of the plan view dimensions of the CL	torisional added mass method E25, normal added mass method [79] + equ. for k_ϕ/k_z -ratio
E27	<i>Torsional deflection via Lorentz force</i> Jeon et al. [40] Sect. 9.7	Applying a static magnetic field and driving an electric current across the two legs of the CL	✓	✓	L	L	Requires a static magn. field and an electrically conductive pattern that allows the flow of el. current across the two legs of a triangular CL	Unknown; magn. field strength and area of the current loop if the abs. value of k_ϕ is calc., k_z if the rel. value is calc.

9.9 Concluding Remarks

From the synopsis of the major characteristics of the demonstrated lateral force calibration methods, the following conclusions can be drawn:

- Frequently, the wedge method seems to be used, also for the purpose of reference when a novel method is evaluated. Provided a suitable substrate including sloped features, the measurement can be achieved within reasonable time. Points of concern are large slope angles and the involvement of a friction law. The former point can be addressed by the fabrication of shallower slopes or can be partly compensated by a low scan speed; the latter point is of a more fundamental nature. The equations used for evaluation involve Amontons' law, although it is not generally applicable to nanoscale contacts. Furthermore, the frictional characteristics of the tip-substrate contact can change with the tip properties, such as its shape and micro-roughness.
- Methods avoiding the direct application of a lateral force via the tip apex, such as the lever methods, have the advantage of reducing the risk of tip damage; however, they need to deal with the added uncertainty in the effective length of the torsional arm. In the case of an LFM measurement, the lateral force is applied via the tip apex, and any calibration method deviating from this case relies on assumptions concerning the tip height and shape. Although the lever methods do not couple to the tip apex and suffer from an added uncertainty, their directness provides the opportunity to cross-check results from more indirect methods. Commercial availability of dedicated CLs with integrated crossbars could reduce the amount of experimental efforts significantly.
- Similarly to the diamagnetic levitation method, the MEMS-based devices employ a suspended platform. Essentially, the suspended platform approach allows for straight calculation of the CL torsional spring constant. Issues to be addressed are the calibration of the suspension system, the detection of the lateral platform displacement upon interaction with the AFM tip, and the topographic profile of the platform surface.
- The diamagnetic levitation method seems promising due to the comparatively low reported uncertainties and due to the capability for quantification of mechanical crosstalk. However, the requirement of a strong magnetic field is bound to reduce the compatibility of the method with commercial AFM heads. Yet, for labs with a focus on measurement science and technology it could be worthwhile to evaluate the method.
- Due to material heterogeneities, asymmetries, or a positional offset, the CL shear centre is likely to be different from the geometrical centre of the CL. Such an offset results in mechanical crosstalk. Although a non-zero crosstalk is likely to occur for most CLs, its quantification seems to be widely neglected. However, a thorough calibration method for lateral forces should provide a comprehensive quantification, including crosstalk

coefficients. Knowledge of these coefficients would help select cantilevers that are most appropriate for quantitative LFM measurements.

Acknowledgements

The author would like to thank C.A. Clifford, D. Roy, and I.S. Gilmore for their helpful comments. This work is supported by the National Measurement System of the UK Department for Business, Innovation & Skills through the Chemical and Biological Metrology Programme. This article is reproduced with the permission of Her Majesty's Stationery Office (Crown copyright, 2009).

References

1. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, *Phys. Rev. Lett.* **59** 1942 (1987)
2. S.S. Sheiko, *Adv. Polym. Sci.* **151** 61 (2000)
3. S.S. Perry, *MRS Bull.* **29** 478 (2004)
4. S.S. Perry, W.T. Tysoe, *Tribol. Lett.* **19** 151 (2005)
5. M. Munz, E. Schulz, H. Sturm, *Surf. Interface Anal.* **33** 100 (2002)
6. J. Raczowska, R. Montenegro, A. Budkowski, K. Landfester, A. Bernasik, J. Rysz, P. Czuba, *Langmuir* **23** 7235 (2007)
7. B. Bhushan, T. Kasai, G. Kulik, L. Barbieri, P. Hoffmann, *Ultramicroscopy* **105** 176 (2005)
8. G. Papastavrou, S. Akari, *Nanotechnology* **10** 453 (1999)
9. S. Sun, K.S.L. Chong, G.J. Leggett, *Nanotechnology* **16** 1798 (2005)
10. H. Liu, B. Bhushan, *Ultramicroscopy* **100** 391 (2004)
11. J.A. Williams, H.R. Le, *J. Phys. D Appl. Phys.* **39** R201 (2006)
12. C.J. Wright, I. Armstrong, *Surf. Interface Anal.* **38** 1419 (2006)
13. T. Pettersson, N. Nordgren, M.W. Rutland, A. Feiler, *Rev. Sci. Instrum.* **78** 093702 (2007)
14. S.-K. Kim, M.-H. Jung, H.-W. Kim, S.-G. Woo, H. Lee, *Nanotechnology* **2005** 16 (2227)
15. M. Munz, D.C. Cox, P.J. Cumpson, *Phys. Stat. Sol. A* **205** 1424 (2008)
16. X. Ling, H.-J. Butt, M. Kappl, *Langmuir* **23** 8392 (2007)
17. D.V. Vezenov, A. Noy, L.R. Rozsnyai, C.M. Lieber, *J. Am. Chem. Soc.* **1997** 119 (2006)
18. D.A. Smith, S.D. Connell, C. Robinson, J. Kirkham, *Anal. Chim. Acta* **479** 39 (2003)
19. H.J. Butt, B. Cappella, M. Kappl, *Surf. Sci. Rep.* **59** 1 (2005)
20. C.D. Frisbie, L.F. Rozsnyai, A. Noy, M.S. Wrighton, C.M. Lieber, *Science* **1994** 265 (2007)
21. R.G. Cain, M.G. Reitsma, S. Biggs, N.W. Page, *Rev. Sci. Instrum.* **72** 3304 (2001)
22. A. Feiler, P. Attard, I. Larson, *Rev. Sci. Instrum.* **71** 2746 (2000)
23. C.A. Clifford, M.P. Seah, *Appl. Surf. Sci.* **252** 1915 (2005)
24. D.F. Ogletree, R.W. Carpick, M. Salmeron, *Rev. Sci. Instrum.* **67** 3298 (1996)

25. M. Varenberg, I. Etsion, G. Halperin, *Rev. Sci. Instrum.* **74** 3362 (2003)
26. E. Tocha, H. Schonherr, G.J. Vancso, *Langmuir* **22** 2340 (2006)
27. G. Toikka, R.A. Hayes, J. Ralston, *J. Adhes. Sci. Technol.* **11** 1479 (1997)
28. M.G. Reitsma, *Rev. Sci. Instrum.* **78** 106102 (2007)
29. G. Bogdanovic, A. Meurk, M.W. Rutland, *Colloids Surf. B* **19** 397 (2000)
30. R.J. Cannara, M. Eglin, R.W. Carpick, *Rev. Sci. Instrum.* **77** 053701 (2006)
31. S. Ecke, R. Raiteri, E. Bonaccorso, C. Reiner, H.-J. Deiseroth, H.-J. Butt, *Rev. Sci. Instrum.* **72** 4164 (2001)
32. M.A.S. Quintanilla, D.T. Goddard, *Rev. Sci. Instrum.* **79** 023701 (2008)
33. R.G. Cain, S. Biggs, N.W. Page, *J. Colloid Interface Sci.* **227** 55 (2000)
34. D.B. Asay, S.H. Kim, *Rev. Sci. Instrum.* **77** 043903 (2006)
35. H. Xie, J. Vitard, S. Haliyo, S. Régnier, M. Boukallel, *Rev. Sci. Instrum.* **79** 033708 (2008)
36. Q. Li, K.-S. Kim, A. Rydberg, *Rev. Sci. Instrum.* **77** 065105 (2006)
37. P.J. Cumpson, J. Hedley, C.A. Clifford, *J. Vac. Sci. Technol. B* **23** 1992 (2005)
38. Y. Ando, N. Shiraiishi, *Rev. Sci. Instrum.* **78** 033701 (2007)
39. C.P. Green, H. Lioe, J.P. Cleveland, R. Proksch, P. Mulvaney, J.E. Sader, *Rev. Sci. Instrum.* **75** 1988 (2004)
40. S. Jeon, Y. Braiman, T. Thundat, *Appl. Phys. Lett.* **84** 1795 (2004)
41. B. Bhushan, *Handbook of Micro/Nanotribology*, 2nd edn. (CRC Press, Boca Raton, FL, 1999), Chap. 1
42. J.A. Ruan, B. Bhushan, *J. Tribol – Trans ASME* **116** 378 (1994)
43. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids* (Clarendon Press, Oxford 1964)
44. E. Meyer, R. Lüthi, L. Howald, M. Bammerlin, M. Guggisberg, H.J. Güntherodt, *J. Vac. Sci. Technol. B* **14** 1285 (1996)
45. C. Putman, R. Kaneko, *Thin Solid Films* **273** 317 (1996)
46. U.D. Schwarz, O. Zwörner, P. Köster, R. Wiesendanger, *Phys. Rev. B* **56** 6987 (1997)
47. M. Munz, B. Cappella, H. Sturm, M. Geuss, E. Schulz, *Adv. Polym. Sci.* **164** 87 (2003)
48. Draft of the standard “Surface chemical analysis – Vocabulary – Part 2: Terms for the scanned probe microscopies”, ISO TC 201/SC 1 – ISO/DIS 18115–2, version 29 Aug 2008
49. D. Choi, W. Hwang, E. Yoon, *J. Microsc.* **228** 190 (2007)
50. Bergmann-Schaefer *Lehrbuch der Experimentalphysik. Bd. 1: Mechanik, Akustik, Wärme*, 9th edn (Walter de Gruyter, Berlin, 1974), Chap. 5
51. J.M. Neumeister, W.A. Ducker, *Rev. Sci. Instrum.* **65** 2527 (1994)
52. R.K. Gupta, *J. Microelectromech. Syst.* **9** 380 (2000)
53. P. Bilas, L. Romana, B. Kraus, Y. Bercion, J.L. Mansot, *Rev. Sci. Instrum.* **75** 415 (2004)
54. J.L. Hazel, V.V. Tsukruk, *J. Tribol* **120** 814 (1998)
55. C.A. Clifford, M.P. Seah, *Nanotechnology* **16** 1666 (2005)
56. J.L. Hazel, V.V. Tsukruk, *Thin Solid Films* **339** 249 (1999)
57. D. Sarid *Scanning Force Microscopy – With Applications To Electric, Magnetic, And Atomic Forces* (Oxford University Press, New York, 1991), Chap. 1
58. D.-A. Mendels, M. Lowe, A. Cuenat, M.G. Cain, E. Vallejo, D. Ellis, F. Mendels, *J. Micromech. Microeng.* **16** 1720 (2006)
59. W. Liu, K. Bonin, M. Guthold, *Rev. Sci. Instrum.* **78** 063707 (2007)

60. Y. Song, B. Bhushan, *J. Phys. Condens. Matter* **20** 225012 (2008)
61. J.E. Sader, C.P. Green, *Rev. Sci. Instrum.* **75** 878 (2004)
62. E. Liu, B. Blanpain, J.P. Celis, *Wear* **192** 141 (1996)
63. E. Tocha, J. Song, H. Schonherr, G.J. Vancso, *Langmuir* **23** 7078 (2007)
64. A. Meurk, I. Larson, L. Bergstrom, *Mater. Res. Soc. Symp. Proc.* **522** 427 (1998)
65. M.W. Such, D.E. Kramer, M.C. Hersam, *Ultramicroscopy* **99** 189 (2004)
66. R. Piner, R.S. Ruoff, *Rev. Sci. Instrum.* **73** 3392 (2002)
67. P. Prunici, P. Hess, *Ultramicroscopy* **108** 642 (2008)
68. M. Varenberg, I. Etsion, G. Halperin, *Rev. Sci. Instrum.* **74** 3569 (2003)
69. G. Michal, C. Lu, A.K. Tieu, *Nanotechnology* **19** 455707 (2008)
70. B. Bhushan, J.A. Ruan, *J. Tribol Trans ASME* **116** 389 (1994)
71. Supporting Information of Ref. [16]
72. F. Wang, X. Zhao, *Rev. Sci. Instrum.* **78** 043701 (2007)
73. HMX-10 type cantilevers by Veeco Instruments Inc
74. O. Sahin, S. Magonov, C. Su, C.F. Quate, O. Solgaard, *Nat. Nanotechnol.* **2** 507 (2007)
75. M.A. Lantz, S.J. O'Shea, A.C.F. Hoole, M.E. Welland, *Appl. Phys. Lett.* **70** 970 (1997)
76. R.W. Carpick, D.F. Ogletree, M. Salmeron, *Appl. Phys. Lett.* **70** 1548 (1997)
77. K. Fukuzawa, S. Terada, M. Shikida, H. Amakawa, H. Zhang, Y. Mitsuya, *J. Appl. Phys.* **101** 034308 (2007)
78. C.T. Gibson, G.S. Watson, S. Myhra, *Nanotechnology* **7** 259 (1996)
79. J.P. Cleveland, S. Manne, D. Bocek, P.K. Hansma, *Rev. Sci. Instrum.* **64** 403 (1993)
80. N. Morel, P. Tordjeman, M. Ramonda, *J. Phys. D Appl. Phys.* **38** 895 (2005)
81. M.D. Simon, L.O. Heflinger, A.K. Geim, *Am. J. Phys.* **69** 702 (2001)
82. P.A. Dunne, J. Hilton, J.M.D. Coey, *J. Magn. Magn. Mater.* **316** 273 (2007)
83. L. Huang, C. Su, *Ultramicroscopy* **100** 277 (2004)
84. O. Sahin, N. Erina, *Nanotechnology* **19** 445717 (2008)
85. J.E. Sader, J.W.M. Chon, P. Mulvaney, *Rev. Sci. Instrum.* **70** 3967 (1999)
86. M. Reinstaedtler, U. Rabe, V. Scherer, J.A. Turner, W. Arnold, *Surf. Sci.* **532** 1152 (2003)
87. N. Mullin, J. Hobbs, *Appl. Phys. Lett.* **92** 053103 (2008)

Part II Characterization

Simultaneous Topography and Recognition Imaging

A. Ebner, L.A. Chtcheglova, J. Preiner, J. Tang, L. Wildling, H.J. Gruber, and P. Hinterdorfer

Summary. Atomic force microscopy (AFM) has developed into a key technique for the investigation of biological samples. In contrast to other microscopy methods, high lateral resolution down to the nanometer scale and measurements under physiological conditions are possible. Additionally, the piconewton force sensitivity allows accurate data collection for single-molecule interactions. This chapter describes the combination of high-resolution imaging and single-molecule interaction measurements. In the so-called topography and recognition imaging (TREC) mode, the scanning AFM tip is upgraded into a molecular sensor by anchoring a ligand to the tip. Enhanced electronics, including a recently developed feedback loop, allow measurement of the sample topography while simultaneously mapping ligand-binding sites. This results in topography images recorded alongside with recognition images, thereby allowing accurate allocation of the binding sites with a lateral resolution of one to a few nanometers. TREC has been successfully used for recognition imaging on isolated proteins, native and artificial membranes, and cells.

Key words: Atomic force microscopy, Molecular recognition, Recognition imaging, Single molecule, Tip functionalization, TREC.

Abbreviations

AFM	Atomic force microscope/microscopy
CFTR	Cystic fibrosis transmembrane conductance regulator
MAC	Magnetic AC mode
PEG	Poly(ethylene glycol)
RBC	Red blood cell
TREC	Topography and recognition imaging
MRFS	Molecular Recognition Force Spectroscopy

10.1 Introduction

The potential of the atomic force microscope (AFM) to measure ultralow forces at high lateral resolution has paved the way for molecular recognition studies [1]. The AFM can resolve nm-sized details and obtain high-resolution structural details on biological samples such as proteins, nucleotides, membranes, and cells in physiological, aqueous environments [2]. In particular, the specimen can be imaged in its native state, thus there is, no need to stain, freeze, coat, or crystallize biological samples. In addition, because of its force detection sensitivity, AFM has opened up the possibility of measuring inter- and intramolecular forces of biomolecules on the single-molecule level [3–6]. These capabilities have led to the development of the method described in this chapter that enables us to investigate the interaction of a single ligand molecule with its cognate receptor while simultaneously recording a high-resolution topography image.

Single molecular interaction forces are typically studied in force spectroscopy experiments [7–14]. In a force-distance cycle, an AFM tip carrying ligands is brought in contact with a surface that contains the respective cognate receptors, so that a receptor/ligand bond is formed. This bond is subsequently broken at a characteristic measurable unbinding force by retracting the tip from the surface. Such experiments allow for estimation of affinity, rate constants, and structural data of the binding pocket [5, 12, 15]. Comparison of these values with those obtained from ensemble-average techniques and binding energies [16, 17] is of particular interest. Furthermore, these force spectroscopy experiments provide insight into the molecular dynamics of the receptor–ligand recognition process [12] and even render mapping of the interaction potential possible [18]. Similar experimental strategies were used for studying the elastic properties of polymers by applying external forces [19–21] and investigating unfolding–refolding kinetics of filamentous proteins in pull–hold–release cycles [6, 19], and, more recently, to determine unfolding free energies of various membrane proteins [22, 23].

In the first attempt of localizing antigenic sites via force spectroscopy, force-distance cycles with tips that were functionalized with antibodies were performed during linear lateral scans on a surface to which the cognate receptor, human serum albumin, was covalently attached [5, 15]. Binding probabilities were determined in dependence on the lateral position, resulting in binding profiles for single human serum albumin molecules that showed a maximum. This allowed for the determination of lateral positions of antigenic sites with 1.5 nm accuracy. A complete 2-D map of forces had previously been generated to study adhesion forces of polymer surfaces and cells [24–26]. This measuring principle, called force volume mode, was first applied to specific ligand–receptor interactions [27] by performing force-distance cycles with a biotinylated tip in defined areas of a micrometer-sized photo-patterned streptavidin surface. Topography images were simultaneously acquired from the contact region of the force-distance cycles. A higher correlation between

specific molecular recognition and topography approaching the single-molecule level was revealed using intercellular adhesion molecules on the surface [28] and the cognitive antibodies on the tip. Force volume has also been applied to cells [29–31]. Albeit exciting biological questions were addressed in all these studies, lateral resolutions and/or data acquisition times attainable by AFM imaging modes were not achieved.

The velocity of recognition force mapping is mainly restricted by the hydrodynamic forces acting on the cantilever during force-distance cycles. The experimental time for recording a high-resolution force map (512×512 pixels) is ~ 7.3 h (20 Hz sampling frequency of force-distance cycles, forth and back line scan) for commercially available cantilevers (length $\sim 100 \mu\text{m}$). Ultrashort cantilevers (length $\sim 5 \mu\text{m}$) enable ~ 25 times faster z -scans in aqueous environment [32,33]. A further enhanced speed of recording recognition maps with 3 nm lateral resolution was introduced [34,35] by Raab et al. using the magnetic AC mode (MAC mode) [36], which is also known as gentle imaging tool for obtaining high-resolution images of DNA [37,38] and soft biological membranes [39] in aqueous environments. A magnetically oscillated AFM tip that carried a tethered antibody was scanned along a surface to which the antigen was bound [34] and the recognition sites were detected from the amplitude reduction arising from the antibody–antigen interactions. Since the oscillation frequency used with these cantilevers is ~ 5 kHz (slightly below the resonance frequency of the cantilever in buffer), data acquisition can be 250 times faster than in conventional force mapping, and 10 times faster than in force mapping using ultrashort cantilevers.

Now, almost 10 years after this pioneering work, a steady refinement of these strategies has led to the simultaneous topography and recognition imaging-technique (TREC) [40–42]. TREC has become an indispensable tool for a quick local mapping of receptor-binding sites in molecular biology and bionanotechnology. In recent years, the technique has been applied to model receptor–ligand pairs [41–43], remodeled chromatin structures [44], protein lattices [45], isolated cell membranes [46], and cells [47–49].

This chapter describes the principles of Topography and Recognition (TREC) Imaging. Several protocols for anchoring different ligands to tips and probe surfaces are given. A variety of ligands reaching from small molecules, such as biotin and hormones, to large biomolecules, such as proteins and virus particles, can be bound. The physical properties beyond the measuring principles are explained. On the basis of this, we give a strategy for the optimization of the operating conditions for improved data recording. The last part of the chapter shows representative examples of TREC studies, covering the whole field of applications done so far that demonstrate the versatility of this technique.

10.2 AFM Tip Chemistry

Attachment of ligand molecules onto the measuring tip of an AFM converts it into a biospecific molecular sensor with which complementary receptor molecules can be detected on a sample surface. For use in molecular recognition force spectroscopy (MRFS), AFM tips can be functionalized in many different ways: Proteins can be attached to unmodified silicon nitride or silicon tips or to silanized tips by nonspecific adsorption or via short cross-linkers. By analogy, gold-coated tips can be functionalized with proteins, peptides, and nucleic acids via short thiol linkers. In all these cases, the ligands are very closely tied to the surface and can hardly reorient, thus the encounter of a tip-bound ligand with a surface-bound receptor rarely leads to a specific binding event. One possibility to overcome this problem is to use a high number of ligands on the tip. The more elegant way is to tether few ligands to the AFM tip via long, flexible polymer chains. Some groups have applied carboxymethyl-amylose or carboxymethyl-dextran for this purpose [27, 50–57], but poly(ethylene glycol) (PEG) chains with a length between 2 and 60 nm have been used more frequently in MRFS studies [5, 10, 11, 15, 50–53, 58–64, 67–76] (Wildling, in preparation).

While PEG chains between AFM tips and ligand molecules are merely advantageous for use in MRFS, they are inherently necessary for TREC measurements, as explained in Sect. 10.3. Attachment of ligands to AFM tips via PEG chains is done in three steps (Fig. 10.2): In the first step, amino groups are generated on the tip surface. Next, PEG chains are attached with one end to the amino groups on the tip. In the last step, a ligand molecule is coupled to the free-tangling end of PEG.

With respect to aminofunctionalization, it was proven to be essential to use methods that do not increase roughness and/or stickiness of the tip surface. This was accomplished by two different methods (Fig. 10.1): (1) esterification

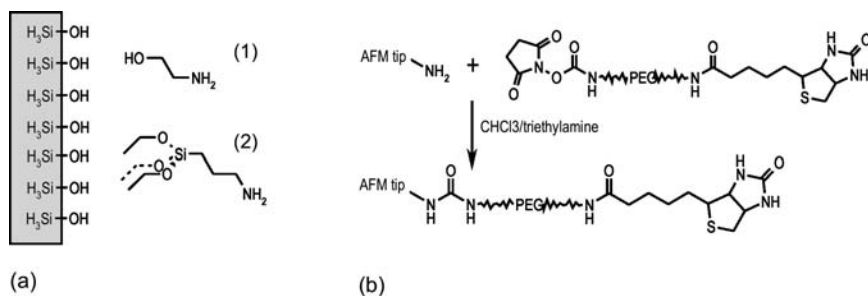


Fig. 10.1. (a) Functionalization of superficially oxidized silicon nitride tips with amino groups: (1) esterification with ethanolamine and (2) silanization with 3-aminopropyltriethoxysilane from the gas phase. (b) Attachment of presynthesized ligand-PEG conjugates to aminofunctionalized AFM tips, with the aminoreactive NHS-ester of biotin-PEG-NHS

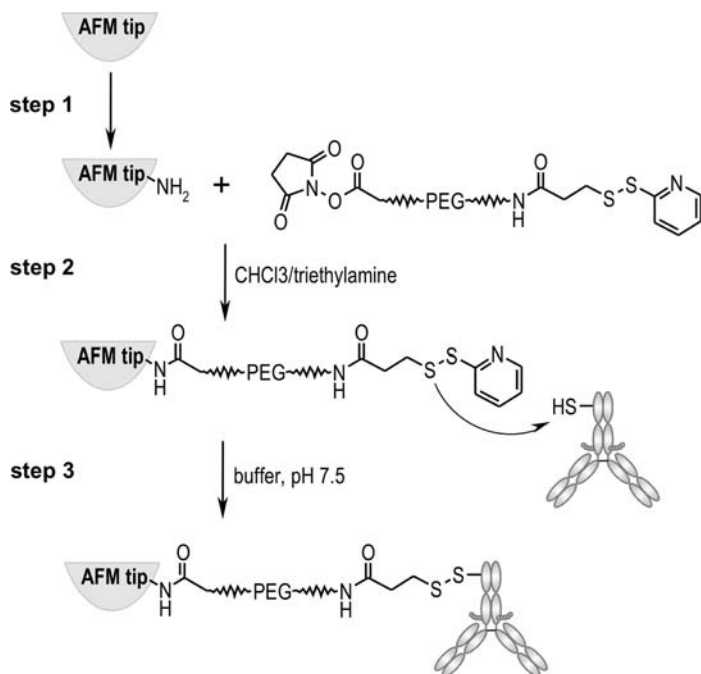


Fig. 10.2. Tethering of thiol proteins to aminofunctionalized tips: (step 1) Functionalization of silicon nitride tips with amino groups, (step 2) coupling of PDP-PEG-NHS by amide bond formation and (step 3) linking of disulfide protein. Alternatively, tip-PEG-S-S-pyridine can be reduced to tip-PEG-SH, followed by coupling of a thiol-reactive molecule

of the superficial silicon oxide layer with ethanolamine hydrochloride in DMSO [59, 77] or (2) gas phase silanization with 3-aminopropyltriethoxysilane.

Attachment of PEG to the amino groups on the tip is always done by amide bond formation, for which, all PEG linkers possess an activated COOH group, in the form of an *N*-hydroxysuccinimide ester (NHS ester), as exemplified by “NHS-PEG-PDP” in Fig. 10.2.

In Fig. 10.2, the last step involves disulfide bond formation between a thiol protein (HS-protein) and the pyridyl-S-S group on the outer end of the PEG chain. Figure 10.3 shows that disulfide bond formation (Fig. 10.3a) is one of many options to link proteins or other sensor molecules to PEG. The disulfide method is ideal for proteins that contain one or few endogenous cysteine residues. Unfortunately, this does not apply to antibodies that are most frequently used as sensor molecules on AFM tips. One way to solve this problem is to prederivatize antibodies with a small thiol linker [61]; however, this requires chemistry skills, and a high consumption of antibodies. We therefore introduced a new benzaldehyde linker (Fig. 10.3b) that can link underderivatized antibodies and other proteins via their endogenous lysine residues, of which 80–90 are found per antibody molecule [78].

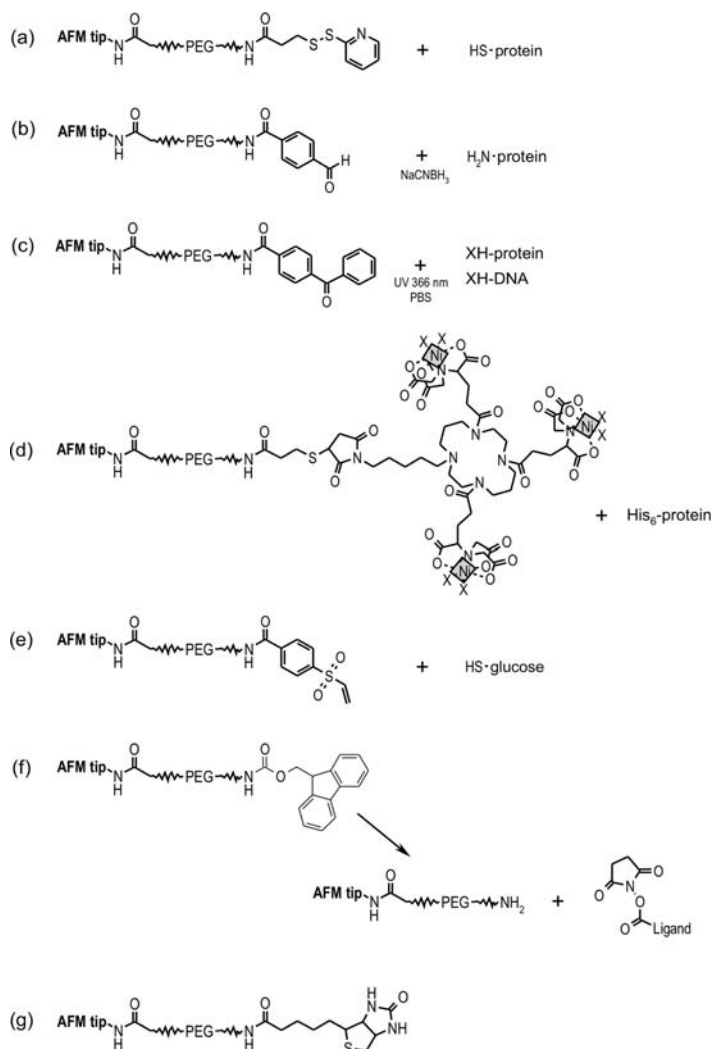


Fig. 10.3. (a) Tethering of thiol proteins to aminofunctionalized tips via PDP-PEG-NHS in two separate steps: amide bond formation and disulfide coupling. (b) Use of heterobifunctional aldehyde-PEG-NHS for flexible attachment of underivatized protein to aminofunctionalized AFM tips, the C=N double bond is usually fixed by reaction with sodium cyanoborhydride. (c) Coupling of underivatized proteins to aminofunctionalized AFM tips via benzophenone-PEG-NHS by amide bond formation and photolinking. (d) Attachment of His₆-protein to tris-NTA-PEG-tip. (e) Tethering of small thiol ligands to aminofunctionalized AFM tips via vinylsulfone-PEG-NHS. (f) Coupling of small molecules to an AFM tip via Fmoc-PEG-NHS. Deprotection of Fmoc with piperidin, DMF and coupling of small bioligands via nucleophilic substitution on the carboxylic group. (g) Attachment of presynthesized ligand-PEG conjugates to aminofunctionalized AFM tips, with the amino-reactive NHS-ester of biotin-PEG-NHS

The most general coupling of proteins can be done via benzophenone-PEG-NHS (Fig. 10.3c). The benzophenone group can be excited by UV light to the reactive triplet state many times and the excited state has a lifetime of several minutes [79], thus the ability of coupling with XH residues ($-\text{CH}_3$, $-\text{NH}_2$, and $-\text{SH}$) is maintained over a long time during irradiation [80]. Moreover, this coupling method is applicable not only to proteins but also to DNA and in no case it requires prederivatization of the sensor molecule.

A convenient alternative to covalent coupling is noncovalent linking of His₆-tagged proteins to a Tris-NTA function on the PEG chain (Fig. 10.3d). This is particularly fortunate because most genetically engineered proteins are expressed with His₆ tags anyway. Even with a single NTA function on the PEG, the binding strength of the NTA-Ni²⁺-His₆ bridge was found to be significantly larger than the unbinding forces of typical ligand-receptor pairs [13, 81]. Nevertheless, the Tris-NTA function is preferred because the three parallel Ni²⁺ bridges ensure irreversible attachment of the His₆ protein to the tip [82].

Functionalization of tips with small ligands is more critical than with proteins or other large biomolecules because blocking of the receptor recognition site on the ligand through the chemical coupling process must be avoided. Several strategies were developed to cope with these problems. Small thiol ligands, such as thio-glucose, can be tethered to the AFM tip via a vinylsulfon group (Fig. 10.3e) [64]. For the attachment of biotin (vitamin H) and hormones with COOH functions, a new cross-linker with a fluoren-9-yl-methoxycarbonyl-protected (Fmoc) amine was used (*see* Fig. 10.3f [83]).

In the special case of biotin, one can also take advantage of the fact that biotin-PEG-NHS can be purchased in different lengths and thus directly attached to tip $-\text{NH}_2$ in a single step (*see* Fig. 10.3g). This system composed of tip-bound biotin and surface-bound avidin can be used for various purposes: (1) as a start-up system for beginners in MRFS (Riener 2003), (2) as a test system for TREC (*see* Sect. 10.3), and (3) to quantitatively estimate the number of tip-bound PEG chains (per μm^2) in order to control the efficiency of different aminofunctionalization methods [59, 74].

10.3 Operating Principles of TREC

The operating principle of TREC is based on MAC mode AFM [84]. Moreover, a molecule tethered to the AFM tip [61, 74, 85] via a flexible PEG linker is capable of binding to a cognate molecule (receptor) on a sample surface (*cf.* Sect. 10.2). This functionalized tip is oscillated close to its resonance frequency while scanning line by line across the surface. The binding sites are evident from the reduction in the oscillation amplitude, as a result of specific recognition during the lateral scan. Special signal processing in combination with a modified feedback loop [58] provides a recognition image simultaneously acquired alongside the topography image. In a nutshell, the separation

of topographical and recognition signals is achieved by splitting the cantilever oscillation amplitude into lower and upper parts (with respect to the cantilevers resting position) containing solely information about topography and recognition, respectively. These parts are then used to record topography (lower parts) and recognition image (upper parts) at the same time.

This chapter explains the measuring principles of TREC and shows how the operating conditions have to be adjusted [Preiner, J., Ebner, A., Chtchelglova, L., Zhu, R., and Hinterdorfer, P. (2009). Simultaneous topography and recognition imaging: physical aspects and optimal imaging conditions. *Nanotechnology* 20, 215103.] The most important requisites for TREC can be summarized to the following points:

1. An AFM cantilever with a low Q factor in liquid.
2. A ligand molecule coupled via a flexible linker to the AFM cantilever tip.
3. Receptor molecules (complementary to the ligand on the tip) immobilized on a surface.
4. A special feedback loop, to separate topography and recognition data.
5. An appropriately chosen oscillation amplitude.
6. An appropriately chosen driving frequency.
7. Specificity proof of the detected interactions.

For the following discussion of this procedure, we use the well-described avidin–biotin interaction that has been shown to be a good starting point for new TREC users, since it is (1) robust and reliable, (2) well known in terms of binding properties, and (3) easy to prepare with commercially available components. Furthermore, (4) avidin can be simply adsorbed to mica, which is an ideal substrate for AFM imaging, and finally (5) the biotin PEG tether is commercially available and can be attached to amino-functionalized AFM tips in one step. Before being applied to TREC, the tips are usually examined for containing a functional biotin residue by force spectroscopy experiments on a mica surface densely covered with avidin molecules. Force spectroscopy with an oscillating tip offers an important tool for adjusting the amplitude for TREC measurements. In simultaneously recorded amplitude-distance and force-distance cycles (Fig. 10.4), the oscillation amplitude and the cantilever bending force can be concurrently investigated. This combination yields the length of the stretched cross-linker, which is important for the adjustment of the amplitude (as will be discussed in Sect. 10.3) for TREC imaging. The amplitude of the cantilever oscillation during approach (Fig. 10.4a, trace, grey line) remains constant (= free amplitude, A_0) until the tip comes into physical contact with the surface. From thereon, the amplitude linearly decreases until it is fully reduced to zero, when the tip is in full contact with the sample. On retraction of the cantilever from the surface (retrace, solid line), the amplitude starts to increase and reaches the free amplitude again when the tip loses contact with the surface. During further retraction, an additional amplitude reduction occurs, originating from the stretching of the PEG-tether molecule, which has bound to an avidin molecule on the mica surface via its biotin

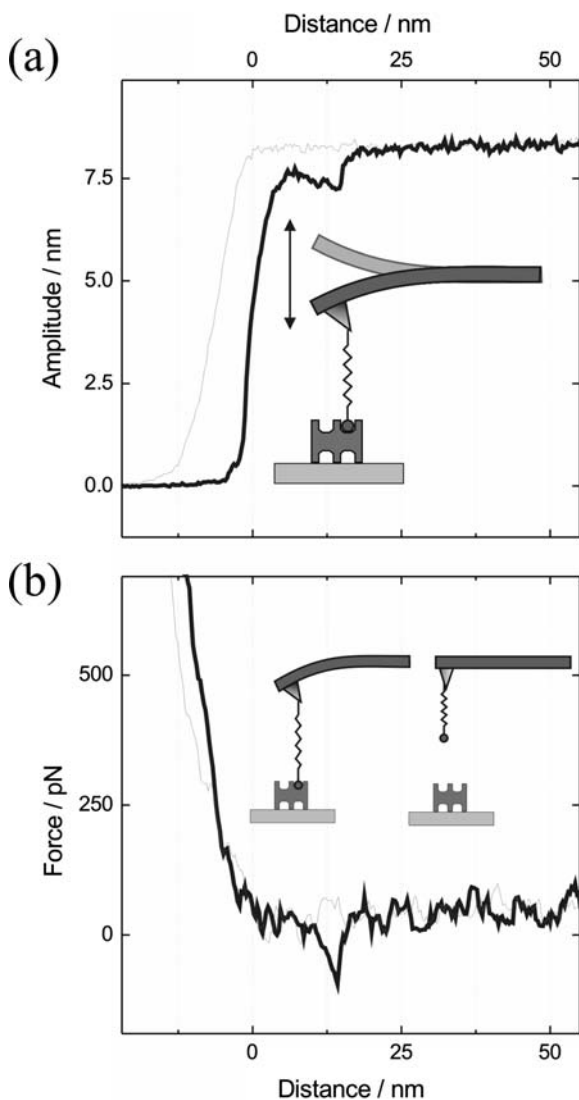


Fig. 10.4. Raw data of an amplitude-distance and a force-distance cycle showing a single avidin–biotin unbinding event. **(a)** In the amplitude-distance cycle, a single unbinding event is shown at 14 nm, resulting in an amplitude reduction of 1 nm at 8 nm free amplitude. **(b)** The simultaneously acquired force-distance cycle exhibits the same unbinding event with an unbinding force of 80 pN. Figure taken from [58]

end. Finally, the steady increase of the tip-surface distance leads to dissociation of the avidin–biotin bond, accompanied by a sharp increase (~ 1 nm at 14 nm tip-surface distance) of the oscillation amplitude, restoring the free amplitude. The shape of the unbinding event in the simultaneously recorded

force-distance cycle (Fig. 10.4b, retrace, black line) is dominated by the non-linear stretching behavior of the tether on tip retraction from the surface (Sect. 10.2, Fig. 10.1). In contrast to the amplitude-distance cycle, the force-distance cycle monitors the cantilever deflection, which directly translates into a force (F) according to Hook's law ($F = k \cdot \Delta z$) with k being cantilever spring constant and Δz being cantilever deflection. Avidin–biotin unbinding results in a sudden jump to zero deflection (retrace, Fig. 10.4b). Cantilevers that show unbinding events in the force- and amplitude-distance cycles can be subsequently used for TREC imaging.

10.3.1 Half-Amplitude Versus Full-Amplitude Feedback

First, we discuss and explore the necessity of a special feedback loop used for TREC imaging to measure the topography unbiased. During recording of an image in dynamic mode AFM with a bare AFM tip, the oscillating cantilever is scanned linewise across the sample surface. Hereby, the cantilevers movement slightly deviates from an ideal sinusoidal oscillation. These deviations usually arise from the sample topography [86] and are initiated during the downward swing of the cantilever oscillation, as the tip hits the sample surface. In this conventional mode, the peak-to-peak value of the oscillating amplitude (“full amplitude” – FA) is utilized as feedback parameter (“FA feedback”) to obtain the sample topography. To this end, the FA is held constant by the feedback loop through the voltage applied to the piezo-actuator that controls the z -distance between the cantilever base and the surface [87, 88]. These voltage signals are proportional to the height of the topographical features of the sample and thus yield all information for the topography image. However, this holds true only for perturbations exclusively caused by repulsive tip-sample interactions, which will ideally (in highly overdamped liquids environment) only influence the bottom peaks of the oscillation. In contrast, when a tip-tethered ligand binds to its receptor on the sample surface (i.e., when molecular recognition occurs), the tether molecule will be stretched during the upward movement of the cantilever. The resulting loss in energy will in turn cause the top peaks of the oscillations to be lowered. As a direct consequence, in FA feedback with a ligand-functionalized AFM tip, both topographical features and molecular recognition will affect the value of the FA. This will therefore result in inaccurate height values of the topographical features [34]. Consequently, a feedback parameter solely depending on the sample topography has to be used in order to unequivocally separate topography from recognition.

Low quality (Q)-factor cantilevers in combination with a proper chosen driving frequency enable a temporal separation of perturbations in the sinusoidal cantilever movement caused by topographical features (affecting the bottom peaks of the oscillations) from molecular recognition events (affecting the top peaks of the oscillations) [89]. Therefore, it appears beneficial to use the lower part of the oscillation amplitude as a feedback parameter. This

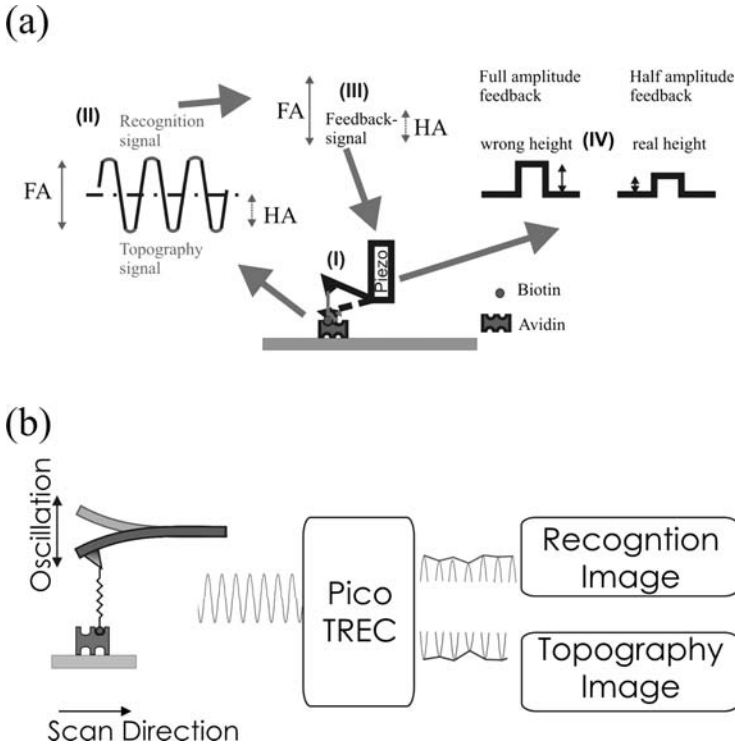


Fig. 10.5. (a) Sketch of the different feedback loops. (I) When the tip tethered biotin has bound to surface attached avidin, the oscillatory movement (II) of the cantilever is altered in two ways: First, the topography is sensed at the bottoms of the oscillation (grey) leading to a reduction of the amplitude in this part, and, second, the stretching of the PEG-linker molecule reduces the movement in the upper parts of the oscillation. The corresponding HA and FA feedback parameters are sketched in (III). In contrast to the HA parameter, the FA parameter is biased by the additional amplitude reduction due to molecular recognition. Consequently, the piezo actuator of the scanner has to additionally withdraw in FA-feedback, leading to an increased and hence incorrect height of the avidin molecules (IV). Figure taken from [Preiner, J., Ebner, A., Chtcheglova, L., Zhu, R., and Hinterdorfer, P. (2009). Simultaneous topography and recognition imaging: physical aspects and optimal imaging conditions. *Nanotechnology* 20, 215103.]. (b) The principle of TREC imaging: The cantilever oscillation is split into lower and upper parts, resulting in simultaneously acquired topography and recognition images. Figure taken from [58]

has been technically realized with the installation of a “half-amplitude” (HA) feedback loop. A detailed comparison of FA and HA feedback loops is depicted in Fig. 10.5a.

The technical implementation of this mechanism is illustrated in Fig. 10.5b. The time-resolved deflection signal of the oscillating cantilever is low-pass

filtered to remove the thermal noise, DC-offset leveled, and amplified, before it is split into the lower and upper parts of the oscillations. The signal passes a trigger threshold on each path, and the lower and upper peaks of each oscillation period are determined by means of sample and hold analysis. Successive peaks result in a staircase function, which is subsequently filtered and fed into the microscope controller. The lower parts (minima of oscillations) are used for driving the feedback loop (HA feedback) so as to record the topographical image. The upper parts (maxima of oscillations) provide the data for the recognition image and do not influence the piezo-actuator movement.

To show the difference of the FA and HA feedback systems experimentally, the well-known avidin–biotin system was chosen. We imaged a mica substrate, moderately covered with avidin molecules using an IgG-biotin functionalized AFM tip (see Chap. 1). After the adjustment of the oscillation amplitude to a value slightly smaller than the contour length of the tether molecule [free amplitude ($A_0 = 14.3$ nm) and choosing an amplitude set point of $A_{sp}/A_0 \sim 0.95$ and a driving frequency of 13 kHz (close to the cantilever resonance frequency (for details, see Sects. 10.3.2 and 10.3.3)], recognition events between the tip-tethered IgG biotin and the avidin molecules on the mica substrate were successfully imaged. This experiment was first conducted in FA feedback and the recorded topography image is depicted in Fig. 10.6a. With the exception of one bigger accumulation of avidin molecules in the center of the image, mainly single avidin molecules were observed. The dashed lines (1–10) in this image mark the lateral location of the cross sections, shown in Fig. 10.6b. The recognition image (Fig. 10.6c) depicts the location of recognition events (dark spots, corresponding to an amplitude reduction in the top peaks of oscillation) on the scanned area. For the analysis of the topographical heights of the avidin molecules, only cross sections through single molecules recognized by the IgG biotin were used (i.e., molecules at locations that also exhibit a dark spot in the recognition image). The mean height of the avidin molecules determined from these cross sections was 4.2 ± 0.3 nm, and the corresponding mean amplitude reduction in the recognition image was 0.64 ± 0.04 nm. This experiment was then repeated with the same amplitude and frequency settings, but in HA feedback. The corresponding images are depicted in Fig. 10.6d–f. The very same molecules were recognized by the IgG biotin tethered to the tip, but the analysis of the corresponding cross sections yielded different results. The topographical height of the avidin molecules recorded in HA feedback was 3.1 ± 0.3 nm, that is 1 nm less than in FA feedback. Analysis of the mean values of the corresponding amplitude reductions (Fig. 10.6f), yielding a mean and standard deviation (SD) of 0.31 ± 0.02 nm.

These experiments clearly demonstrate that in case of molecular recognition between the ligand tethered to the tip and the receptor molecule on the surface, the conventionally used feedback parameter (FA) is affected by the stretching of the tether molecule. This results in an additional movement (~ 1 nm) of the z -piezo actuator that increases the force on the cantilever

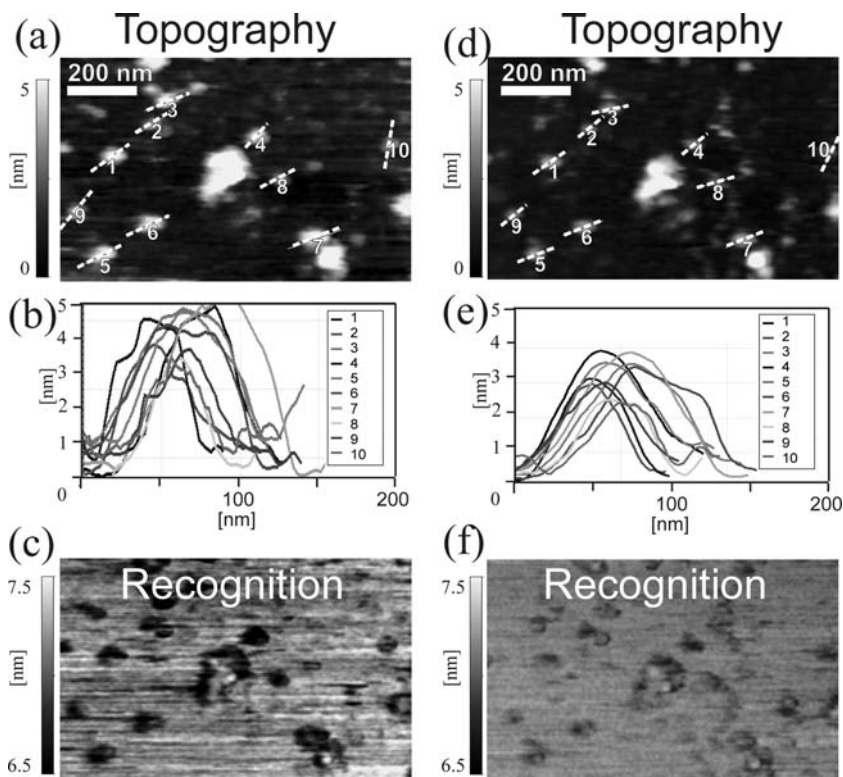


Fig. 10.6. (a) Topography image recorded with FA feedback. (b) Cross sections indicated in (a). (c) Recognition image, simultaneously recorded with (a). (d) Topography image recorded in HA feedback. (e) Cross sections indicated in (d). (f) Recognition image, simultaneously recorded with (d). Scansize: $1 \times 0.5 \mu\text{m}$. Figure taken from [Preiner, J., Ebner, A., Chtcheglova, L., Zhu, R., and Hinterdorfer, P. (2009). Simultaneous topography and recognition imaging: physical aspects and optimal imaging conditions. *Nanotechnology* 20, 215103.]

in the upward movement. As a consequence, the amplitude reduction (the recognition signal) is increased in FA-feedback (0.64 nm) compared with HA-feedback (0.31 nm). Although the recognition signal is more pronounced using FA-feedback, these experiments clearly show that only TREC imaging using a HA-feedback allows an accurate determination of the surface topography.

10.3.2 Adjusting the Amplitude

Stretching of the linker molecule on recognition of the receptor through the ligand is characterized by a parabolic force increase until the linker is completely extended. [21] (Fig. 10.4b). Only linker extensions slightly smaller than the fully extended length exert a measurable force on the cantilever in the top

peaks of oscillations so as to generate a clear recognition signal. On the contrary, for a strong and pronounced recognition signal, it is important that during line scans, a formed ligand–receptor bond persists until the tip has laterally moved away from the receptor molecule. This sets the upper limit of the oscillation amplitude, as the ligand must not dissociate from the receptor during an oscillation period. Thus, a small gap of proper linker extensions defines three different regimes of oscillation amplitudes as shown in Fig. 10.7a. In the first regime (I) the ligand (biotin) molecule has bound to the receptor (avidin) on the surface. The oscillation amplitude is too small, so that the linker does not exert an efficient force on the cantilever in the upward swing and no recognition signal is generated. In the second regime (II), the amplitude is increased, resulting in an efficient stretching of the linker without detaching the biotin from the avidin in the upward swing. This results in a proper recognition signal. In the third regime (III), the amplitude is too high, causing the biotin to unbind from the avidin in every top peak of oscillation immediately following the binding. Hence, no recognition signal is generated. In order to assure a sufficient linker stretching without detaching the biotin from the avidin when scanning laterally over the avidin molecule, the amplitude has to be adjusted to regime (II) (Fig. 10.7a). This amplitude setting will result in a clear recognition signal.

The regimes (I)–(III) were investigated with a single IgG-biotin functionalized AFM tip scanned over a mica surface, moderately covered with single avidin molecules. Figure 10.7c shows the topography (left) and simultaneously recorded recognition image taken with a free amplitude of 14.3 nm and $A_{\text{sp}}/A_0 \sim 0.95$ in HA feedback. The dark spots in the recognition image clearly reveal recognition events between the IgG-biotin tethered to the tip and the avidin molecules on the mica surface. Then, the free amplitude was decreased and increased several times (for the corresponding amplitude vs distance curves cf. Fig. 10.7b) in the range between 11.4 and 20.4 nm. The excitation frequency was 13 kHz and the amplitude/free amplitude ratio was always set to $A_{\text{sp}}/A_0 \sim 0.95$ in HA feedback. Recognition signals could only be observed at amplitudes of 14.3 and 16 nm (Fig. 10.7c, e, g, i, and l), indicating the amplitude boundaries for the observation of molecular recognition, that is the regime (II) as described above. A lower amplitude [regime (I)] results in an insufficient stretching of the linker (amplitudes of 11.4 and 12.7 nm, respectively; Fig. 10.7d and h), whereas a higher amplitude [regime (III)] averts continuous binding until the cantilever veers away laterally from the avidin molecules (amplitudes of 17.4, 18.1, and 12.7 nm, respectively; Fig. 10.7f, j, and k). Both conditions do not yield recognition signals. Generally, the ideal amplitude regime for observation of recognition differs from one functionalized cantilever to the other. It depends on the length of the linker molecule, on the exact location of the linker molecule on tip apex, and the size of the attached molecule. It typically lies in the range of 10–20 nm.

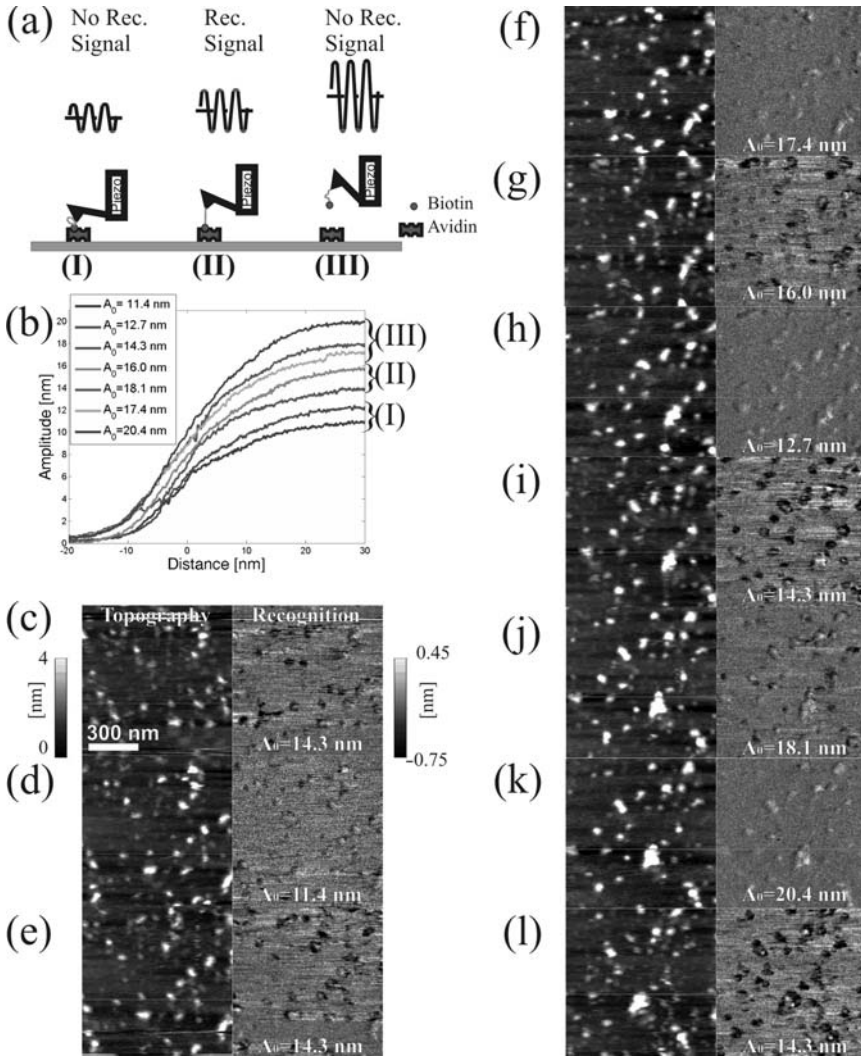


Fig. 10.7. Amplitude adjustment. (a) Three different amplitude regimes can be identified. (b) Amplitude–distance curves of the free amplitudes (varied from 11.4 to 20.4 nm) used to identify the different regimes sketched in (a). (c)–(l) Time series of subsequent images, recorded with free amplitudes (A_0) as indicated. The investigated amplitude range can be divided into three parts [cf. (a) and (b)]: (I) no recognition signal between 11.4 and 12.7 nm, since the linker is stretched to less; (II) recognition signal between 14.3 and 16 nm, linker stretching is efficient to gain a signal; (III) no recognition signal from 17.4 to 20.4 nm, since the amplitude is too high for the biotin to bind continuously to avidin. The data displayed in the recognition image correspond to the difference of the actual amplitude and the unperturbed free amplitude of the upward swing (similar to the amplitude error image). Figure taken from [Preiner, J., Ebner, A., Chtcheglova, L., Zhu, R., and Hinterdorfer, P. (2009). Simultaneous topography and recognition imaging: physical aspects and optimal imaging conditions. *Nanotechnology* 20, 215103.]

10.3.3 Adjusting the Driving Frequency

Artifacts arising from the finite feedback speed can mask the recognition image. A minimization of these unwanted signals can be achieved by a thorough adjustment of the driving frequency. To understand this mechanism, it is advisable to examine the physical origin of these artifacts, which are often termed as “topographical cross talk.” When the cantilever is oscillated close to a sample surface in liquid, the equation of motion can be approximated (neglecting higher flexural modes) [86] by,

$$\frac{\ddot{x}}{\omega_0^2} + \frac{\dot{x}}{\omega_0 Q} + x = \frac{F}{k} \cos(\omega t) + \frac{F_{ts}(x)}{k}, \quad (10.1)$$

where x denotes the coordinate of the tip deflection and dots represent temporal derivatives. F , k , ω_0 , ω , Q and F_{ts} are the excitation force amplitude, the cantilevers spring constant, the natural angular resonance frequency, the angular driving frequency, the cantilever quality factor, and the nonlinear tip-sample interaction force, respectively. For a qualitative understanding of the origin of topographical cross talk in the recognition image, it is useful to restrict the analysis to $A_{sp}/A_0 \sim 1$. Then the nonlinear tip-sample interaction forces that usually introduce anharmonic contributions to the cantilever movement [86] can be neglected and Eq. (10.1) reduces to a simple forced harmonic oscillator with damping. This equation has a steady solution of the form $x_\omega(t) = A_{sp} \cos(\omega t + \varphi)$. When the feedback is not operating infinitely fast, small perturbations from this movement (denoted in the following as $\tilde{A}(t)$) can arise at the closest tip-sample distance during the downward movement of the cantilever. In this case, the actual movement of the cantilever is a superimposition of the steady state solution and the small perturbation, $x(t) = A_{sp} \cos(\omega t + \varphi) + \tilde{A}(t)$. Inserting this approach into the oscillator equation, we get a differential equation for the time evolution of the perturbation:

$$\frac{\ddot{\tilde{A}}}{\omega_0^2} + \frac{\dot{\tilde{A}}}{\omega_0 Q} + \tilde{A} = 0. \quad (10.2)$$

Starting with an initial perturbation $\tilde{A}(0)$ at time $t = 0$, the closest tip-sample separation, the perturbation evolves in time t according to the solution of the damped harmonic oscillator equation:

$$\tilde{A}(t) = \tilde{A}(0) \exp\left(\frac{-t}{\tau}\right) \frac{\sin(2\pi f_0 \sqrt{1 - 1/4Q^2} t + \varphi)}{\sin(\varphi)}. \quad (10.3)$$

After a half oscillation period ($t_{1/2} = 1/(2f_{drive})$), at the following top peak, the perturbation has decayed to:

$$\tilde{A}(t_{1/2}) = \tilde{A}(0) \exp\left(-\frac{f_0 \pi}{2Q f_{drive}}\right) \frac{\sin(\pi f_0 \sqrt{1 - 1/4Q^2} / f_{drive} + \varphi)}{\sin(\varphi)}. \quad (10.4)$$

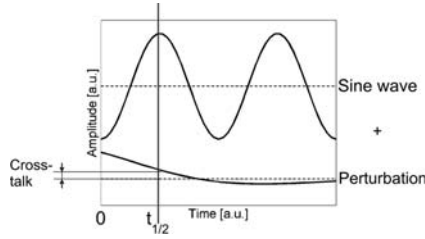


Fig. 10.8. The actual movement of a cantilever weakly interacting with a sample surface can be approximated by a superimposition of a sinewave and small perturbations arising from the finite feedback speed. These perturbations are initiated at the minimum of the oscillation (at the very left) and ring down until the maximum of the cantilever oscillation. The recognition signal is generated from these maxima, and depending on the driving frequency, these perturbation are still to a certain amount included in the maxima

$\tilde{A}(t_{1/2})$ describes the amount of amplitude error transduced to the recognition image, since the recognition image is generated from these top peaks (i.e., at $t_{1/2}$, cf. Fig. 10.8). Therefore, by determining the average $\langle \tilde{A}(t_{1/2}) / \tilde{A}(0) \rangle$ from experimental data recorded at different driving frequencies [i.e., $\tilde{A}(t_{1/2})$ from the recognition image minus A_{sp} ; and $\tilde{A}(0)$ from the amplitude error image], Eq. (10.4) can be used to obtain Q and f_0 for the conditions when the tip is weakly interacting with the sample surface.

This contribution of the amplitude error to the contrast in the recognition image was experimentally investigated with the same cantilever used for the recording of Figs. 10.6 and 10.7. To exclusively record the cross talk in the recognition image, the signals caused by molecular recognition had to be suppressed. This was achieved by adjusting the free amplitude to a value (~ 23.5 nm) that was too high for the biotin to continuously bind to avidin, as shown in Fig. 10.7 [regime (III)]. The ratio $A_{sp}/A_0 \sim 0.95$ was kept constant (by adjusting the drive) in HA-feedback and the excitation frequency was varied in the range of 7–15 kHz. Figure 10.9a–e illustrates that the cross talk in the recognition image (middle panel) caused by the amplitude error increases with the excitation frequency, whereas the topography (left panel) and amplitude error images (right panel) remain the same. The cross talk reaches the noise level at an excitation frequency of ~ 9 kHz. A quantitative analysis of this experiment is shown in Fig. 10.9f. Here, the average value of the features in the recognition image normalized by the corresponding amplitude error image, $\langle \tilde{A}(t_{1/2}) / \tilde{A}(0) \rangle$, is plotted versus the respective driving frequency (mean values with SDs). In accordance with the images (Fig. 10.9a–e), the topographical cross talk in the recognition image increases with the driving frequency. The dashed line in Fig. 10.9f corresponds to a fit of Eq. (10.4) to the experimental data points. From the fit, we obtained $Q = 0.69$ and $f_0 = 11.3$ kHz, the Q -factor and resonance frequency for the cantilever gently interacting with the samples surface. For low driving frequencies, the topographical cross talk

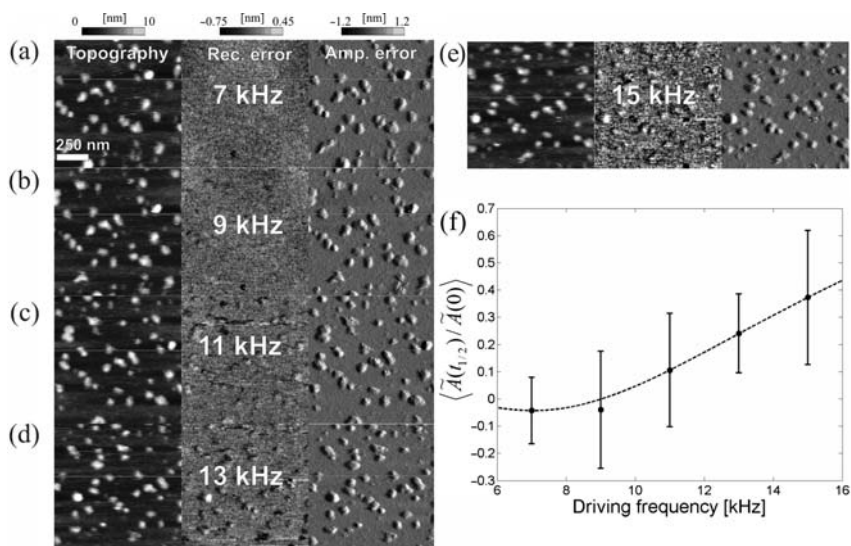


Fig. 10.9. Influence of the excitation frequency on image cross talk. Topography (*left*), corresponding recognition image (*middle*), and amplitude error image (*right*) recorded at the indicated driving frequency (a)–(e). By adjusting the drive, the free amplitude and the set point amplitude were kept constant during the recording. The data displayed in the recognition image corresponds to the difference of the actual amplitude and the unperturbed free amplitude of the upward swing (similar to the amplitude error image). Figure taken from Preiner et al. (submitted for publication). (f) Average of the cross talk signals from (a)–(e) normalized by the corresponding amplitude error image as a function of the driving frequency (dots; mean standard \pm deviation). The dashed line corresponds to a fit of the harmonic model (10.4) to the experimental data. Figure taken from [Preiner, J., Ebner, A., Chtcheglova, L., Zhu, R., and Hinterdorfer, P. (2009). Simultaneous topography and recognition imaging: physical aspects and optimal imaging conditions. *Nanotechnology* 20, 215103.]

is allowed to decay further (i.e., $t_{1/2}$ is longer) until the detection of recognition, compared with higher driving frequencies. In order to minimize the cross talk in the recognition image, the driving frequency should be adjusted to a value smaller than the cantilevers resonance frequency at the surface.

10.3.4 Proofing the Specificity of the Detected Interactions

Specificity Proof by Competitive Inhibition

Specificity proof experiments allow for unambiguous distinction of signals arising from molecular recognition and signals induced by nonspecific tip-sample interactions. The latter can be attributed to nonspecific tip-surface adhesion or artifacts caused by the finite feedback speed when the topography changes rapidly. The usual and most commonly used specificity proof for molecular

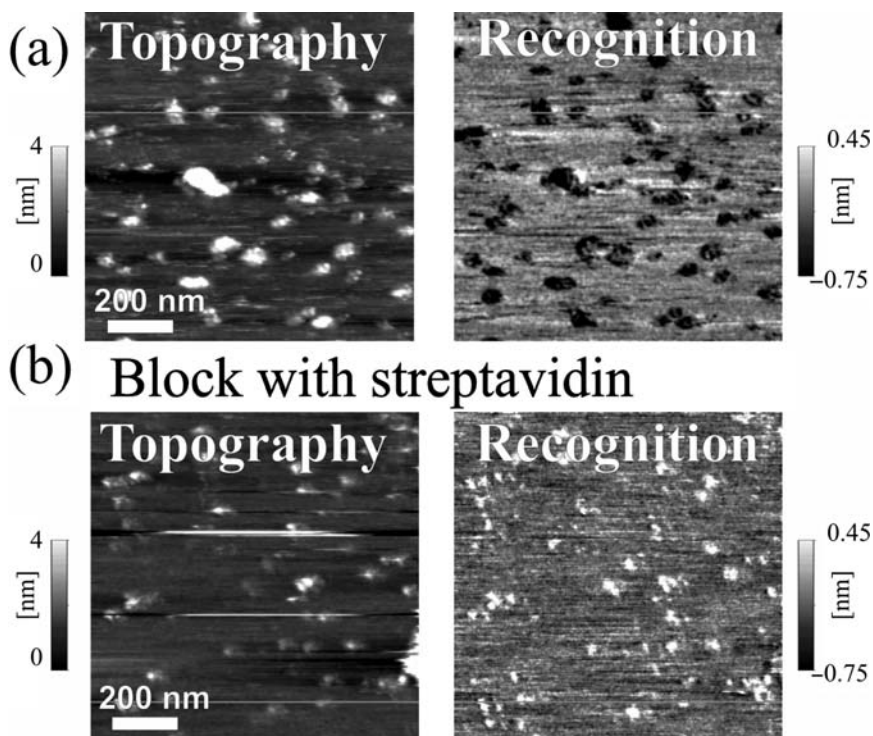


Fig. 10.10. Recognition and specificity check experiment. (a) Topography (left) and recognition image (right) before adding free streptavidin. (b) Topography (left) and recognition image (right) after addition of streptavidin. The dark recognition spots disappeared from the recognition image, showing the specificity of the block. The data displayed in the recognition image correspond to the difference of the actual amplitude and the unperturbed free amplitude of the upward swing [similar to the amplitude (error) image]. Figure taken from [Preiner, J., Ebner, A., Chtcheglova, L., Zhu, R., and Hinterdorfer, P. (2009). Simultaneous topography and recognition imaging: physical aspects and optimal imaging conditions. *Nanotechnology* 20, 215103.]

recognition events is performed by addition of free receptors into the bath solution so as to block the ligand molecule on the tip (tip block) [46,90,91]. In the case of avidin–biotin system, tip blocking is typically achieved by adding streptavidin at a high concentration (1 mg/ml) into the measuring cell. Before the tip block (Fig. 10.10a), almost all avidin molecules on the surface were recognized by the biotin on the tip (dark spots in the recognition image, right panel). After addition of streptavidin, the recognition spots vanished (Fig. 10.10b), thus indicating the specificity of the interaction.

Specificity Proof by Amplitude Variation

Specificity proof experiments appear most convincing if they are performed on the very same position. As this usually involves addition of free ligands (or receptors) to the imaging buffer during scanning, the perturbation caused thereby often induces a significant thermal drift, resulting in a loss of the previously scanned area. In addition, it can cause instabilities in the feedback leading to serious tip-damage. Moreover, after block, the receptor–ligand system is irreversibly disrupted and is unlikely to be restored for further successful TREC experiments. To overcome this limitation, the algorithm of choosing the right amplitude regime described in Fig. 10.7 is appropriate for controlling the reliability of the recognition signals without addition of a blocking agent. After establishing the right amplitude from regime (II), switching to regime (I) and (III) by decreasing or increasing the free amplitude (while keeping the ratio $A_{\text{sp}}/A_0 \sim 0.95$ constant), respectively, leads to disappearance of the recognition spots and therefore proves their specific nature.

10.4 Applications of TREC: Single Proteins, Membranes, and Cells

Determining the distribution of specific binding sites on biological samples with high spatial accuracy is an important task in many fields of biological science [92]. TREC has become an indispensable tool for high-resolution receptor mapping, which was demonstrated for a variety of molecular species on different biological surfaces. It has been successfully used to localize single antigenic binding sites on protein lattices [93], to confirm the preservation of the functionality of single antibodies on their adsorption to organic semiconductor surfaces [91], and for mapping histones within remodeled chromatin [90]. The technique was also applied to cells, for the localization of chloride channels on red blood cell (RBC) membranes [46] and the detection of adhesion proteins [94], potassium channels [95], and growth factor receptors on mammalian cells [49]. This chapter gives an overview of the most recent TREC applications.

10.4.1 Antibiotin Antibodies Adsorbed to an Organic Semiconductor

The combination of electronic devices like organic thin film transistors and nanoscale biological objects could open new approaches for the detection and measurement of properties of a diversity of biological entities. A very common material in the field of organic thin film transistors, used due to its great chemical and mechanical stability, is the molecular conjugated organic semiconductor pentacene. In fact, it is the active material in electronic paper

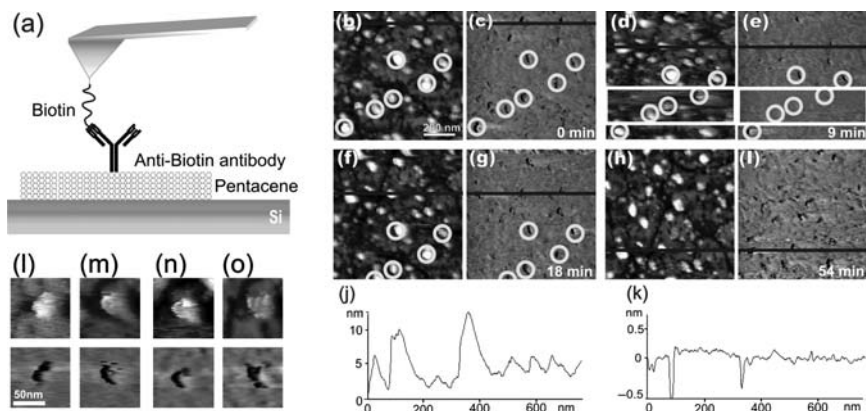


Fig. 10.11. (a) Scheme of the simultaneous topography and recognition AFM imaging of antibody adsorbed on pentacene islands. Time Series of topography (b), (d), (f), (h) and recognition image (c), (e), (g), (i) of antibody adsorbed on pentacene-islands obtained using a biotin-functionalized tip. (Gray scale range: topography images: 0 to 13 nm; recognition images: -0.5 to 0.7 nm). (j) and (k), show the cross section indicated by the lines in (b) and (c). (l)–(o) Topography (*top*) and recognition images (*bottom*) of antibody adsorbed on pentacene islands. Different shapes of recognition spots were observed that can be attributed to different orientations of the Fab fragments of the adsorbed antibodies. (Gray scale range: topography images: 0 to 13 nm; recognition images -0.5 to 0.7 nm). Figures taken from [91]

applications and can also be used to fabricate new types of conductance sensors. In order to generate specific addressable binding sites on such substrates, it is of great interest how biological molecules, capable of serving as binding sites, behave upon attachment to the organic semiconductor. This paragraph shows how the TREC technique was applied to investigate the functionality of antibody, adsorbed on pentacene islands in buffer solution [91] (Fig. 10.11).

A time series of pentacene islands after incubation with antibody recorded with a biotin-functionalized tip is shown in Fig. 10.11b–i. Antibodies with a height of ~ 5 nm are clearly observed in the topographical images (Fig. 11b, d, f, and h). The corresponding recognition images, Fig. 10.11c, e, g, and i, reveal a constant recognition pattern throughout the whole observation time of 54 min, which is evident from the black lines indicating the same cross sections (Fig. 10.11j and k). Because of the thermal drift (~ 11 nm/min), the cross sections are steadily shifted in each image. Within the framed area in Fig. 10.11d and e, the set point-amplitude was increased up to a point, where the actual peak–peak amplitude exceeds the PEG-linker length, which resulted in a disappearance of the recognition spots (circles, compare with Fig. 10.11f–i). The biotin on the tip was no longer able to bind continuously to an antibody on the surface. The increased peak–peak

amplitude leads to a rupture of the biotin from the antibody in each oscillation cycle. For this reason, the dwell time close to the antibodies on the surface is reduced, which in turn reduces the binding probability. Before and after the set point-amplitude increase (i.e., outside the framed area), the antibodies were recognized by the biotin-functionalized tip (circles).

In order to prove the specificity of the recognition signals, the cantilever used for recording the previous images was immersed into a phosphate-buffered saline solution with high streptavidin concentration (10 mg/ml, for 20 min). After this treatment, the biotin molecule on the tip was blocked by a streptavidin molecule in solution [91]. In general, IgG antibodies like the antibiotin antibodies have two flexibly linked Fab fragments, each carrying a specific binding site for biotin at the end. Antibodies are very flexible molecules and can therefore adopt many different conformations when adsorbed to a surface. Depending on the actual conformation (i.e., the angle between the Fabs) on the surface, these binding sites can be accessible for biotin. A closer look on the observed recognition spots (Fig. 10.11l–o, bottom) and the corresponding topography images (Fig. 10.11l–o, top) reveals, although blurred due to the finite PEG linker length, various different spot shapes. Consequently, these different shapes can be attributed to the conformations of the antibodies on the surface leading to different distances between the two binding sites on each antibody. In some cases (Fig. 10.11l), the binding sites were far away from each other and appeared spatially separated. This example demonstrates that biological molecules, such as IgG antibodies, can be adsorbed on organic semiconductors without losing their functionality of binding their conjugate ligand. Since the antibodies can serve as specific addressable binding sites, these findings could be exploited for the specific attachment of various biological entities (cells, viruses, etc.) to organic semiconductors.

10.4.2 Bacterial S-Layer Lattices

Besides the detection and verification of single receptor-binding sites, recognition imaging can be used for more complex systems like membranes and cells. In a recent study, bacterial surface layer (S-layer) crystals with genetically fused affinity tags were investigated using recognition imaging [93]. Such S-layers represent a self-assembling system that has been optimized over billions of years in biological evolution [96–98]. These proteins have an intrinsic ability to form isoporous two-dimensional crystals *in vitro* in suspension, on lipid films, on liposomes, and on solid supports [98, 99]. This allows generation of well-defined surface coatings by a simple bottom up approach. Additionally, the fused affinity tag *Streptag-II* with eight-amino acid introduces single reactive binding sites, resulting in a functional and addressable nanoarray. An overview image of a formed square lattice of S-layer SbpA *Bacillus Sphaericus* can be seen in Fig. 10.12. *Strep-Tactin*, a genetically engineered streptavidin variant with higher affinity toward *Streptag-II* compared

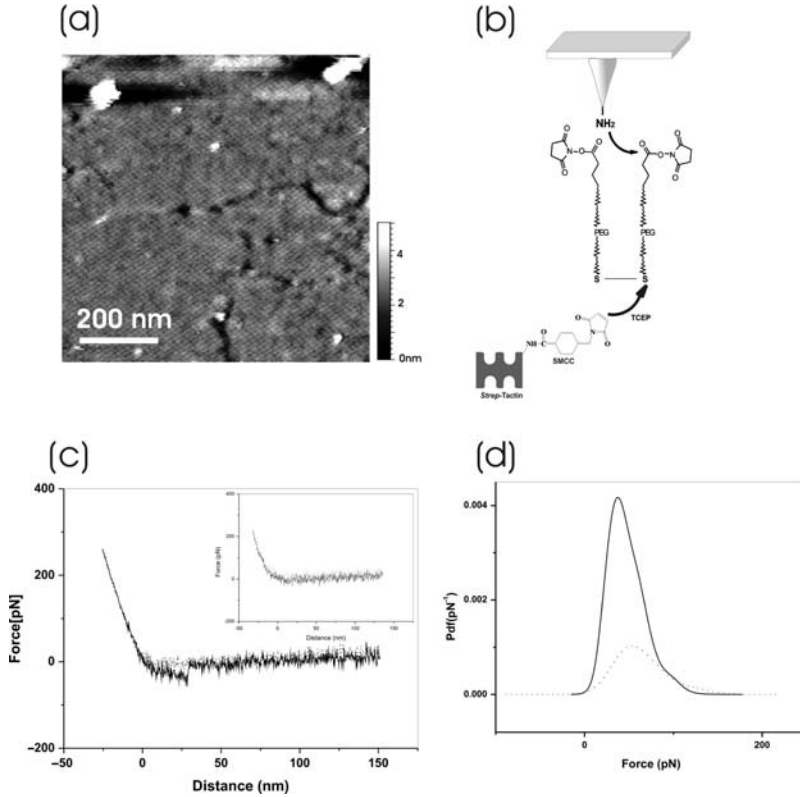


Fig. 10.12. Surface and tip proof for TREC experiments on bacterial S-layers. (a) The topography nicely shows the recrystallized S-layer lattice. (b) Using a short heterobifunctional cross-linker, *Strep-Tactin* was bound covalently to the AFM tip. (c) In molecular recognition force spectroscopy experiments, the surface and tip chemistry was investigated showing specific recognition. The inset reflects a force-distance cycle without specific interaction between *Streptag-II* and *Strep-Tactin*. (d) The distribution of the unbinding force is shown in absence (*solid line*) and presence of (*dotted line*) free *Streptag-II*. Figure taken from [93]

with streptavidin was used as specific receptor. The receptor molecule was tethered to the previously amino-functionalized tip via a rather short heterobifunctional PEG-cross-linker with an extended length of ~ 2 nm as shown in Fig. 10.12b. For recognition imaging experiments, cocrystallized S-layer proteins of SbpA-*Streptag-II* with wild-type SbpA proteins were used, resulting in an adjustable density of recognition sites. In Fig. 10.13a, the topography of the 2D crystal is shown. The center-to-center spacing is 14 nm and no detectable difference between tagged and untagged S-layer proteins is observable. This is in clear contrast to the simultaneously acquired recognition image (Fig. 10.13b). Here, the *Strep-Tactin* on the AFM tip recognizes only the S-layer proteins fused with the *Streptag-II*. The dark spots in the recognition

image indicate the location of the recognition events between Streptag-II and *Strept*-Tactin. They obviously reflect the distribution of the fusion protein rSbpA-*Streptag*-II in the cocrystallized S-layer lattice of rSbpA-*Streptag*-II and wild-type SbpA protein at a molar ratio of 1:7. The specificity of the recognition signals was proven by performing recognition imaging on the very same position before and after blocking the *Strept*-Tactin on the tip. This was done by injecting free *Streptag*-II peptides into the AFM fluid cell during imaging, resulting in a complete abolishment of the recognition signals in the recognition image (whereas the topography remains unaffected). For the investigation of the obtainable limit of the lateral resolution of TREC imaging, a smaller scan area was chosen. This allowed precisely localizing the position of the *Streptag*-II at the grid using a “center of mass” approach. Although the diameter of recognition spots was in the range of two times the PEG linker length plus the protein (in total ~ 7 nm), the position of the recognition spots was here determined with accuracy below 1 nm, impressively demonstrating the lateral precision achievable in recognition imaging. By comparing the positions of the recognition sites calculated from the recognition image recorded in opposite direction, a root mean square distance of 1.5 nm was calculated, explainable by the asymmetry of the cantilever movement and the position of the PEG tether on the tip.

To account for the effect of the scanning direction, the position of Streptag-II was defined as mean position of corresponding recognition sites calculated for both scanning directions. The overlay of the topographical image (Fig 10.13a) with the recognition spots of the simultaneously acquired recognition image (Fig. 10.13b) showed that 88% of the *Streptag*-II was found positioned on the corner of the square lattice (Fig 10.13c). The white spots indicate the position of *Streptag*-II (in correspondence to the dark spots on the

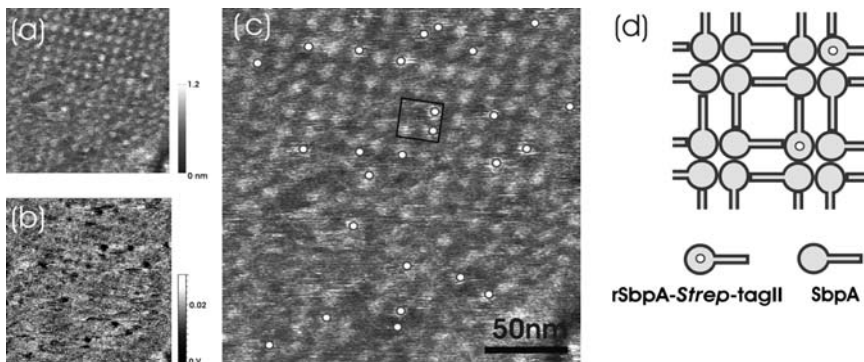


Fig. 10.13. (a) Topography and (b) recognition images of an S-layer protein crystal (cocrystallized proteins wild-type SbpA and rSbpA-streptag-II in a molar ratio of 7:1). The superimposition (c) allows to directly localize the recognition sites on the topography as schematically drawn in (d). Figure taken from [93]

recognition image). The possible positions of the affinity tag of a mixed (wild type and Streptag-II modified) SbpA S-layer crystal is given in Fig. 10.13d. In addition to the recognition imaging experiments, the functional activity of this *Streptag-II* rSbpA S-layer was proven by direct coupling of Strep-Tactin to S-layers formed either by rSbpA-*Streptag-II* alone or by a mixture of rSbpA-*Streptag-II* and wild-type SbpA at molecular ratio of 1:7 resulting in a high or low surface density of *Strep-Tactin* on the S-layer surface.

This study demonstrates the potential of recognition imaging on 2D nanoarrays, which are expected to play an important role as biotemplates, as platforms for molecular building blocks. Using short cross-linkers, a lateral resolution of ~ 1 nm can be achieved in recognition imaging experiments.

10.4.3 RBC Membranes

In another study, TREC was used to determine the densities of cystic fibrosis transmembrane conductance regulators (CFTRs) [46]. CFTR is a cAMP-regulated chloride channel playing an important role in salt and fluid movement across the epithelia [100, 101]. CFTR plays a key role in cystic fibrosis (CF), the most common genetic disease of Caucasians. In this disease, the predominant mutation F508del is the deletion of the phenylalanine at the position 508. The resulting protein (F508del-CFTR) is misfolded, which leads to an impaired trafficking to the plasma membrane. In earlier studies, it was shown that the absolute number of CFTR in the membrane is significantly reduced [102, 103]. Nevertheless, because of the limited lateral resolution, conclusions of the number of CFTR molecules localized in the plasma membrane have only been drawn indirectly [104–106]. Recent studies strongly indicate that this CFTR is not only present in epithelia but also in human erythrocytes [107–109]. Here, recognition imaging first allowed the direct mapping of CFTR on plasma membranes with a sufficient resolution. Erythrocytes obtained from venous blood [from healthy (non-CF) persons and persons with the homozygous F508del mutation (CF)] were spread over poly-L-lysine-coated glass. After shear opening and washing, isolated erythrocyte membranes remained on the glass surface with diameters of 5–10 μm and total heights of ~ 70 nm showing protrusions in the range of 25 nm. Before mapping the membranes, anti-CFTR antibodies were tethered to the tip using the NHS-PEG-aldehyde cross-linker (see tip Sect. 10.2). This sensor-tip was then tapped over the erythrocyte membranes, thereby recognizing the CFTRs. Possible anti-CFTR–CFTR complex formations resulted in dark spots in the recognition image. The topography of RBC membranes of non-CF (Fig. 10.14a) as well as of CF (Fig. 10.14d) patients revealed similar structures protruding out of the membranes with 10–12 nm in height, representing the membrane proteins. In these experiments, the lateral resolution was lower compared with the experiments on the bacterial S-layers due to the longer cross-linker and the softer surface. The lateral resolution in the RBC measurements was limited by the tip apex as well as by the ~ 8 -nm-long cross-linker and the

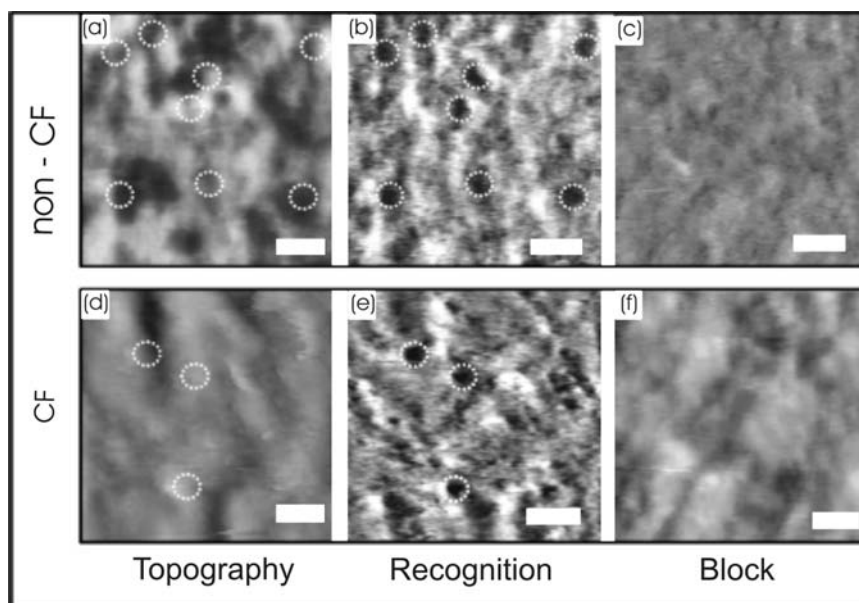


Fig. 10.14. Topography on non-CF (a) and CF (d) isolated erythrocyte membranes and corresponding recognition images in non-CF (b) and CF (e) membranes, respectively. As specificity proof, (c) and (f) represent recognition images after the addition of free anti-CFTR measurements reflecting a complete disappearance of all recognition sites. Scale bar is 200 nm, z scale 80 nm. Figure taken from [46]

bound antibody, resulting in a broadening effect of ~ 25 nm. In contrast to the topography, the simultaneously acquired recognition images showed differences between healthy (Fig. 10.14b) and CF positive (Fig. 10.14e) patients. The observed recognition spot sizes of 32 ± 5 nm (CF) and 33.5 ± 8 nm (non-CF) allowed a direct counting of the CFTRs in the membrane. Although a differentiation between single CFTRs and CFTR di- and trimers is not possible within one recognition spot, a clearly decreased number of recognition spots (1.9 ± 0.9 spots/ μm^2) was observed on CF membranes compared with non-CF membranes (6.2 ± 1.29 spots/ μm^2). Extrapolated to one erythrocyte, the number of 800 CFTRs per erythrocyte of non-CF patients and 250 for CF positive patients is somewhat higher than previously published data using quantum dots [109], proving the high resolution of recognition imaging. The specificity of the antibody CFTR complex formation was successfully proven in a control experiment by blocking the antibody CFTR interaction. This was done by adding an excess of free anti-CFTR antibodies that resulted in a complete abolishment of the recognition spots in both non-CF (Fig. 10.14c) and CF membranes (Fig. 10.14f).

In conclusion, the identification of CFTR on plasma membranes allowed a direct evaluation of the distribution of single proteins and their organization

within the heterogeneous environment of the plasma membrane. Because of the high lateral resolution of recognition imaging, the clear decrease of CFTRs in CF positive erythrocytes was determined.

10.4.4 Cells

Mapping of receptor-binding sites on cellular surfaces is a challenging task in molecular cell biology. This information can be generally obtained from the widespread use of techniques such as immunostaining (or immunocytochemistry) or some sophisticated optical techniques such as stimulated emission depletion microscopy [110], nearfield scanning optical microscopy [111], or single-molecule optical microscopy [112]. Such methods based on optical signals require either fluorophore labeling or genetic modification of cell surface proteins. The lateral resolution in these studies ranges from a few tens of nanometers to 200 nm. However, in optical studies, no information about topography is attainable. Nowadays, AFM representing a nonoptical microscopy offers an exceptional solution to obtain topography images with nanoscale resolution and single molecular interaction forces of biological specimens (e.g., proteins, DNA, membranes, cells, etc.) under or near physiological conditions and without the need for rigorous sample preparation or labeling [113]. Therefore, spatial mapping of molecular recognition sites can be obtained by performing AFM adhesion force mapping using the force–volume technique. However, dynamic recognition mapping (TREC) is faster and offers better lateral resolution than adhesion force mapping [40, 90, 92]. Because of the continuous progress in the technical aspects of the AFM and tip functionalization procedures, the investigations of receptor–ligand interactions on living cells at the single-molecule level have become possible. Because cells represent systems of more complex composition, organization, and processing in space and time than proteins, the application of the TREC technique to eukaryotic cell membranes containing functional domains enriched in (glycol) sphingolipids, cholesterol, and specific transmembrane proteins is a challenging task.

The first TREC studies on cells were performed on microvascular endothelial cells from mouse myocardium (MyEnd) to locally identify vascular endothelial (VE)-cadherin binding sites and colocalize the receptor position with membrane topographical features [94] (Fig. 10.15). VE-cadherin belongs to the widespread family of cadherins, transmembrane glycoproteins known to play an important role in calcium-dependent homophilic cell-to-cell adhesion [114]. VE-cadherin is located at intercellular junctions in essentially all types of endothelium, where VE-cadherin molecules are clustered and linked through their cytoplasmic domain to the actin cytoskeleton. The cadherin *cis*-dimer, which is formed by association of two extracellular domains in physiological Ca^{2+} -concentration (1.8 mM), represents a basic structural functional unit to promote a homophilic bond between cells [12, 114].

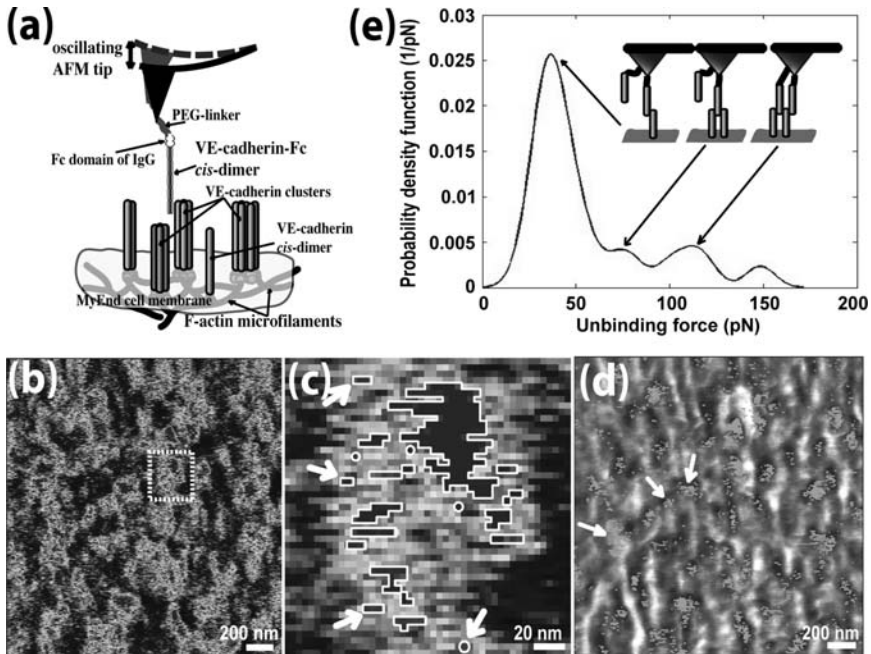


Fig. 10.15. Mapping VE-cadherin domains on vascular endothelial cell surface. (a) Scheme of dynamic recognition imaging to visualize VE-cadherin binding sites (here active single *cis*-dimers and/or clusters rich in *cis*-dimers) on a gently fixed MyEnd cell surface. (b) Recognition image of VE-cadherin domains represents an amplitude reduction due to a physical connection between VE-cadherin on the AFM tip and VE-cadherin molecules on the cell surface when specific trans-interaction occurs (shown as dark grey spots). (c) Example of recognition spots magnified from recognition image (d) Superimposition of recognition map of Vecadherin domains (in grey) onto the corresponding topography image. Color scale (dark grey to white) is 0–12 nm. (e) Corresponding force distribution recorded in Ca^{2+} -rich conditions

To avoid lateral diffusion of VE-cadherin and increase the stability of the cell membrane, cells were gently fixed with glutaraldehyde that is suitable not only to prevent the lateral mobility of receptors on the cell surface but also to maintain the cell volume and mostly preserve the filamentous structure on the cell cortex [94, 115]. TREC measurements were performed with magnetically coated AFM tips, functionalized with a recombinant VE-cadherin-Fc *cis*-dimer via a soft and long (~ 8 nm) PEG-linker (refer to Sect. 10.2 for tip chemistry). All images were taken in Hank's balanced salt solution (1.8 mM Ca^{2+}) at ambient temperature with the same experimental conditions as for model protein systems (e.g., cantilever spring constant 0.1 N/m , Q -factor ~ 1 , resonance frequency $\sim 7 \text{ kHz}$, scanning speed $\sim 3 \mu\text{m/s}$, and feedback loop coupled to the minima of the oscillations). In order to obtain the proper recognition image with high efficiencies and repeatability

(>90%), the oscillation amplitude was adjusted to be less than the extended PEG-linker length. As a consequence, a recognition signal corresponds to the amplitude reduction due to a binding between VE-cadherin molecules on the AFM tip and the cell surface when specific trans-interaction occurs (seen as dark grey spots in recognition image). These dark spots were distributed nonuniformly and correspond to microdomains with dimensions from ~ 10 to ~ 100 nm (Fig. 10.15). Figure 10.15b illustrates high recognition efficiency, which remains practically unaltered on subsequent rescans. The addition of 5 mM EDTA (Ca^{2+} -free conditions) leads to the disappearance of almost all binding events in the recognition image, whereas no change in the topography image has been observed. A closer look at some recognition spots reveals that they consist of one-to-two large domains (50–80 nm) surrounded by smaller domains (10–20 nm) or even single-molecule spots (typically 1–4 pixels long, 1 pixel ~ 4 nm) (Fig 10.15c). Taking into consideration the size of the VE-cadherin *cis*-dimer (diameter 3 nm) and the free orientation of PEG-linker leading to specific binding even before/after (8 nm) the binding site position, the dimensions of the single sites meet the expectation. More than 600 single events were recognized and ~ 6000 active *cis*-dimers were estimated over the scanned area ($4 \mu\text{m}^2$).

The shapes and positions of VE-cadherin domains were correlated to topographical features of MyEnd cell surfaces (Fig 10.15d). The topography of a scanned MyEnd cell surface area represents a complex picture of linear and branched filamentous structures, likely representing filaments of the peripheral actin belt and some globular features as well. Interestingly, a few VE-cadherin domains were found directly on top of filaments. Nevertheless, most domains were located near and between filaments, indicating that at this stage of cell maturation (day 1 or 2 after seeding), clustering of VE-cadherin was incomplete. Additionally, force curves (MRFS) were accumulated ($n = 500$, pulling velocity of 800 nm/s, encounter duration of 100 ms) before and after blocking experiment at the same scan area with the same VE-cadherin-Fc tip. The force distribution of cadherin–cadherin dissociation illustrates multiple force peaks of one-, two-, and threefold binding with a force quantum of 40 pN (Fig 10.15e). This characteristic force fingerprint was similar to an isolated VE-cadherin system [12]. The unbinding events were abolished by addition of 5 mM EDTA (reduction of binding probability from 30% to 1%). Thus, the force spectroscopy data plainly confirms that the specific domains are rich in active VE-cadherin *cis*-dimers.

This work has demonstrated a major advantage of TREC over optical approaches to cells with a spatial topographical and recognition resolution of ~ 5 nm. At present, no other microscopic techniques are able to directly provide both structural information of a biological sample and related functional information at such high spatial resolution.

Dynamic recognition imaging (i.e., TREC) and single-molecule force spectroscopy have been recently introduced as a novel way to investigate the properties of voltage-gated channels in cells [95]. Generally, the information

about the structure and function of different voltage-gated channels in living cells (including hERG K^+ channels) were gained from patch-clamp investigations. The AFM techniques have been exploited to identify a new receptor site(s) for ergtoxin-1 (ErgTx1) in voltage-sensing domain of the human ether-à-go-go-related (hERG) K^+ channel [116], with the goal to extend the understanding of the microscopic mechanism by which ErgTx1 blocks K^+ channels. The inhibition of hERG K^+ channels is the major cause of long QT syndromes inducing fatal cardiac arrhythmias. ErgTx1 belongs to scorpion-toxins, which are K^+ channel blockers, and binds to hERG channel with 1:1 stoichiometry and high affinity ($K_d \sim 10$ nM). Nevertheless, patch-clamp recordings recently demonstrated that ErgTx1 does not establish complete blockade of hERG currents, even at high ErgTx1 concentrations [117, 118]. Such phenomenon is supposed to be consistent with highly dynamic conformational changes of the outer pore domain of hERG. Thus, TREC has been exploited to locally identify extracellular binding sites of hERG K^+ channels on gently fixed hERG HEK-293 cells. Measurements were started by scanning of the whole cell surface with subsequent zooming into small areas $2 \times 2 \mu\text{m}^2$. TREC images were acquired with MAC tips coated with antibody anti-Kv11.1 (against epitope tags present on the hERG subunits) via PEG-linker (Fig. 10.16a). The recognition map represents an amplitude reduction due to specific binding between anti-Kv11.1 on the tip and epitope tags on the cell surface (dark spots) (Fig. 10.16b). These dark spots (amplitude reduction up to 2 nm) are distributed nonuniformly (Fig. 10.16c) and reflect microdomains with dimensions from ~ 30 up to ~ 350 nm, with a mean \pm SD of 99 ± 81 nm ($n = 25$) for the long domain axis. During several subsequent rescans, recognition maps of hERG channels remain unchanged. Then, ErgTx1 was very slowly ($50 \mu\text{l}/\text{min}$) injected into the fluid cell while scanning the same sample. After the first and second injection of ErgTx1 (concentration of ~ 400 nM), no visual changes in the recognition maps have been observed. However, the recognition clusters partly disappeared after the third injection of ErgTx1 ($\sim 1 \mu\text{M}$), whereas no change in the topography image has been observed (Fig. 10.16d). The specific binding between anti-Kv11.1 and cellular surface was abolished when free ErgTx1 molecules bound to the hERG channels and thus blocked the antibody access to interact with epitope tags on hERG subunits. The topography of a scanned cell surface areas shows a complex picture of linear and branched filamentous structures with some globular features. Most domains are found to be located near or between filaments (Fig. 10.16c). TREC results suggest that ErgTx1 does not only interact with the extracellular surface of the pore domain (S5–S6), but might interact with the voltage sensing domains (S1–S4) of the hERG K^+ channel.

In addition to these TREC measurements, AFM force-distance cycles with a tip carrying an epitope-specific antibody (anti-Kv11.1) were performed on living as well as on gently fixed hERG HEK-293 cells. Both studies on living and fixed cells conducted to similar results (force distributions and binding probabilities). The anti-Kv11.1 (hERG)-extracellular antibody is known to

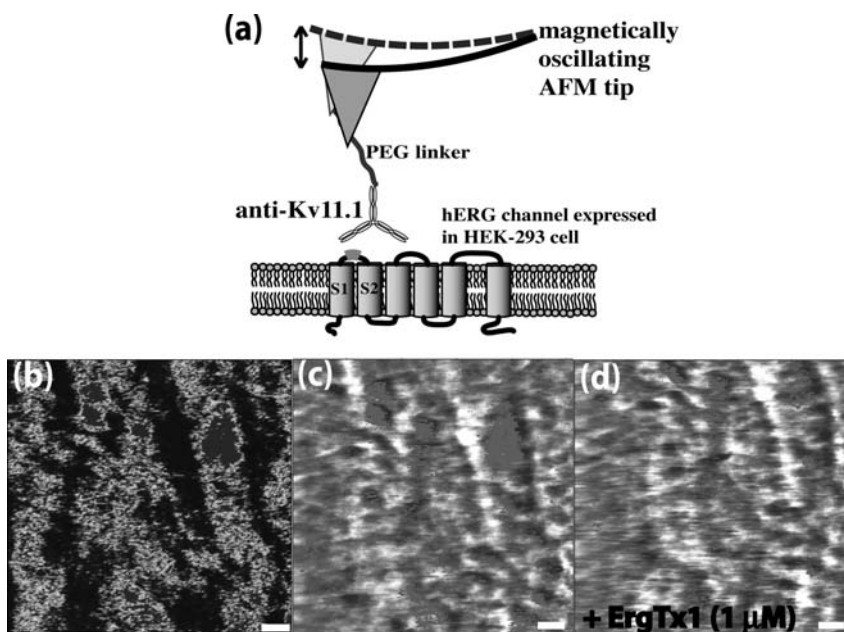


Fig. 10.16. Dynamic recognition images on a hERG HEK-293 cell surface. (a) Schematic representation of recognition imaging to visualize hERG K⁺ channels [here binding sites of extracellular epitope between S1 and S2 domains of hERG subunit (*shown in light gray*)] on a gently fixed hERG HEK-293 cell surface. (b) TREC maps obtained on a hERG HEK-293 cell surface with anti-Kv11.1 – functionalized tip. The most pronounced recognition clusters are indicated by white lines. (c) Correlation between recognition maps (in grey) and the corresponding simultaneously recorded topography images. Color scale (dark grey to white) is 0–12 nm. (d) The presence of 400 nM ErgTx1 had practically no effect on recognition map (data not shown), whereas the recognition clusters partly disappeared after the third injection of ErgTx1 (~1 μ M). Blocking experiments do not affect membrane topography. Scale bars on all images are 170 nm

bind to the voltage sensor domain (S1–S2 region) of HERG K⁺ channel. The specific binding of the antibody to the extracellular part of hERG channel was characterized by a unique unbinding force. To confirm the specificity of this binding, blocking experiments were carried out by injecting either free antibodies or free peptide antigens. In both cases, almost no unbinding events were observed. Binding probabilities (probability to record an unbinding event in force-distance cycles) from several experiments were also quantified. The binding probability of ~30% was calculated for the interaction between anti-Kv11.1-extracellular antibody and hERG HEK-293 cells (Fig. 10.17a). When free anti-Kv11.1 antibodies or free peptide antigens were present in solution, the binding probability drastically decreased to the level of ~2% (Fig. 10.17a). In an empirical probability density function (pdf) of the unbinding forces, the maximum of the distribution was found to be 45 ± 9 pN (Fig. 10.17b). As

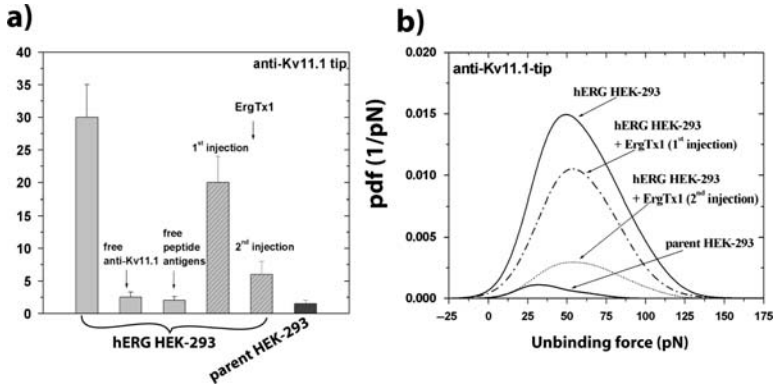


Fig. 10.17. Recognition of hERG K^+ channels on the surface of intact cells by an AFM tip carrying anti-Kv11.1. (a) Quantitative comparison of binding probabilities of anti-Kv11.1-coated tip on live hERG HEK-293 cells (*left light gray*) in the absence and presence of either free anti-Kv11.1 (*middle light gray*) or free antigen peptides (*right light gray*); binding probability of anti-Kv11.1-tip on parent HEK-293 cells is shown in *black*. The effect of ErgTx1 on the binding probabilities anti-Kv11.1-tip on hERG HEK-293 (*gray*). Values are mean \pm SEM, $n = 2,000\text{--}4,000$. (b) Probability density functions (*pdf*) observed in the absence (*solid lines*) and in the presence of ErgTx1 [*dot-dashed* (300 nM) and *dashed lines* (1 μ M)]. Areas are scaled to the corresponding binding probabilities

another indicator of the specificity, a very low binding probability ($\sim 1.5\%$) with a force peak of ~ 25 pN were found for the parent HEK-293 cells not expressing hERG K^+ channels. These results illustrate that the extracellular part of hERG K^+ channels expressed in living cells can be specifically detected at the molecular level by using epitope-specific antibodies. The eventual effects on antibody binding induced by ErgTx1 were additionally investigated. Force curves were accumulated before and after ErgTx1 multiple injections at the same scan area with the same functionalized tip. In the presence of ErgTx1 at different concentrations, the peak force for force distributions remained, whereas the binding probability between antibody and living hERG HEK 293 cells decreased dramatically after multiple ErgTx1 injections. These findings indicate additionally a possibly new binding site of ErgTx1 in the voltage sensor domain of hERG K^+ channel.

Thus, it has been demonstrated that the combination of dynamic molecular recognition imaging (TREC) with single molecular force spectroscopy is a suitable method to obtain information about the structure and function of hERG K^+ channels on the cellular surface. Both methods exploit AFM tips with a very low surface density of ligands (~ 400 molecules per μm^2) so as to allow detection of single molecular events. Functionalization of AFM tips with anti-Kv11.1 (hERG)-extracellular antibody enabled them to detect binding sites of hERG (in S1–S2 domains) on the surface of cells expressed hERG channels.

The main outcome of this study reveals that voltage sensing domain (S1-S4) of hERG K⁺ channel might be one of the binding sites of ErgTx1.

This chapter illustrates the great potential of TREC for the investigation and localization of membrane proteins on the cell surface with several piconewtons force resolution and a few nanometers positional accuracy.

10.5 Conclusion

TREC has evolved into an imaging technique that yields fine structural details together with the biochemical composition of native biological samples as well as artificial, nano-technological engineered surfaces. The major advantage of TREC over optical approaches to cells is its spatial topographical and recognition resolution of 1–5 nm. Dynamic recognition imaging allows to detect single molecular interactions, and thus to visualize, identify, and quantify local receptor-binding sites and assign their locations to the topographical features of cell surface. TREC shows a high potential to be used for many adherent cells or extracted cellular membranes to locally identify receptor binding sites. On top of that, TREC is a promising tool for the localization of functional groups on various nano-technological devices based on biological building blocks.

Reliable protocols are available for attaching biomolecules or cells on the AFM tips and on supporting surfaces. Also, procedures to optimize the imaging parameters for TREC are now well established and should make measurements more easy, robust, accurate, and reliable. Nevertheless, it is fair to say that accurate data collection and interpretation often remain delicate and require strong expertise, especially when dealing with complex specimens like living cells. The main tasks are those associated with the quality of tip and support surface chemistries and with their possible alteration during data acquisition. Thus, a detailed understanding of the principles of TREC and its limitations is essential before users start their first experiment.

Acknowledgements

This work was supported by the European community projects BioLightTouch (FP6-2004-NEST-C-1-028781), Imunanomap (MRTN-CT-2006-035946), and SMW (MNT-2007-213717), the Austrian Science Foundation (F 3503-B11), and the Austrian Christian Doppler Society.

References

1. G. Binnig, C.F. Quate, C. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).
2. B.P. Jena, S.J. Cho, *Methods Cell. Biol.* **68**, 33 (2002).
3. E.L. Florin, V.T. Moy, H.E. Gaub, *Science* **264**, 415 (1994).

4. G.U. Lee, A.C. Chrisey, R.J. Colton, *Langmuir* **10**, 354 (1994).
5. P. Hinterdorfer, W. Baumgartner, H.J. Gruber, K. Schilcher, H. Schindler, *Proc. Natl. Acad. Sci. USA* **93**, 3477 (1996).
6. A.F. Oberhauser, P.E. Marszalek, H.P. Erickson, J.M. Fernandez, *Nature* **393**, 181 (1998).
7. E. Evans, K. Ritchie, *Biophys. J.* **72**, 1541 (1997).
8. J. Fritz, A.G. Katopodis, F. Kolbinger, D. Anselmetti, *Proc. Natl. Acad. Sci. USA* **95**, 12283 (1998).
9. R. Ros, F. Schwesinger, D. Anselmetti, M. Kubon, R. Schafer, A. Pluckthun, L. Tiefenauer, *Proc. Natl. Acad. Sci. USA* **95**, 7402 (1998).
10. T. Strunz, K. Oroszlan, R. Schafer, H.J. Guntherodt, *Proc. Natl. Acad. Sci. USA* **96**, 11277 (1999).
11. F. Schwesinger, R. Ros, T. Strunz, D. Anselmetti, H.J. Guntherodt, A. Honegger, L. Jermutus, L. Tiefenauer, A. Pluckthun, *Proc. Natl. Acad. Sci. USA* **97**, 9972 (2000).
12. W. Baumgartner, P. Hinterdorfer, W. Ness, A. Raab, D. Vestweber, H. Schindler, D. Drenckhahn, *Proc. Natl. Acad. Sci. USA* **97**, 4005 (2000).
13. F. Kienberger, G. Kada, H.J. Gruber, V. Pastushenko, C. Riener, M. Trieb, H.-G. Knaus, H. Schindler, P. Hinterdorfer, *Single Mol.* **1**, 59 (2000).
14. C. Yuan, A. Chen, P. Kolb, V.T. Moy, *Biochemistry* **39**, 10219 (2000).
15. P. Hinterdorfer, K. Schilcher, W. Baumgartner, H.J. Gruber, H. Schindler, *Nanobiol.* **4**, 177 (1998).
16. V.T. Moy, E.L. Florin, H.E. Gaub, *Science* **266**, 257 (1994).
17. A. Chilkoti, T. Boland, B.D. Ratner, P.S. Stayton, *Biophys. J.* **69**, 2125 (1995).
18. R. Merkel, P. Nassoy, A. Leung, K. Ritchie, E. Evans, *Nature* **397**, 50 (1999).
19. M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, *Science* **276**, 1109 (1997).
20. P.E. Marszalek, A.F. Oberhauser, Y.P. Pang, J.M. Fernandez, *Nature* **396**, 661 (1998).
21. F. Kienberger, V.P. Pastushenko, G. Kada, H.J. Gruber, C. Riener, H. Schindler, P. Hinterdorfer, *Single Mole.* **1**, 123 (2000).
22. J. Preiner, H. Janovjak, C. Rankl, H. Knaus, D.A. Cisneros, A. Kedrov, F. Kienberger, D.J. Muller, P. Hinterdorfer, *Biophys. J.* **93**, 930 (2007).
23. G. Hummer, A. Szabo, *Proc. Natl. Acad. Sci. USA* **98**, 3658 (2001).
24. H.A. Mizes, K.G. Loh, R.J.D. Miller, S.K. Ahuja, E.F. Grabowski, *Appl. Phys. Lett.* **59**, 2901 (1991).
25. M. Radmacher, J.P. Cleveland, M. Fritz, H.G. Hansma, P.K. Hansma, *Biophys. J.* **66**, 2159 (1994).
26. C.A.J. Putman, K.O. Vanderwerf, B.G. Degrooth, N.F. Vanhulst, J. Greve, *Appl. Phys. Lett.* **64**, 2454 (1994).
27. M. Ludwig, W. Dettmann, H.E. Gaub, *Biophys. J.* **72**, 445 (1997).
28. O.H. Willemsen, M.M. Snel, K.O. van der Werf, B.G. de Groot, J. Greve, P. Hinterdorfer, H.J. Gruber, H. Schindler, Y. van Kooyk, C.G. Figdor, *Biophys. J.* **75**, 2220 (1998).
29. M. Grandbois, W. Dettmann, M. Benoit, H.E. Gaub, *J. Histochem. Cytochem.* **48**, 719 (2000).
30. P.P. Lehenkari, G.T. Charras, A. Nykanen, M.A. Horton, *Ultramicroscopy* **82**, 289 (2000).
31. N. Almqvist, R. Bhatia, G. Primbs, N. Desai, S. Banerjee, R. Lal, *Biophys. J.* **86**, 1753 (2004).

32. M.B. Viani, T.E. Schaffer, A. Chand, M. Rief, H.E. Gaub, P.K. Hansma, J. Appl. Phys. **86**, 2258 (1999).
33. T.E. Schaffer, Y.K. Jiao, Biophys. J. **80**, 303A (2001).
34. A. Raab, W. Han, D. Badt, S.J. Smith-Gill, S.M. Lindsay, H. Schindler, P. Hinterdorfer, Nat. Biotechnol. **17**, 901 (1999).
35. H. Schindler, D. Badt, P. Hinterdorfer, F. Kienberger, A. Raab, S. Wielert-Badt, V. Pastushenko, Ultramicroscopy **82**, 227 (2000).
36. W.H. Han, S.M. Lindsay, T.W. Jing, Appl. Phys. Lett. **69**, 4111 (1996).
37. W. Han, M. Dlakic, Y.J. Zhu, S.M. Lindsay, R.E. Harrington, Proc. Natl. Acad. Sci. USA **94**, 10565 (1997).
38. W. Han, S.M. Lindsay, M. Dlakic, R.E. Harrington, Nature **386**, 563 (1997).
39. F. Kienberger, C. Stroh, G. Kada, R. Moser, W. Baumgartner, V. Pastushenko, C. Rankl, U. Schmidt, H. Muller, E. Orlova, C. LeGrimellec, D. Drenckhahn, D. Blaas, P. Hinterdorfer, Ultramicroscopy **97**, 229 (2003).
40. C.M. Stroh, A. Ebner, M. Geretschlager, G. Freudenthaler, F. Kienberger, A.S.M. Kamruzzahan, S.J. Smith-Gill, H.J. Gruber, P. Hinterdorfer, Biophys. J. **87**, 1981 (2004).
41. A. Ebner, F. Kienberger, G. Kada, C.M. Stroh, M. Geretschlager, A.S. Kamruzzahan, L. Wildling, W.T. Johnson, B. Ashcroft, J. Nelson, S.M. Lindsay, H.J. Gruber, P. Hinterdorfer, Chem. Phys. Chem. **6**, 897 (2005).
42. C.M. Stroh, A. Ebner, M. Geretschlager, G. Freudenthaler, F. Kienberger, A.S. Kamruzzahan, S.J. Smith-Gill, H.J. Gruber, P. Hinterdorfer, Biophys. J. **87**, 1981 (2004).
43. J. Preiner, N. Losilla, A. Ebner, P. Annibale, F. Biscarini, R. Garcia, P. Hinterdorfer, Nano Lett. **9**, 571 (2009).
44. C. Stroh, H. Wang, R. Bash, B. Ashcroft, J. Nelson, H. Gruber, D. Lohr, S.M. Lindsay, P. Hinterdorfer, Proc. Natl. Acad. Sci. USA **101**, 12503 (2004).
45. J. Tang, A. Ebner, H. Badelt-Lichtblau, C. Vollenkle, C. Rankl, B. Kraxberger, M. Leitner, L. Wildling, H.J. Gruber, U.B. Sleytr, Nano Lett. **8**, 4312 (2008).
46. A. Ebner, D. Nikova, T. Lange, J. Haberle, S. Falk, A. Dubbers, R. Bruns, P. Hinterdorfer, H. Oberleithner, H. Schillers, Nanotechnology **19** (2008).
47. L. Chtcheglova, J. Waschke, L. Wildling, D. Drenckhahn, P. Hinterdorfer, Biophys. J. (2007).
48. L. Chtcheglova, F. Atalar, U. Ozbek, L. Wildling, A. Ebner, P. Hinterdorfer, Pflügers Archiv Eur. J. Physiology (2008).
49. S. Lee, J. Mandic, K.V. Vliet, Proc. Natl. Acad. Sci. **104**, 9609 (2007).
50. K.C. Langry, T.V. Ratto, R.E. Rudd, M.W. McElfresh, Langmuir **21**, 12064 (2005).
51. T.V. Ratto, K.C. Langry, R.E. Rudd, R.L. Balhorn, M.J. Allen, M.W. McElfresh, Biophys. J. **86**, 2430 (2004).
52. B. Cross, F. Ronzon, B. Roux, J.P. Rieu, Langmuir **21**, 5149 (2005).
53. P. Desmeules, M. Grandbois, V.A. Bondarenko, A. Yamazaki, C. Salesse, Biophys. J. **82**, 3343 (2002).
54. A. Touhami, B. Hoffmann, A. Vasella, F.A. Denis, Y.F. Dufrene, Langmuir **19**, 1745 (2003).
55. A. Touhami, M.H. Jericho, T.J. Beveridge, Langmuir **23**, 2755 (2007).
56. L. Schmitt, M. Ludwig, H.E. Gaub, R. Tampe, Biophys. J. **78**, 3275 (2000).
57. M. Conti, G. Falini, B. Samori, Angewandte Chemie **112** (2000).

58. A. Ebner, F. Kienberger, G. Kada, C.M. Stroh, M. Geretschlager, A.S.M. Kamruzzahan, L. Wildling, W.T. Johnson, B. Ashcroft, J. Nelson, S.M. Lindsay, H.J. Gruber, P. Hinterdorfer, *Chem. Phys. Chem.* **6**, 897 (2005).
59. C.K. Riener, C.M. Stroh, A. Ebner, C. Klampfl, A.A. Gall, C. Romanin, Y.L. Lyubchenko, P. Hinterdorfer, H.J. Gruber, *Analytica Chimica Acta* **479**, 59 (2003).
60. R. Ros, F. Schwesinger, D. Anselmetti, M. Kubon, R. Schäfer, A. Plückthun, L. Tiefenauer, *Proc. Natl. Acad. Sci. USA*, **95**, 7402.
61. A.S.M. Kamruzzahan, A. Ebner, L. Wildling, F. Kienberger, C.K. Riener, C.D. Hahn, P.D. Pollheimer, P. Winklehner, M. Holzl, B. Lackner, D.M. Schorkl, P. Hinterdorfer, H.J. Gruber, *Bioconjug. Chem.* **17**, 1473 (2006).
62. T. Haselgrubler, A. Amerstorfer, H. Schindler, H.J. Gruber, *Bioconjug. Chem.* **6**, 242 (1995).
63. S. Wielert-Badt, P. Hinterdorfer, H.J. Gruber, J.-T. Lin, D. Badt, B. Wimmer, H. Schindler, R.K.-H. Kinne, *Biophys. J.* **82**, 2767 (2002).
64. T. Puntheeranurak, L. Wildling, H.J. Gruber, R.K.H. Kinne, P. Hinterdorfer, *J. Cell Sci.* **119**, 2960 (2006).
65. T. Puntheeranurak, B. Wimmer, F. Castaneda, H.J. Gruber, P. Hinterdorfer, R.K.H. Kinne, *Biochemistry* **46**, 2797 (2007).
66. M. Morpurgo, F.M. Veronese, D. Kachensky, J.M. Harris, *Bioconjug. Chem.* **7**, 363 (1996).
67. W. Grange, T. Strunz, I. Schumakovitch, H.J. Guntherodt, M. Hegner, *Single Mole.* **2**, 75 (2001).
68. L. Lin, H. Wang, Y. Liu, H. Yan, S. Lindsay, *Biophys. J.* **90**, 4236 (2006).
69. T. Strunz, K. Oroszlan, I. Schumakovitch, H. Guntherodt, M. Hegner, *Biophys. J.* **79**, 1206 (2000).
70. H.G. Linde, *J. Appl. Polym. Sci.* **40**, 613 (1990).
71. R. Nevo, C. Stroh, F. Kienberger, D. Kaftan, V. Brumfeld, M. Elbaum, Z. Reich, P. Hinterdorfer, *Nat. Struc. Biol.* **10**, 553 (2003).
72. C.K. Riener, F. Kienberger, C.D. Hahn, G.M. Buchinger, I.O.C. Egwim, T. Haselgrubler, A. Ebner, C. Romanin, C. Klampfl, B. Lackner, *Analytica Chimica Acta* **497**, 101 (2003).
73. B. Bonanni, A.S.M. Kamruzzahan, A.R. Bizzarri, C. Rankl, H.J. Gruber, P. Hinterdorfer, S. Cannistraro, *Biophys. J.* **89**, 2783 (2005).
74. A. Ebner, L. Wildling, A.S.M. Kamruzzahan, C. Rankl, J. Wruss, C.D. Hahn, M. Holzl, R. Zhu, F. Kienberger, D. Blaas, P. Hinterdorfer, H.J. Gruber, *Bioconjug. Chem.* **18**, 1176 (2007).
75. E.J. Hukkanen, J.A. Wieland, D.E. Leckband, R.D. Braatz, **2005**, pp. 3265.
76. J.A. Wieland, A.A. Gewirth, D.E. Leckband, *J. Biologic. Chem.* **280**, 41037 (2005).
77. A. Ebner, P. Hinterdorfer, H.J. Gruber, *Ultramicroscopy* **107**, 922 (2007).
78. R. Avci, M. Schweitzer, R.D. Boyd, J. Wittmeyer, A. Steele, J. Toporski, W. Beech, F.T. Arce, B. Spangler, K.M. Cole, D.S. McKay, *Langmuir* **20**, 11053 (2004).
79. R. Horn, S. Ding, H.J. Gruber, *J. General Physiol.* **116**, 461 (2000).
80. K.L. Prime, G.M. Whitesides, 252, 1164 (1991).
81. C. Verbelen, H.J. Gruber, Y.F. Dufrene, *J. Molecul. Recog.* **20** (2007).
82. J.L. Tang, A. Ebner, N. Ilk, H. Lichtblau, C. Huber, R. Zhu, D. Pum, M. Leitner, V. Pastushenko, H.J. Gruber, U.B. Sleytr, P. Hinterdorfer, *Langmuir* **24**, 1324 (2008).

83. L. Wildling, P. Hinterdorfer, K. Kusche-Vihrog, Y. Treffner, H. Oberleithner, *Pflugers Arch.* 2008, in press.
84. W. Han, S.M. Lindsay, T. Jing, *Appl. Phys. Lett.* **69**, 4111 (1996).
85. A. Ebner, L. Wildling, R. Zhu, C. Rankl, T. Haselgrubler, P. Hinterdorfer, H.J. Gruber, *Stm And Afm Studies On (Bio)Molecular Systems: Unravelling The Nanoworld* **285**, 29 (2008).
86. J. Preiner, J.L. Tang, V. Pastushenko, P. Hinterdorfer, *Phys. Rev. Lett.* **99** (2007).
87. Q. Zhong, D. Inniss, K. Kjoller, V.B. Elings, *Surf. Sci.* **290**, L688 (1993).
88. Y. Martin, C.C. Williams, H.K. Wickramasinghe, *J. Appl. Phys.* **61**, 4723 (1987).
89. M. Lantz, Y.Z. Liu, X.D. Cui, H. Tokumoto, S.M. Lindsay, *Interface Anal.* **27**, 354 (1999).
90. C. Stroh, H. Wang, R. Bash, B. Ashcroft, J. Nelson, H. Gruber, D. Lohr, S.M. Lindsay, P. Hinterdorfer, *Proc. Natl. Acad. Sci. USA* **101**, 12503 (2004).
91. J. Preiner, N. Losilla, A. Ebner, P. Annibale, F. Biscarini, R. Garcia, P. Hinterdorfer, *Nano Lett.* **9**, 571 (2009).
92. P. Hinterdorfer, Y.F. Dufrene, *Nat. Method.* **3**, 347 (2006).
93. J.L. Tang, A. Ebner, H. Badelt-Lichtblau, C. Vollenkle, C. Rankl, B. Kraxberger, M. Leitner, L. Wildling, H.J. Gruber, U.B. Sleytr, N. Ilk, P. Hinterdorfer, *Nano Lett.* **8**, 4312 (2008).
94. L.A. Chtcheglova, J. Waschke, L. Wildling, D. Drenckhahn, P. Hinterdorfer, *Biophys. J.* **93**, L11 (2007).
95. L.A. Chtcheglova, F. Atalar, U. Ozbek, L. Wildling, A. Ebner, P. Hinterdorfer, *Pflugers Archiv-Eur. J. Physiol.* **456**, 247 (2008).
96. U.B. Sleytr, P. Messner, D. Pum, M. Sara, *Angewandte Chemie – Int. Ed.* **38**, 1035 (1999).
97. U.B. Sleytr, M. Sara, D. Pum, B. Schuster, *Prog. Surface Sci.* **68**, 231 (2001).
98. U.B. Sleytr, C. Huber, N. Ilk, D. Pum, B. Schuster, E.M. Egelseer, *Fems. Microbiol. Lett.* **267**, 131 (2007).
99. U.B. Sleytr, E.M. Egelseer, N. Ilk, D. Pum, B. Schuster, *Febs J.* **274**, 323 (2007).
100. E.M. Schwiebert, D.J. Benos, C.M. Fuller, *Am. J. Med.* **104**, 576 (1998).
101. G.M. Braunstein, R.M. Roman, J.P. Clancy, B.A. Kudlow, A.L. Taylor, V.G. Shylonsky, B. Jovov, K. Peter, T. Jilling, Ismailov, II, D.J. Benos, L.M. Schwiebert, J.G. Fitz, E.M. Schwiebert, *J. Biologic. Chem.* **276**, 6621 (2001).
102. N. Kalin, A. Claass, M. Sommer, E. Puchelle, B. Tummler, *J. Clin. Invest.* **103**, 1379 (1999).
103. F. Dupuit, N. Kalin, S. Brezillon, J. Hinnrasky, B. Tummler, E. Puchelle, *J. Clin. Invest.* **96**, 1601 (1995).
104. C. Haws, W.E. Finkbeiner, J.H. Widdicombe, J.J. Wine, *Am. J. Physiol.* **266**, L502 (1994).
105. M.J. Hug, I.E. Thiele, R. Greger, *Pflugers Archiv-Eur. J. Physiol.* **434**, 779 (1997).
106. D. Penque, F. Mendes, S. Beck, C. Farinha, P. Pacheco, P. Nogueira, J. Lavinha, R. Malho, M.D. Amaral, *Lab. Invest.* **80**, 857 (2000).
107. K.M. Sterling, S. Shah, R.J. Kim, N.I.F. Johnston, A.Y. Salikhova, E.H. Abraham, *J. Cell. Biochem.* **91**, 1174 (2004).
108. R.S. Sprague, M.L. Ellsworth, A.H. Stephenson, M.E. Kleinhenz, A.J. Lonigro, *Am. J. Physiol.-Heart Circulat. Physiol.* **275**, H1726 (1998).

109. T. Lange, P. Jungmann, J. Haberle, S. Falk, A. Duebbers, R. Bruns, A. Ebner, P. Hinterdorfer, H. Oberleithner, H. Schillers, *Mole. Memb. Biol.* **23**, 317 (2006).
110. K.I. Willig, S.O. Rizzoli, V. Westphal, R. Jahn, S.W. Hell, *Nature* **440**, 935 (2006).
111. M. Koopman, A. Cambi, B.I. de Bakker, B. Joosten, C.G. Figdor, N.F. van Hulst, M.F. Garcia-Parajo, *Febs Lett.* **573**, 6 (2004).
112. T. Schmidt, G.J. Schutz, W. Baumgartner, H.J. Gruber, H. Schindler, *Proc. Natl. Acad. Sci. USA* **93**, 2926 (1996).
113. J.K. Horber, M.J. Miles, *Science* **302**, 1002 (2003).
114. P.A. Vincent, K.Y. Xiao, K.M. Buckley, A.P. Kowalczyk, *Am. J. Physiol.-Cell Physiol.* **286**, C987 (2004).
115. H. Oberleithner, S.W. Schneider, L. Albermann, U. Hillebrand, T. Ludwig, C. Riethmüller, V. Shahin, C. Schäfer, H. Schillers, *J. Memb. Biol.* **196**, 163 (2003).
116. G.B. Gurrola, B. Rosati, M. Rocchetti, G. Pimienta, A. Zaza, A. Arcangeli, M. Olivotto, L.D. Possani, E. Wanke, *Faseb J.* **13**, 953 (1999).
117. L. Pardo-Lopez, M. Zhang, J. Liu, M. Jiang, L.D. Possani, G.N. Tseng, *J. Biologic. Chem.* **277**, 16403 (2002).
118. A.M. Torres, P. Bansal, P.F. Alewood, J.A. Bursill, P.W. Kuchel, J.I. Vandenberg, *Febs Lett.* **539**, 138 (2003).

Structural and Mechanical Mechanisms of Ocular Tissues Probed by AFM

Noël M. Ziebarth, Felix Rico, and Vincent T. Moy

Summary. In recent years, the atomic force microscope (AFM) has become an important tool in ophthalmic research. It has gained popularity largely because AFM is not restricted by the diffraction limits of light microscopy and can be applied to resolve images with molecular resolution. AFM is a minimally invasive technique and can be used to visualize molecular structures under near-physiological conditions. In addition, the AFM can be employed as a force apparatus to characterize the viscoelastic properties of biomaterials on the micron level and at the level of individual proteins. In this article, we summarize recent AFM studies of ocular tissues, while highlighting the great potential of AFM technology in ophthalmic research. Previous research demonstrates the versatility of the AFM as high resolution imaging technique and as a sensitive force apparatus for probing the mechanical properties of ocular tissues. The structural and mechanical properties of ocular tissues are of major importance to the understanding of the optomechanical functions of the human eye. In addition, AFM has played an important role in the development and characterization of ocular biomaterials, such as contact lenses and intraocular lenses. Studying ocular tissues using Atomic Force Microscopy has enabled several advances in ophthalmic research.

Key words: Atomic force microscopy, Nanoindentation, Ophthalmology.

11.1 Introduction

In recent years, the atomic force microscope (AFM) has become an important tool in ophthalmic research. It has gained popularity in biological applications largely because AFM is not restricted by the diffraction limits of light microscopy and can be applied to resolve images of biological samples with molecular resolution. Moreover, unlike electronic microscopy (EM), AFM is a minimally invasive technique and can be used to visualize molecular structures within biological specimens under near-physiological conditions [1, 2]. In addition, the AFM can be employed as a force apparatus to characterize the viscoelastic properties of biomaterials on the micron level and at the level

of individual proteins [3]. This chapter is divided into two sections. The first section will provide an overview of AFM technology in the studies of soft biomaterials. This is followed by a review of recent advances in ophthalmology stemming from AFM research.

The structural and mechanical properties of ocular tissues are of major importance to the understanding of the optomechanical functions of the human eye. The optics of the human eye serve two primary functions: (1) to focus incoming light onto the retina and (2) to regulate the amount of light reaching the photoreceptors [4, 5]. Focusing is achieved with a positive double lens arrangement that includes the cornea and the lens. The cornea is a relatively rigid structure and is responsible for 75% of refractive power of the eye. Light crossing the lens is further bent on its way to the retina. The lens is a highly elastic structure that is suspended by zonule fibers attached to the ciliary muscles. The ability of the ciliary muscle to control the curvature of the lens permits the eye to adjust its focus on objects located at different distances.

The main structural components of the crystalline lens are the inner lens fibers and the surrounding lens capsule. The lens fibers form the bulk of the lens. They are formed by concentric tightly packed layers of long, thin, transparent cells, much like the layers of an onion. The lens capsule is a thick basement membrane (4–20 μm) [6] that encapsulates the lens fibers. The released tension on the lens capsule during accommodation enables it to contract, thus allowing the lens to relax and thicken. This change in lens shape is characterized by an increase in the anterior and posterior curvatures of the lens, thereby increasing its refractive power.

Changes in structural and mechanical properties have been implicated as the underlying mechanism for loss of accommodation as a person ages [7–9]. This condition is referred to as presbyopia when it leads to the loss of near vision [10, 11]. Probable causes of presbyopia are multifactorial. There is strong evidence to suggest that it is caused by an inability of the lens to change shape when the ciliary muscle changes zonular tension [7].

11.2 Atomic Force Microscopy

11.2.1 Principle of Operation

Overview

The AFM was originally designed to characterize the topography of hard material surfaces [12]. It was rapidly improved to enable operation in liquids, which made characterization of biological samples possible. The AFM makes use of a sharp tip at the end of a flexible cantilever to scan and probe the sample surface. The cantilever, and in particular the cantilever tip, is the most relevant part of the AFM because it interacts with the sample and

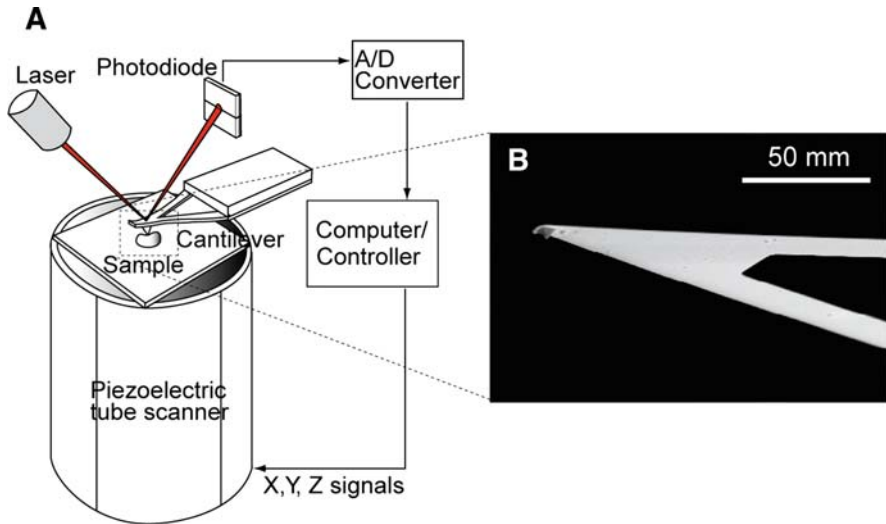


Fig. 11.1. (a) Principle of atomic force microscope (AFM) operation. (b) Close-up of AFM cantilever

determines the maximum resolution obtainable (Fig. 11.1). The interaction forces between the tip and the sample are usually determined by monitoring the deflection of the cantilever, modeled as a very sensitive spring with known stiffness. The deflection, and thus the force, is usually monitored by focusing a laser beam on the back side of the cantilever and recording changes in the reflected beam angle using a segmented photodiode. Various imaging modes have been developed during the last two decades, the most popular being *contact mode*, *intermittent contact* or “*tapping*” mode, and *force mapping* or “*force volume*” mode. Force mapping mode takes advantage of the capacity of the AFM to obtain mechanical measurements of samples by the acquisition of force–displacement (F – z) curves. These curves are obtained by oscillating the AFM cantilever in the vertical direction, perpendicular to the sample surface, and monitoring its deflection as it interacts with the sample surface. Both imaging and force modes have been extensively used to investigate the topography, adhesion, and viscoelastic properties of many biological samples [13–64],¹ including ophthalmic tissues [65–88].

¹ A special issue entitled *Atomic force microscopy enters physiology* appeared recently on Pflügers Archive European Journal of Physiology (Volume 456, Number 1, April 2008). A more general issue dedicated to *Nanotechnology at the interface of cell biology, materials science and medicine* was also recently published on Nanotechnology (Volume 19, Number 38, September 2008) with various articles focused on AFM. In 2007 an issue dedicated to *Atomic force microscopy in life sciences and medicine* appeared in the Journal of Molecular Recognition (Volume 20, Issue 6, December 2007).

Imaging Mode

AFM images are three-dimensional topographical representations of a sample surface. In imaging mode, the cantilever tip is scanned along the sample surface while keeping the interaction between the tip and the sample constant. This is accomplished by using a feedback loop that continuously corrects the relative position of the cantilever and the sample.

In contact mode, the most widely used AFM mode, the cantilever tip is kept in continuous contact with the sample surface, maintaining constant cantilever deflection while scanning. The position corrections applied by the piezoelectric mechanism to keep this deflection constant constitute the actual three-dimensional topography of the sample surface, also known as the height image. Contact mode enables the acquisition of AFM images with nanometer resolution in the horizontal plane and subnanometer resolution in the vertical axis; however, the lateral shear forces exerted on the sample can be quite high when dealing with high aspect ratio samples. A gentler imaging mode is the intermittent contact or “tapping” mode, in which the cantilever is oscillated at its resonant frequency. In this mode, the amplitude of oscillation of the free end of the cantilever is the control parameter. As a freely oscillating cantilever tip approaches the sample surface, its amplitude of oscillation decreases before contact is made. This behavior is due to a thin layer of water that permanently covers surfaces under standard conditions as well as to drag forces and electrostatic, steric, or van der Waals interactions between the tip and the sample in liquid. To keep the cantilever amplitude of oscillation constant during scanning of the sample, the position of the cantilever relative to the sample is continuously corrected, which translates into a three-dimensional replica of the sample surface. Although intermittent contact does not produce AFM images with as high resolution as contact mode, lateral drag forces that may damage the sample are considerably reduced [46, 89–94].

Force Mode

A less generalized AFM mode, force mode, enables the investigation of the mechanical properties of the sample as well as the interaction forces between the cantilever tip and the sample surface (refer to the works by Cappella et al. [95] and Butt et al. [18] for excellent, extensive reviews on force mode AFM). This mode consists of monitoring cantilever deflection (i.e., force), whereas the tip approaches or retracts from the sample surface at constant velocity. The resulting force–displacement curve provides information about long-range interaction forces, contact elastic forces, and/or adhesion forces between the tip and the sample. This mode has been used to better understand the underlying theory behind high-resolution imaging, to characterize the cellular adhesion forces between protein-coated tips and living cells at both single molecule and whole cell levels, and to determine the viscoelasticity of living cells and biological samples, including ophthalmic tissues [5, 14, 17, 20, 21, 23, 27, 34, 36, 37, 39, 41–43, 45, 48, 49, 51, 52, 54, 59–61, 64, 96–118].

Force Mapping Mode

Force mapping or force volume mode can be seen as a combination of both imaging and force modes. It consists of collecting a series of force curves at different points of the sample and obtaining a map of the adhesion or the elastic properties of the sample combined with topographic information. This mode has been used to determine the distribution of adhesion receptors over cells surface [64, 96, 105–107] and to map the stiffness variability within living cells [54, 59–61].

11.2.2 Instrumentation

The main components of an AFM include a cantilever with a tip that probes the sample surface, piezoelectric elements to move the tip relative to the sample, a segmented photodiode to monitor the cantilever deflection, and a feedback system linked to a computer to control the system. Since its invention in 1986 [12], AFM instrumentation has evolved, especially during the first decade. One of the most relevant advances was the introduction of the optical lever deflection method, which facilitated the generalized application of the AFM to study biological samples under liquid conditions [108, 109]. Another important improvement was the combination of simultaneous atomic force and inverted optical microscopy, which enables visualization of biological samples, such as living cells, and accurate positioning of the AFM tip on specific regions [110]. Other advances include the use of position detectors (strain gauge or capacitive) to monitor the actual position of the piezoelectric elements, to correct for inherent hysteresis and creep effects, and to allow more accurate repositioning of the tip, especially with the long ranges used when dealing with whole cells or parts of tissues. Recent advances in the biological field include the differential force microscope, which uses an additional reference lever to minimize drift between the tip and the sample for long time measurements [111], and the horizontal force probe to enable lateral force measurements with lateral visualization of the system [112, 113].

The election of the most adequate AFM system for each intended experiment is also crucial. Thus, long-range piezoelectric elements with position sensors are adequate for cell–cell adhesion measurements [104, 114, 115], whereas small open-loop piezoelectric scanners with high resonant frequency are required for high-speed AFM [116].

The most relevant part of an AFM is the cantilever tip, because it interacts with the sample and determines the maximum resolution obtainable. There is a wide variety of different AFM cantilevers with various tips commercially available. The selection of the appropriate probe, taking into account the sample properties as well as the tip stiffness and geometry, is crucial for the reliability of the experiment. For example, to measure the mechanical properties of biological materials, it is important to choose a soft cantilever, with an unsharpened or spherical cantilever tip to minimize damage to the sample.

Uncoated cantilevers are required to reduce thermal drift in the case of long experimental time.

Several types of cantilevers are available commercially with different tip geometries and dimensions. Sharpened conical or pyramidal tips have apex dimension on the order of few nanometers and have been used to image reconstituted proteins and native membranes with high, subnanometer resolution [68, 69, 76, 83, 90, 91, 117–119]. Unsharpened pyramidal tips, instead, are more appropriate to image living cells and map their mechanical properties [52, 54, 55]. Spherical silica or polystyrene beads have been also used as microindenters to measure the rheology of living cells under various conditions and to study membrane fusion processes [48, 49, 51, 120, 121]. Custom made cylindrical tips, with a diameter ranging from 100 nm to 2 μm , have been also used as nanodissecting needles to probe the mechanical properties of living cells and gels [52, 100, 122].

11.2.3 Mechanical Measurements

In addition to providing high-resolution images of surface topography, AFM can be used for elasticity measurements [45, 49, 64, 70, 88, 102, 123]. In this mode, the AFM tip is lowered onto the sample at a set rate. The piezoelectric mechanism lowers the cantilever by a preset amount, the “modulation amplitude” (usually in the micron range), thereby indenting the sample. In response, the cantilever deflects (bends) an amount dependent on the geometry of the tip and on the softness of the sample: the harder the sample, the more the cantilever deflects (Fig. 11.2); the softer, the more the tip indents. The deflection of the cantilever is proportional to the force that the AFM probe tip exerts on the sample (i.e., $F = -kd$, where k is the spring constant of the cantilever and d is the deflection of the cantilever) per Hooke’s law. The recorded cantilever deflection–displacement curves can be used to derive the force-indentation curves for the sample when the deflection response of the cantilever on a hard surface is known.

The geometry of the tip determines the area of contact between the tip and the sample. Because the most common tip geometries are conical or pyramidal, the area of contact will depend on the indentation itself, leading to a nonlinear force-indentation relationship. Thus, knowing the geometry of the tip and the corresponding force-indentation relationship, one is able to determine the stiffness of the sample. To reliably determine the viscoelasticity of soft biological samples, the tip geometry should be blunt, unsharpened and smooth to prevent sample damage and high strains [41–43, 98, 99, 124–126].

Two different approaches are commonly used to determine the elastic or the viscoelastic properties of samples: quasistatic and oscillatory. In quasistatic mode, high amplitude, low-frequency oscillations are produced, which result in force-indentation (F – δ) curves that can be modeled by an appropriate contact elastic model. Low speed is required to accomplish quasistatic conditions, where viscous effects are minimized and the drag effects

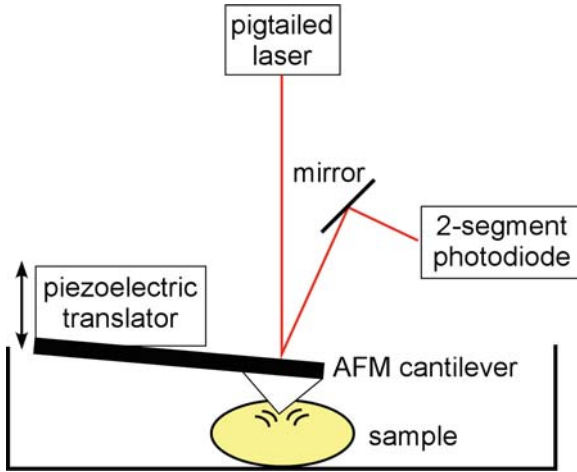


Fig. 11.2. Interaction of the atomic force microscopy cantilever with the sample. Dependent on the mechanical properties of the sample, the cantilever will undergo a combination of deflection (bending) and indentation

on the cantilever are negligible. In addition, a noncontact region is also recommended to accurately determine the initial point of contact between the tip and the sample surface [99]. To minimize adhesion artifacts, the contact model is fitted to the approaching trace of the force curves. A review of the various tip geometries with the corresponding contact elastic modes and their advantages and disadvantages in probing soft biological samples is provided in Table 11.1.

The oscillatory mode enables estimation of the viscoelastic properties of a sample, the complex shear modulus, and its frequency response [$G^*(\omega)$]. In this mode, low-amplitude oscillations (50–100 nm) at single or multiple frequencies are applied at a fixed or slowly varying indentation point. The range of frequencies is limited by thermal drifts and by the piezo and electronic bandwidth [40, 64, 98, 131]. To analyze the data, it is necessary to expand in Taylor terms the appropriate contact model (Table 11.1), $F(\delta)$, for low oscillations around a fixed indentation (δ_0). Using the relationship between elastic modulus (E), rigidity modulus (G), and Poisson ratio (ν), $E = 2G(1 - \nu)$, solving for G and transforming into the frequency domain, it is possible to obtain an expression like the following for the cylindrical contact model with radius a :

$$G^*(\omega) = \frac{1 - \nu}{4a} \frac{F(\omega)}{\delta(\omega)}$$

ω being the angular frequency ($\omega = 2\pi f$), and $F(\omega)$ and $\delta(\omega)$, the Fourier transforms of force and indentation, respectively. The expression of the complex shear modulus obtained above can be then decomposed into an elastic,

Table 11.1. Contact elastic models for various common AFM cantilever tip geometries

Tip geometry	Contact model, $F(\delta)$	Advantages	Disadvantages
Sphere of radius R	$F = \frac{4E\sqrt{R}}{3(1-\nu^2)}\delta^{3/2}$ [127]	Global stiffness, low strains, lower minimum indentation (~ 200 nm)	Low lateral resolution (>500 nm), limited imaging
Flat-ended cylinder of radius a	$F = \frac{2Ea}{(1-\nu^2)}\delta$ [52]	Linear response, constant area of contact, lower minimum indentation (~ 200 nm)	Low lateral resolution (a), limited imaging, costly fabrication procedure
Cone of semi-included angle θ	$F = \frac{2E \tan \theta}{\pi(1-\nu^2)}\delta^2$ [128]	Local stiffness, higher lateral resolution (~ 100 nm), imaging	Higher indentations required (>300 nm)
n -sided regular pyramid of semi-included angle θ	$F \approx \frac{2E}{1-\nu^2} \frac{\tan \theta}{n \sin \frac{\pi}{n}}\delta^2$ [51, 129]	Local stiffness, high lateral resolution (~ 100 nm), imaging	Higher indentations required (>300 nm)
Blunted n -sided pyramidal tip	[51]	Local stiffness, high lateral resolution (~ 100 nm), imaging, lower minimum indentation (~ 200 nm)	Suggested SEM image of tip
Sphere of radius R indenting thin samples	[39, 124, 130]	Thin regions, local stiffness, Mahaffy et al. also determined Poisson ratio	Low lateral resolution (>500 nm), limited imaging

E , Young's modulus, ν , Poisson ratio, F , force, and δ , indentation

in-phase, real part and a viscous, out-of-phase, imaginary part, $G^*(\omega) = G'(\omega) + iG''(\omega)$.

11.3 Atomic Force Microscopy in Ophthalmology

11.3.1 Cornea

Structure

The outer shell of the eye consists of the sclera and the cornea. The sclera is the outer white part of the eye that provides strength to protect the eye from

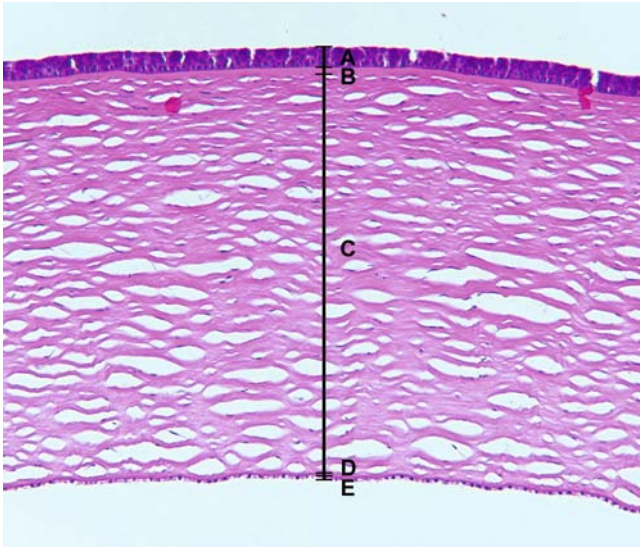


Fig. 11.3. Histological section of a human cornea. (a) Corneal epithelium, (b) Bowman's membrane, (c) stroma, (d) Descemet's membrane, and (e) corneal endothelium. Image used with permission of Sander Dubovy, M.D. and Carlos Medina, M.D., Bascom Palmer Eye Institute, University of Miami Miller School of Medicine

external stressors as well as to withstand the intraocular pressure. The cornea is the transparent, avascular structure through which light rays pass (Fig. 11.3) [132]. The cornea contributes approximately 66% of the total refractive power of the eye [132]. The cornea transmits wavelengths from 310 to 2,500 nm with minimal light scattering or reflection. These unique optical characteristics are due mainly to the structure and composition of the stroma, the middle layer of the cornea that contributes 90% of total corneal thickness (Fig. 11.3) [132]. The stroma consists of nearly 250 layers of collagen fibers, extracellular matrix, and keratocytes. The collagen fibers are arranged parallel to the corneal surface except at the edge (the limbus), where they run circumferentially to form a ring around the cornea. This collagen ring gives the cornea its curved shape responsible for focusing incoming light rays [132].

Studies have shown that the collagen fibers of the stroma are composed primarily of type I collagen, but also contain types III, V, VI, XII, and XIV [78]. The refractive index of the collagen fibers is 1.411, whereas the refractive index of the extracellular matrix is 1.365. Despite this disparity in optical properties, the corneal stroma scatters less than 10% of the incident light [132]. The transparency and minimal light scattering of the cornea is attributed to the uniform arrangement of the collagen fibers. In addition, it is believed that the spacing of the fibrils should be less than half the wavelength of the incoming light. The morphology of the stroma has been investigated using various

microscopic techniques, including scanning electron microscopy and atomic force microscopy, to determine its role in the transparency of the cornea.

The collagen arrangement in the sclera is often used to contrast with that of the corneal stroma. Although the composition of the sclera is comparable to that of the corneal stroma, the optical properties are vastly different [78]. Comparing the collagen fibril direction, diameter, and organization of the cornea and sclera is necessary to understand both corneal transparency and sclera opacity [78].

The morphology of the cornea and sclera was first investigated using atomic force microscopy in 1995 [71]. Fresh samples were taken from bovine eyes and imaged using an AFM (ExplorerTM, Topometrix, Santa Clara, CA) in contact mode with silicon nitride cantilevers with pyramidal tips. All imaging was performed in air. The images revealed regular axial periodicity of the collagen fibrils, with the diameter of each fibril ranging from 35 to 48 nm. Fullwood et al. [71] also reported the presence of cross-bridge structures connecting individual fibrils. In contrast, the collagen fibrils of the sclera were considerably larger, measuring up to 900 nm in diameter. Cross-bridge structures were also observed in the sclera, but some fibrils were found to be in direct contact [71]. Studies that imaged isolated bovine corneal and scleral fibrils using tapping mode in air found similar differences [40, 85, 86]. Interestingly, these same qualitative results were found in a study using AFM in contact mode to investigate the corneal and sclera collagen of the mouse [80].

Meller et al. [79] performed a morphological investigation of human corneal and scleral collagen using an AFM (ExplorerTM; Fig. 11.4) in contact mode with silicon nitride cantilevers with pyramidal tips. This study found that corneal collagen was 48–113 nm in diameter, whereas sclera collagen was 118–1268 nm in diameter. The difference in height between the grooves and ridges in corneal collagen was approximately half that of scleral collagen (0.23 nm vs. 0.42 nm, respectively). Similar results were found by imaging isolated human corneal and sclera collagen fibers [5, 14, 16, 17, 19–23, 27, 29, 31, 34, 36, 37, 39, 80, 87].

The corneal epithelium consists of five to seven layers of cells, and it serves to protect the cornea from the surrounding environment. Junctional complexes between the superficial cells keep this boundary tight [132]. The structure of the rabbit corneal epithelium was studied using an AFM (Nanoscope III, Veeco, Plainview, NY) in both contact and tapping modes [84]. The polygonal epithelial cells with microprojections over their surface were easily distinguished from the images. This study demonstrated the advantages of AFM over scanning electron microscopy imaging of this tissue.

The corneal endothelium is important because it acts as a barrier between the liquid aqueous humor inside the eye and the stroma. Movement of water into the stroma would cause swelling and reduction of its transparency [132]. The endothelium, therefore, serves as a metabolic pump controlling ion transport to the stroma. An AFM (Nanoscope III) was used in contact mode or tapping mode to elucidate the physiology of the corneal endothelium by

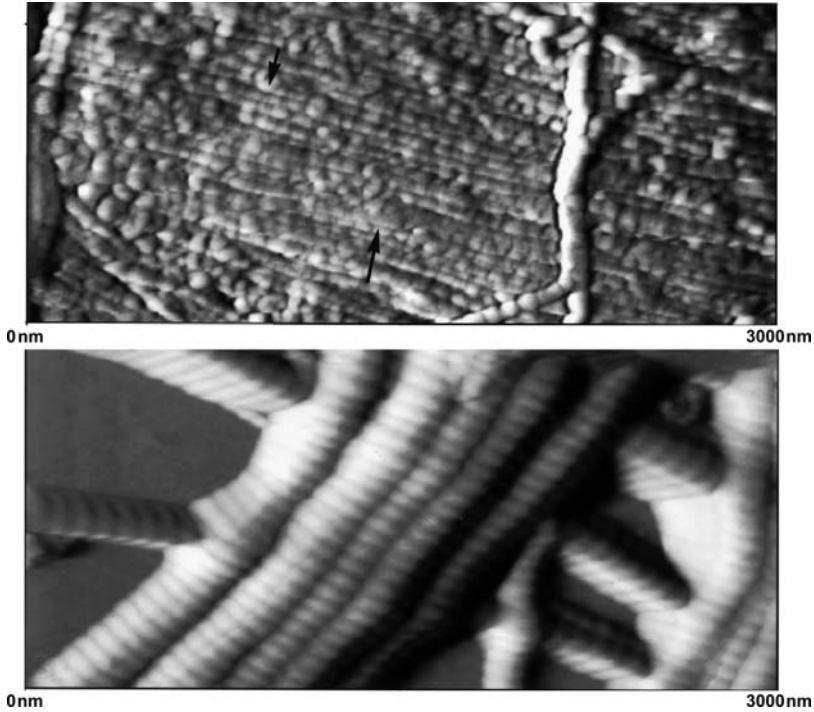


Fig. 11.4. Collagen fibrils of the cornea (*top*) and sclera (*bottom*). The corneal fibrils have a beaded structure, whereas the sclera fibrils are clearly banded [79]

imaging the surface covering layer of the cells [73]. The tissue preparation techniques required for traditional microscopy techniques, such as transmission electron microscopy and scanning electron microscopy, do not enable preservation of this cellular layer.

The corneal stroma is bounded by two basement membranes, the Bowman's membrane and Descemet's membrane (see Fig. 11.3). These basement membranes play a vital role in regulating the cellular processes of the epithelium and endothelium. AFM in tapping mode was used to investigate the surface topography of these membranes to better understand their effect on normal cell behavior [65]. These studies showed that basement membranes consist of an interwoven meshwork of collagen fibers and pores. These features were larger and less densely packed in the epithelial basement membrane, suggesting fundamental functional differences in the overlying cells [65].

These studies have illustrated the utility of the AFM in imaging the different layers of the cornea and sclera. Collagen fibrils of each of these tissues were investigated at the nanometer scale using quantitative three-dimensional imaging. Because the images were performed on fresh samples, the quantitative data obtained is not affected by dehydration during preparation required by other microscopy techniques. Differences between the structure

and arrangement of collagen fibrils in the cornea and the sclera help to elucidate the relationship between structure and optical properties in these ocular tissues.

Corneal Refractive Surgery

AFM has also been utilized to investigate structural changes in the corneal stroma after ablation with an excimer laser. In the photorefractive keratectomy and laser-assisted in situ keratomileusis procedures, the corneal stroma is reshaped to change the refractive power of the eye. By performing ablation to modify the anterior corneal curvature, a perfectly emmetropic eye can be created, eliminating the need to wear glasses. Studies have shown that the outcome of these surgical procedures is optimal when the ablated surface is as smooth as possible [72, 82, 133, 134]. A smooth surface will ensure proper remodeling of the surface and reduce postsurgical visual abnormalities [72, 82]. An AFM (Topometrix; Autoprobe CP, Veeco, Plainview, NY) was used to image the human [82] and porcine [72] cornea in contact mode before and after ablation with the excimer laser. AFM and scanning electron microscopy provided the same qualitative data [82]; however, AFM was able to provide quantitative three-dimensional data of the samples under fluid [72] without tissue shrinkage [82]. AFM images demonstrated submicron irregularities on the stromal surface after photoablation.

Corneal Transplant Surgery

Corneal transplant surgery requires the use of conventional suturing to connect the donor cornea [77]. Laser welding of the donor cornea to the host eye reduces inflammation and scarring [77]. It is thought that laser welding occurs due to changes in the ultrastructural organization of the collagen fibers at the site [77]. AFM was used to investigate the corneal stroma before and after laser welding to determine if localized restructuring occurred. The images demonstrated a disorganization of the collagen fibers at the welding site; however, AFM images clearly showed that the structure of individual fibers was not affected by the treatment [77].

11.3.2 Contact Lenses

Surface Characterization

In a perfectly emmetropic eye, the combined refractive power of the cornea and lens is sufficient to focus incoming light rays onto the macula of the retina. Ametropia is a deviation from emmetropia, with light rays focused either in front of or behind the macula (myopia and hyperopia, respectively) [135]. Contact lenses are the most popular form of nonsurgical refractive correction. However, the use of contact lenses can result in several complications,

including corneal infections and injury [136]. These complications have been related to the properties of the contact lens material, including soft versus hard lenses, as well as to its interactions with the cornea [137].

Minimizing the surface roughness of the contact lens is vital to its optimization. Surface roughness will result in reduced optical quality and biocompatibility of the lens [138]. AFM was used to characterize the surface roughness of various contact lenses [137–145]. AFM images taken in contact mode of soft contact lenses showed numerous pits, crevices, and peaks due to manufacturing. Soft contact lenses are manufactured using one of three methods: lathe-cutting, spin casting, or cast molding [142]. AFM was used to determine the average root mean square roughness value [142–144, 146]. It was found that a lathed surface produced the greatest surface roughness, and spin casting and cast molding produced the smoothest surfaces (Fig. 11.5). It was also demonstrated using AFM that the surface roughness increased after wear [137, 141].

Bacterial adhesion to contact lenses can cause a serious ocular infection known as microbial keratitis, which can lead to permanent vision loss [147]. Rigid gas permeable contact lenses were imaged using AFM in contact mode

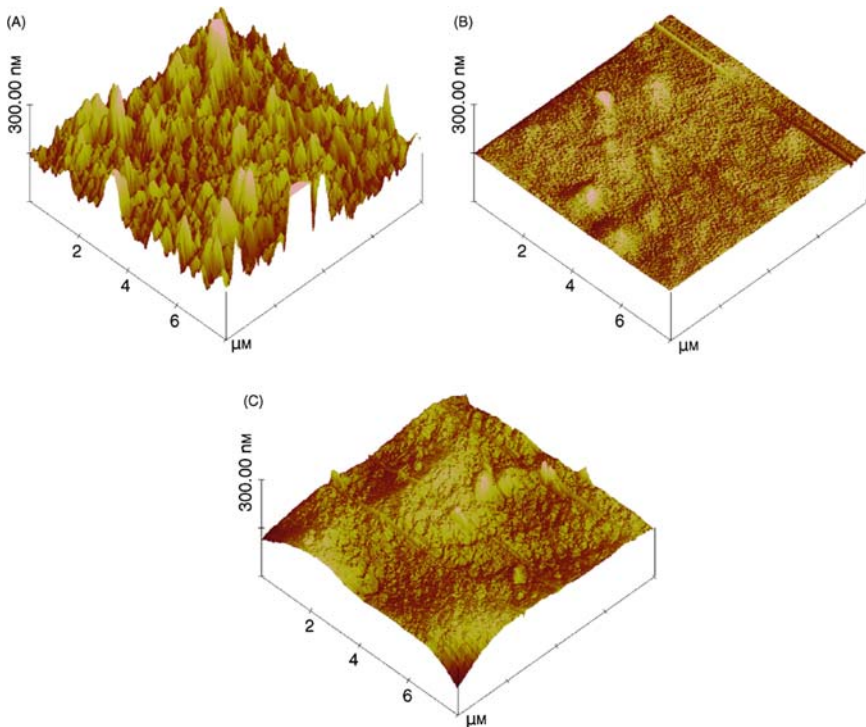


Fig. 11.5. Atomic force microscopy images of contact lens surfaces manufactured using different methods: (a) lathe-cutting, (b) spin casting, and (c) cast-molding [142]

[148]. This study demonstrated lens surface roughness increases with wear, eventually reaching a critical limit when the surface becomes more prone to bacterial adhesion [148]. With the advent of silicone–hydrogel soft contact lenses, the incidence of microbial keratitis decreased fivefold [147], resulting in the hypothesis that hydrogel contact lenses bind fewer bacteria than traditional hydroxyethylmethacrylate (HEMA) contact lenses. It is also thought that lens wear may increase the extent of bacterial adhesion. Unworn and worn hydrogel and HEMA contact lenses treated with *Staphylococcus epidermidis* were imaged with AFM in tapping mode [147]. This study demonstrated that unworn and worn soft contact lenses had the same propensity to bacterial adhesion. Hydrogel contact lenses were less prone to bacterial adhesion than HEMA contact lenses [147].

Biomechanical Properties

Research has shown that the comfort of contact lenses is dependent on the contact lens surface water content. It is known that the water content of a material is directly related to its mechanical properties: as water content increases, hardness decreases. Therefore, measurement of the mechanical properties of contact lenses can help elucidate surface water content [149]. Using an AFM to measure Young's modulus of elasticity of a soft hydrogel contact lens, it was found that the contact lens surface is dry under ambient humidity, but becomes softer with increasing humidity [149]. This indicates that the contact lens surface in contact with the air will be extremely stiff and may affect the interaction of the lens with the eyelid [149].

AFM was also used in adhesion and friction studies of hydrogel contact lenses [150, 151]. It was found that soft contact lenses with a higher water content had significantly lower adhesion and friction values [150, 151].

11.3.3 Lens

Structure

The lens is a transparent, avascular body located between the iris and the vitreous [152]. The lens consists of organized layers of long fiber cells. After light rays entering the eye are bent by the cornea, the lens is responsible for an additional 30% bending [5]. The combined optical effect of the cornea and lens produces emmetropic vision. Development of the lens begins as early as day 28 of human gestation with the formation of the lens placode [153]. Lens growth does not cease after gestation due to the continued proliferation of the lens epithelial cuboidal cells located at the equator. These epithelial cells differentiate into the lens fiber cells that make up the bulk of the lens material [153]. Because the older cells are not shed, they are displaced toward the center of the lens. As these older fiber cells continue to mature, they lose their cellular organelles; consequently, metabolic activity comes to an

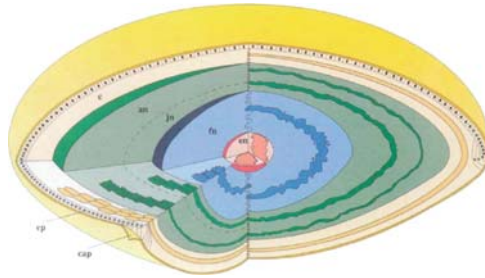


Fig. 11.6. Diagram of the mature lens. The immature, nucleated lens fiber cells are in the outer region of the lens, whereas the mature, enucleated lens fiber cells are towards the center [155]

end [153]. Eventually, the older fiber cells become compacted in the center of the lens, forming a dense nucleus with a decreased concentration of water, a condition known as sclerosis [154]. This cyclic growth process results in an onion-like structure, with the newer, softer fiber cells on the outside, and the older, denser cells in the inner nucleus (Fig. 11.6) [152].

As a person ages, the transparency of the lens is lost due to disruption of lens fiber cell organization [156]. As the opacity increases, it prevents light rays from passing through the lens and focusing on the retina. This condition is known as cataract, and it is the leading cause of blindness worldwide. The ultrastructure of the lens fibers is important in determining their role in cataract formation. An AFM (Pico SPM I, Molecular Imaging, Tempe, AZ) was used in tapping mode to image the healthy lens fiber cell arrangement in the adult New Zealand white rabbit (Fig. 11.7) [66]. This study was expanded to determine differences in the lens fiber cell structure in healthy and cataractous canine lenses [67]. Using AFM images, these studies illustrated that lens fiber cells in the young lens are extremely thin and organized. In later stages of cataract development, however, the fibers lose their organization. This loss of organization could be responsible for the opacity characteristic of cataractous lenses [67].

The organization of lens fiber cells was investigated in even greater detail using an AFM (Nanoscope III) in contact mode [68, 69, 76, 83]. Two types of membrane proteins form junctions between neighboring fiber cells: aquaporin 0 and connexins [68]. Studies have shown that mutations in these two proteins result in cataract formation [157]. To gain greater insight into these proteins and the organization of the membrane, AFM was used to image the sheep and human fiber cell membrane (Fig. 11.8). Imaging these proteins at the nanoscale enabled the identification of differences between healthy and diabetic patients. Diabetic lens membranes lack connexin proteins, which are responsible for ion, metabolite, and waste flux between fibers [76]. The lack of connexin proteins has also been reported in cataract membranes, which could explain how diabetes can induce cataract formation [76].

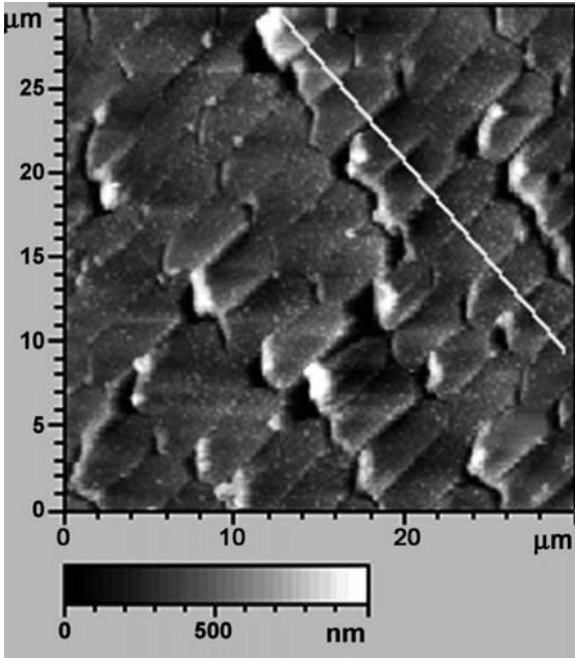


Fig. 11.7. Lens fiber cell arrangement in the cortical region of a healthy adult rabbit [67]

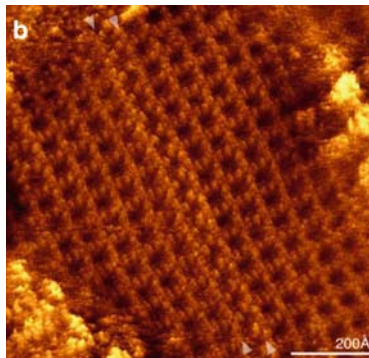


Fig. 11.8. Aquaporin 0 junctional microdomains of a healthy human fiber cell membrane [76]

Mechanics

The lens is significant in ocular physiology because it changes shape to bring objects located at different distances into focus, a dynamic process known as accommodation. The most widely accepted theory of accommodation is the Helmholtz theory, which attributes the increase in refractive power of the lens during near vision to the contraction of the ciliary muscles and relaxation of

the lens suspensory ligaments called the zonules [5]. The released pressure on the lens capsule enables it to contract, resulting in an increased thickness of the lens. This change in shape is characterized by an increase in the anterior and posterior curvatures of the lens, increasing its refractive power and thereby bringing near objects into focus. As a person ages, the ability to accommodate diminishes until it is lost. The near point of the eye gradually shifts farther away, requiring small objects to be placed at greater distances to be visualized clearly [154]. The loss of accommodation is referred to as presbyopia when it leads to the loss of near vision [8, 158].

Because of lens growth and decreased water content with age, the lens hardness increases by several orders of magnitude. Accommodation can only occur if the lens is sufficiently pliable so that the capsule can mold it and subsequently change its shape [5, 158–160]. In the young eye, the modulus of elasticity of the lens is several orders of magnitude less than that of the outer capsule [153–158]. Therefore, the lens is compliant with the molding pressure of the capsule [161–164]. With age, the lens modulus of elasticity increases, and the capsule can no longer change the shape of the lens. The onset of presbyopia may represent the point in an individual's lifetime when Young's modulus of the lens exceeds that of the capsule [161, 163]. The mechanical properties of all parts of the lens, and their changes with age, are vital to a more complete understanding of accommodation and presbyopia [165].

Because lens hardness is thought to be one of the main causes of presbyopia, several researchers have assessed this parameter using dynamic mechanical analysis [166–170]. However, the extensive tissue preparation utilized most likely affected the results. A custom-built AFM was used to assess the biomechanical properties of primate lenses [88]. This study found that lens elasticity does not significantly decrease in young primate lenses, even though significant power loss has been noted in this same age range [171].

Artificial Lenses

Cataract surgery involves the removal of the opaque lens usually followed by the implantation of an artificial lens, called an intraocular lens. Intraocular lenses compensate for the loss of power of the eye that occurs as a result of lenticular removal.

After implantation of the intraocular lens, any lens epithelial cells left in the capsular bag will proliferate. Proliferation of the epithelial cells can result in the formation of a secondary cataract, a condition known as posterior capsular opacification [172]. To restore vision to the patient, a posterior capsulotomy must be performed using an Nd-YAG laser. It has been suggested that surface properties of the implanted intraocular lens can induce a foreign-body reaction. This reaction, in turn, may contribute to the formation of posterior capsular opacification [173]. Commercially available intraocular lenses were examined using an AFM (Autoprobe CP) in contact mode to determine the relationship between surface properties and secondary cataract

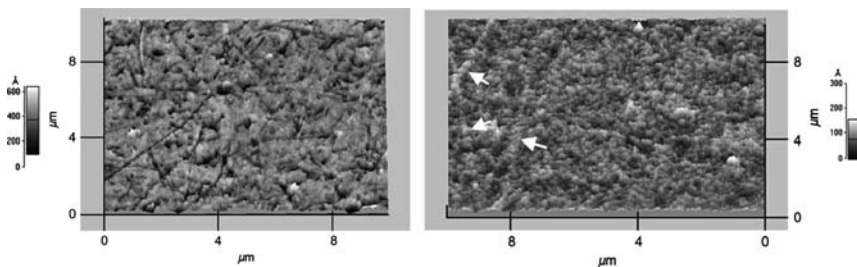


Fig. 11.9. Atomic force microscopy images of the topography of a polymethymethacrylate (PMMA) intraocular lens (*left*) and a silicone intraocular lens (*right*) [172]

formation [172]. This study found statistically significant differences in the surface roughness of the various intraocular lenses imaged. Intraocular lenses made of polymethymethacrylate (PMMA) were the roughest, and are surgically associated with the highest incidence of posterior capsular opacification (Fig. 11.9) [172]. This study suggests that a smoother intraocular lens surface should decrease its cell adhesion and consequently reduce posterior capsular opacification [172].

Studies have shown that the hydrophobic surface of PMMA intraocular lenses can cause adherence of corneal endothelial cells, which can result in a foreign body reaction [174]. Cell adhesion studies were performed on PMMA intraocular lenses coated with hyaluronan using an AFM (Nanoscope III) in force mode [174]. This study found that fibroblast adhesion was much less significant for coated intraocular lenses. However, because hyaluronan dissipates over time, the coating will not be effective in preventing the progression of PCO.

Research has suggested that a modification of the intraocular lens surface could also prevent posterior capsular opacification. The surface of PMMA, silicone, and hydrogel intraocular lenses have been treated and coated with a variety of different methods. To effectively prevent posterior capsular opacification, these coatings should repel both cells and proteins [175]. An AFM (Nanoscope III) was used to image intraocular lenses coated with poly(ethylene glycol) in tapping mode. The images showed that intraocular lenses coated with poly(ethylene glycol) with longer grafted chains were smoother than uncoated intraocular lenses [175]. In vitro studies demonstrated the efficacy of polyethylene glycol in repelling lens epithelial cells.

Intraocular lens surface properties are also important in maintaining the optical quality of the implant. Silicone intraocular lenses tend to have post-operative glistenings, which can cause glare and decreased visual acuity. The surface of an explanted intraocular lens that manifested glistenings was observed using an AFM (Nanoscope III) [176]. These studies revealed significant changes in the surface morphology, which most likely caused the glistenings.

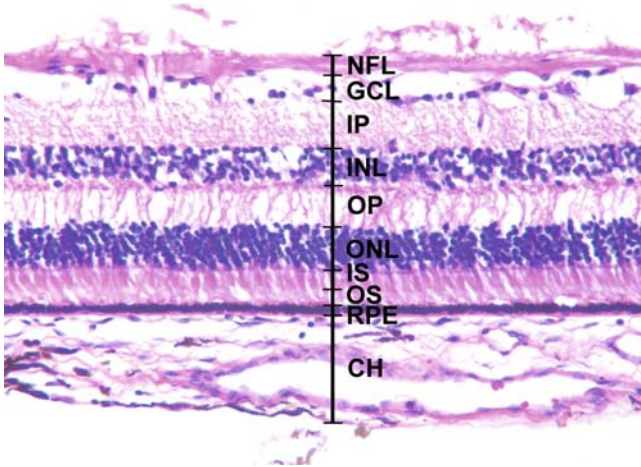


Fig. 11.10. Histological section of a human retina. *NFL* nerve fiber layer, *GCL* ganglion cell layer, *IP* inner plexiform layer, *INL* inner nuclear layer, *OP*, outer plexiform layer, *ONL*, outer nuclear layer, *PL* photoreceptors, *RPE* retinal pigment epithelium, *CH* choroid. Image used with permission of Sander Dubovy, M.D. and Carlos Medina, M.D., Bascom Palmer Eye Institute, University of Miami Miller School of Medicine

11.3.4 Retinal Tissue

Structure

The retina is a structure located at the back of the eye consisting of distinct cellular layers that can be distinguished histologically (Fig. 11.10). The function of the retina is to convert incoming light signals into electrical signals that are transmitted to the visual cortex of the brain [177]. The neural layers consist of ganglion cells, amacrine cells, bipolar cells, horizontal cells, Müller cells, and photoreceptors (rods and cones). The central portion of the retina is known as the macula, which is the site of maximal visual acuity [178]. The central region of the macula, the fovea, occupies 0.02% of the total retinal area, but contains 25% of the total ganglion cells and has a high cone density, both necessary for high visual acuity [178]. The fovea is easily distinguishable because of the recessed vitreal surface (the foveal pit) and high concentration of yellow pigment.

Although the macula has the highest concentration of neural cells in the retina, it has a limited vascular supply. In fact, the foveal pit is an avascular section of the retina [179]. The limited number of blood vessels in the macula contributes to its high visual acuity, because entering light rays are not diffracted by the vasculature. However, this lack of blood supply to the macula drastically reduces the oxygen supply [179]. Photoreceptors require considerable oxygen supply; however, in the macula where the cone density is

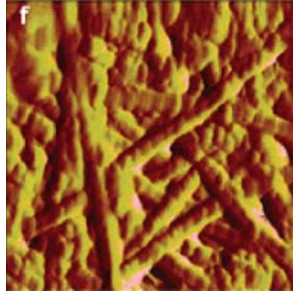


Fig. 11.11. Images of Bruch's membrane of the porcine retina obtained with a commercial atomic force microscope [74]

greatest, the oxygen is lowest. This physiological limitation becomes evident in a person's later years, when there is a greater risk of photoreceptor loss and macular degeneration [179].

Age-related macular degeneration (AMD) is the leading cause of blindness in Americans over 55 years of age [180]. This disease occurs in two distinct forms: neovascular lesion (wet) and atrophic lesion (dry). Wet AMD is characterized by new blood vessels entering the retina from the choroid vasculature, resulting in recurrent hemorrhage [179]. In dry AMD, a reduced blood supply to the macula results in degeneration of the retinal pigment epithelium (RPE) and atrophy of the photoreceptors [179].

Research has suggested that morphological changes in Bruch's membrane with age may cause the structural changes associated with dry AMD [180]. Bruch's membrane separates the RPE from the capillaries in the choroid. With age, this membrane loses its elasticity and subsequently breaks down [180]. Because of this breakdown, Bruch's membrane becomes less effective in exchanging nutrients from the blood supply with waste products in the retina. Disruption of this exchange results in degeneration of RPE cells, and eventually leads to abnormalities associated with age-related macular degeneration [180]. Bruch's membrane has been imaged in tapping mode using an AFM (Multi-Mode Nanoscope IIIa) [74, 75]. The images clearly illustrated the collagen fiber arrangement in the membrane (Fig. 11.11) [74, 75]. Understanding the normal versus degenerated structure is vital to a more complete understanding of the role of Bruch's membrane in AMD [74].

Mechanical Properties

The inner limiting membrane is a retinal basement membrane that separates the retina from the vitreous. This membrane is important in stabilizing the blood vessels and maintaining the integrity of the vitreo-retinal border [70]. Studies have shown that the integrity of the inner limiting membrane is vital to the survival of retinal ganglion cells [181]. Measurement of the mechanical properties of the inner-limiting membrane gives insight into its mechanical stability. Young's modulus of the inner-limiting membrane of the retina was

measured using an AFM (MFP-3D, Asylum Research, Santa Barbara, CA) [70]. These measurements can be used to differentiate normal from diseased states.

11.4 Summary and Conclusions

In this chapter, we summarized recent AFM studies of ocular tissues, while highlighting the great potential of AFM technology in ophthalmic research. These studies demonstrate the versatility of the AFM as a high-resolution imaging technique and as a sensitive force apparatus for probing the mechanical properties of ocular tissues. The studies, however, only begin to address some of the issues in ophthalmology that can be investigated by AFM. A topic that was not discussed in this review is the integration of AFM with other techniques, such as confocal microscopy and optical coherence tomography. Future research carried out with these AFM-based systems will undoubtedly result in new insight into the working of the eye. Moreover, AFM may also play an important role in the development and characterization of eye implants.

Acknowledgements

The authors thank Jean-Marie Parel, Ph.D., and Fabrice Manns, Ph.D., of the Ophthalmic Biophysics Center of Bascom Palmer Eye Institute, for providing their expertise in preparation of this chapter. The authors also thank the following support: NSF MRI 0722372; University of Miami SAC Award; NIH EY14225; NSF Graduate Student Fellowship (NMZ); Rakhi Jain, Ph.D., AMO, CA; Florida Lions Eye Bank; Vision Cooperative Research Centre, Sydney, New South Wales, Australia, supported by the Australian Federal Government through the Cooperative Research Centres Programme; NIH center grant P30-EY014801; Research to Prevent Blindness.

References

1. A. Engel, Y. Lyubchenko, and D. Muller, Atomic force microscopy: a powerful tool to observe biomolecules at work. *Trends Cell Biol.* **9**(2), 77–80(1999).
2. W.F. Heinz, J.H. Hoh, Relative surface charge density mapping with the atomic force microscope. *Biophys. J.* **76**(1 Pt 1), 528–538(1999).
3. H.W. Wu, T. Kuhn, and V.T. Moy, Mechanical properties of 1929 cells measured by atomic force microscopy: effects of anticytoskeletal drugs and membrane crosslinking. *Scanning* **20**, 389–397(1998).
4. A. Glasser, Restoration of accommodation. *Curr. Opin. Ophthalmol.* **17**(1), 12–8(2006).
5. A. Glasser, Accommodation: mechanism and measurement. *Ophthalmol. Clin. North Am.* **19**(1), 1–12(2006).

6. R.I. Barraquer, et al., Human lens capsule thickness as a function of age and location along the sagittal lens perimeter. *Invest. Ophthalmol. Vis. Sci.* **47**(5), 2053–2060(2006).
7. D. Atchison, G. Smith (eds.), *Optics of the Human Eye* (Butterworth Heinemann: Oxford, UK, 2000), pp. 16–18.
8. A Glasser, P.L. Kaufman, Accommodation and presbyopia. In *Adler's Physiology of the Eye. Clinical Application*, 10th edn., edited by P.L. Kaufman, A. Alm (Mosby, St Louis, MO, 2003), pp. 197–233.
9. R. Fisher, The significance of the shape of the lens and capsular energy changes in accommodation. *J. Physiol.* **201**, 21–47(1969).
10. W.N. Charman, The eye in focus: accommodation and presbyopia. *Clin. Exp. Optom.* **91**(3), 207–225(2008).
11. H.J. Wyatt, Some aspects of the mechanics of accommodation. *Vis. Res.* **28**, 75–86(1988).
12. G. Binnig, C.F. Quate, C. Gerber, Atomic force microscope. *Phys. Rev. Lett.* **56**, 930–933(1986).
13. H.X. You, J.M. Lau, S. Zhang, L. Yu, Atomic force microscopy imaging of living cells: a preliminary study of the disruptive effect of the cantilever tip on cell morphology. *Ultramicroscopy* **82**, 297–305(2000).
14. E. A-Hassan, W. Heinz, M.D. Antonik, N.P. D'Costa, S. Nageswaran, C.A. Schoenenberger, J.H. Hoh, Relative microelastic mapping of living cells by atomic force microscopy. *Biophys. J.* **74**, 1564–1578(1998).
15. R. Afrin, A. Ikai, Force profiles of protein pulling with or without cytoskeletal links studied by AFM. *Biochem. Biophys. Res. Commun.* **348**, 238–244(2006).
16. J. Alcaraz, L. Buscemi, M. Grabulosa, X. Trepas, B. Fabry, R. Farre, D. Navajas, Microrheology of human lung epithelial cells measured by atomic force microscopy. *Biophys. J.* **84**, 2071–2079(2003).
17. M. Benoit, H.E. Gaub, Measuring cell adhesion forces with the atomic force microscope at the molecular level. *Cells Tissues Organs* **172**, 174–189(2002).
18. H.J. Butt, B. Cappella, M. Kappl, Force measurements with the atomic force microscope: technique, interpretation and applications. *Surf. Sci. Rep.* **59**, 1–152(2005).
19. G.T. Charras, M.A. Horton, Single cell mechanotransduction and its modulation analyzed by atomic force microscope indentation. *Biophys. J.* **82**, 2970–2981(2002).
20. S.L. Crick, F.C. Yin, Assessing micromechanical properties of cells with atomic force microscopy: importance of the contact point. *Biomech. Model. Mechanobiol.* **6**, 199–210(2007).
21. J. Domke, S. Dannohl, W.J. Parak, O. Muller, W.K. Aicher, M. Radmacher, Substrate dependent differences in morphology and elasticity of living osteoblasts investigated by atomic force microscopy. *Colloids Surf. B Biointerfaces* **19**, 367–379(2000).
22. V. Dupres, F.D. Menozzi, C. Locht, B.H. Clare, N.L. Abbott, S. Cuenot, C. Bompard, D. Raze, Y.F. Dufrene, Nanoscale mapping and functional analysis of individual adhesins on living bacteria. *Nat Methods* **2**, 515–520(2005).
23. A.J. Engler, F. Rehfeldt, S. Sen, D.E. Discher, Microtissue elasticity: measurements by atomic force microscopy and its influence on cell differentiation. *Methods Cell. Biol.* **83**, 521–545(2007).

24. E.L. Florin, M. Rief, H. Lehmann, M. Ludwig, C. Dornmair, V.T. Moy, H.E. Gaub, Sensing specific molecular-interactions with the atomic-force microscope. *Biosens. Bioelectron.* **10**, 895–901(1995).
25. M. Grandbois, W. Dettmann, M. Benoit, H.E. Gaub, Affinity imaging of red blood cells using an atomic force microscope. *J. Histochem. Cytochem.* **48**, 719–724(2000).
26. W. Haberle, J.K.H. Horber, G. Binnig, Force microscopy on living cells. *J. Vac. Sci. Technol. B* **9**, 1210–1213(1991).
27. A. Hategan, R. Law, S. Kahn, D.E. Discher, Adhesively-tensed cell membranes: lysis kinetics and atomic force microscopy probing. *Biophys. J.* **85**, 2746–2759(2003).
28. E. Henderson, P.G. Haydon, D.S. Sakaguchi, Actin filament dynamics in living glial cells imaged by atomic force microscopy. *Science* **257**, 1944–1946(1992).
29. P. Hinterdorfer, Y.F. Dufrene, Detection and localization of single molecular recognition events using atomic force microscopy. *Nat. Methods* **3**, 347–355(2006).
30. J.H. Hoh, C.A. Schoenenberger, Surface morphology and mechanical properties of MDCK monolayers by atomic force microscopy. *J. Cell Sci.* **107**, 1105–1114(1994).
31. K. Hyonchol, H. Arakawa, T. Osada, A. Ikai, Quantification of fibronectin and cell surface interactions by AFM. *Colloid Surf. B-Biointerfaces* **25**, 33–43(2002).
32. H. Kim, H. Arakawa, T. Osada, A. Ikai, Quantification of cell adhesion force with AFM: distribution of vitronectin receptors on a living MC3T3-E1 cell. *Ultramicroscopy* **97**, 359–363(2003).
33. E. Kokkoli, S.E. Ochsenshirt, M. Tirrell, Collective and single-molecule interactions of alpha(5)beta(1) integrins. *Langmuir* **20**, 2397–2404(2004).
34. M. Krieg, Y. Arboleda-Estudillo, P.H. Puech, J. Kafer, F. Graner, D.J. Muller, C.P. Heisenberg, Tensile forces govern germ-layer organization in zebrafish. *Nat. Cell Biol.* **10**, 429–436(2008).
35. R. Lal, B. Drake, D. Blumberg, D.R. Saner, P.K. Hansma, S.C. Feinstein, Imaging real-time neurite outgrowth and cytoskeletal reorganization with an atomic force microscope. *Am. J. Physiol.-Cell Physiol.* **269**, C275–C285(1995).
36. I. Lee, R.E. Marchant, Force measurements on platelet surfaces with high spatial resolution under physiological conditions. *Colloid Surf. B-Biointerfaces* **19**, 357–365(2000).
37. Q.S. Li, G.Y. Lee, C.N. Ong, C.T. Lim, AFM indentation study of breast cancer cells. *Biochem. Biophys. Res. Commun.* (2008).
38. T. Ludwig, R. Kirmse, K. Poole, U.S. Schwarz, Probing cellular microenvironments and tissue remodeling by atomic force microscopy. *Pflugers Arch.* **456**, 29–49(2008).
39. R.E. Mahaffy, S. Park, E. Gerde, J. Kas, C.K. Shih, Quantitative analysis of the viscoelastic properties of thin regions of fibroblasts using atomic force microscopy. *Biophys. J.* **86**, 1777–1793(2004).
40. R.E. Mahaffy, C.K. Shih, F.C. MacKintosh, J. Kas, Scanning probe-based frequency-dependent microrheology of polymer gels and biological cells. *Phys. Rev. Lett.* **85**, 880–883(2000).
41. J.C. Martens, M. Radmacher, Softening of the actin cytoskeleton by inhibition of myosin II. *Pflugers Arch.* **456**, 95–100(2008).

42. A.B. Mathur, A.M. Collinsworth, W.M. Reichert, W.E. Kraus, G.A. Truskey, Endothelial, cardiac muscle and skeletal muscle exhibit different viscous and elastic properties as determined by atomic force microscopy. *J. Biomech.* **34**, 1545–1553(2001).
43. A.B. Mathur, G.A. Truskey, W.M. Reichert, Atomic force and total internal reflection fluorescence microscopy for the study of force transmission in endothelial cells. *Biophys. J.* **78**, 1725–1735(2000).
44. V.T. Moy, E.L. Florin, H.E. Gaub, Adhesive forces between ligand and receptor measured by AFM. *Biophys. J.* **66**, A340(1994).
45. P.H. Puech, A. Taubenberger, F. Ulrich, M. Krieg, D.J. Muller, C.P. Heisenberg, Measuring cell adhesion forces of primary gastrulating cells from zebrafish using atomic force microscopy. *J. Cell Sci.* **118**, 4199–4206(2005).
46. C.A.J. Putman, K.O. Vanderwerf, B.G. Degrooth, N.F. Vanhulst, J. Greve, Viscoelasticity of living cells allows high-resolution imaging by tapping mode atomic-force microscopy. *Biophys. J.* **67**, 1749–1753(1994).
47. C.A.J. Putman, K.O. Vanderwerf, B.G. Degrooth, N.F. Vanhulst, F.B. Segerink, J. Greve, Atomic force microscope with integrated optical microscope for biological applications. *Rev. Sci. Instr.* **63**, 1914–1917(1992).
48. M. Radmacher, Measuring the elastic properties of biological samples with the AFM. *IEEE Eng. Med. Biol. Magn.* **16**, 47–57(1997).
49. M. Radmacher, Studying the mechanics of cellular processes by atomic force microscopy. *Methods Cell Biol.* **83**, 347–372(2007).
50. M. Radmacher, R.W. Tillmann, M. Fritz, H.E. Gaub, From molecules to cells: imaging soft samples with the atomic force microscope. *Science* **257**, 1900–1905(1992).
51. F. Rico, P. Roca-Cusachs, N. Gavara, R. Farre, M. Rotger, D. Navajas, Probing mechanical properties of living cells by atomic force microscopy with blunted pyramidal cantilever tips. *Phys. Rev. E* **72**(2005).
52. F. Rico, P. Roca-Cusachs, R. Sunyer, R. Farre, D. Navajas, Cell dynamic adhesion and elastic properties probed with cylindrical atomic force microscopy cantilever tips. *J. Mol. Recognit.* **20**, 459–466(2007).
53. M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109–1112(1997).
54. C. Rotsch, F. Braet, E. Wisse, M. Radmacher, AFM imaging and elasticity measurements on living rat liver macrophages. *Cell Biol. Int.* **21**, 685–696(1997).
55. C. Rotsch, M. Radmacher, Drug-Induced changes of cytoskeletal structure and mechanics in fibroblasts: an atomic force microscopy study. *Biophys. J.* **78**, 520–535(2000).
56. S.W. Schneider, P. Pagel, C. Rotsch, T. Danker, H. Oberleithner, M. Radmacher, A. Schwab, Volume dynamics in migrating epithelial cells measured with atomic force microscopy. *Pflugers Arch.-Eur. J. Physiol.* **439**, 297–303(2000).
57. S.G. Shroff, D.R. Saner, R. Lal, Dynamic micromechanical properties of cultured rat atrial myocytes measured by atomic-force microscopy. *Am. J. Physiol.-Cell Physiol.* **38**, C286–C292(1995).
58. A. Simon, M.-C. Durrieu, Strategies and results of atomic force microscopy in the study of cellular adhesion. *Micron* **37**, 13(2006).

59. L. Sirghi, J. Ponti, F. Broggi, F. Rossi, Probing elasticity and adhesion of live cells by atomic force microscopy indentation. *Eur. Biophys. J.* **37**, 935–945(2008).
60. M. Stolz, R. Raiteri, A.U. Daniels, M.R. VanLandingham, W. Baschong, U. Aebi, Dynamic elastic modulus of porcine articular cartilage determined at two different levels of tissue organization by indentation-type atomic force microscopy. *Biophys. J.* **86**, 3269–3283(2004).
61. A. Touhami, B. Nysten, Y.F. Dufrene, Nanoscale mapping of the elasticity of microbial cells by atomic force microscopy. *Langmuir* **19**, 4539–4543(2003).
62. H.X. You, L. Yu, Atomic force microscopy imaging of living cells: progress, problems and prospects. *Methods Cell Sci.* **21**, 1–17(1999).
63. A.E. Pelling, M.A. Horton, An historical perspective on cell mechanics. *Pflugers Arch.* **456**, 3–12(2008).
64. P. Roca-Cusachs, I. Almendros, R. Sunyer, N. Gavara, R. Farre, D. Navajas, Rheology of passive and adhesion-activated neutrophils probed by atomic force microscopy. *Biophys. J.* **91**, 3508–3518(2006).
65. G.A. Abrams, S.S. Schaus, S.L. Goodman, P.F. Nealey, C.J. Murphy. Nanoscale topography of the corneal epithelial basement membrane and Descemet's membrane of the human. *Cornea* **19**(1), 57–64(2000).
66. A. Antunes, F.V. Gozzo, M.I. Borella, M. Nakamura, A.M.V. Safatle, P.S.M. Barros, H.E. Toma. Atomic force imaging of ocular tissues: morphological study of healthy and cataract lenses. In *Modern Research and Educational Topics in Microscopy* (Formatex, 2007a).
67. A. Antunes, F.V. Gozzo, M. Nakamura, A.M. Safatle, S.L. Morelhão, H.E. Toma, P.S. Barros. Analysis of the healthy rabbit lens surface using MAC mode atomic force microscopy. *Micron* **38**(3), 286–290(2007b).
68. N. Buzhynskyy, J.F. Girmens, W. Faigle, S. Scheuring, Human cataract lens membrane at subnanometer resolution. *J. Mol. Biol.* **374**, 162–169(2007a).
69. N. Buzhynskyy, R.K. Hite, T. Walz, S. Scheuring. The supramolecular architecture of junctional microdomains in native lens membranes. *EMBO Rep.* **8**, 51–55(2007b).
70. J. Candiello, M. Balasubramani, E.M. Schreiber, G.J. Cole, U. Mayer, W. Halfter, H. Lin. Biomechanical properties of native basement membranes. *FEBS J.* **274**(11), 2897–908(2007).
71. N.J. Fullwood, A. Hammiche, H.M. Pollock, D.J. Hourston, M. Song. Atomic force microscopy of the cornea and sclera. *Curr. Eye Res.* **14**, 529–535(1995).
72. M. Lombardo, M.P. De Santo, G. Lombardo, R. Barberi, S. Serrao. Atomic force microscopy analysis of normal and photoablated porcine corneas. *J. Biomech.* **39**(14), 2719–2724(2006a).
73. S. Lydataki, E. Lesniewska, M.K. Tsilimbaris, C. Le Grimellec, L. Rochette, J.P. Goudonnet, I.J. Pallikaris. Observation of the posterior endothelial surface of the rabbit cornea using atomic force microscopy. *Cornea* **22**(7), 651–664(2003).
74. S.B. Mallick, A. Ivanisevic. Study of the morphological and adhesion properties of collagen fibers in the Bruch's membrane. *J. Phys. Chem. B.* **109**(41):19052–19055(2005).
75. S.B. Mallick, S. Bhagwandin, A. Ivanisevic. Characterization of collagen fibers in Bruch's membrane using chemical force microscopy. *Anal. Bioanal. Chem.* **386**(3), 652–657(2006).

76. S. Mangenot, N. Buzhynskyy, J.F. Girmens, S. Scheuring. Malformation of junctional microdomains in cataract lens membranes from a type II diabetes patient. *Eur. J. Physiol.* (2008).
77. P. Matteini, F. Sbrana, B. Tiribilli, R. Pini. Atomic force microscopy and transmission electron microscopy analyses of low-temperature laser welding of the cornea. *Lasers Med. Sci.*(2008). DOI 10.1007/s10103-008-0617-4.
78. K.M. Meek, N.J. Fullwood. Corneal and scleral collagens – a microscopist’s perspective. *Micron* **32**, 261–272(2001).
79. D. Meller, K. Peters, K. Meller. Human cornea and sclera studied by atomic force microscopy. *Cell Tissue Res.* **288**(1), 111–118(1997).
80. A. Miyagawa, M. Kobayashi, Y. Fujita, M. Nakamura, K. Hirano, K. Kobayashi, Y. Miyake, Surface topology of collagen fibrils associated with proteoglycans in mouse cornea and sclera. *Japan J. Ophthalmol.* **44**, 591–595(2000).
81. A. Miyagawa, M. Kobayashi, Y. Fujita, O. Hamdy, K. Hirano, M. Nakamura, Y. Miyake, Surface ultrastructure of collagen fibrils and their association with proteoglycans in human cornea and sclera by atomic force microscopy and energy-filtering transmission electron microscopy. *Cornea* **20**(6), 651–656(2001).
82. A. Nógrádi, B. Hopp, K. Révész, G. Szabó, Z. Bor, L. Kolozsvari, Atomic force microscopic study of the human cornea following excimer laser keratectomy. *Exp. Eye Res.* **70**, 363–368(2000).
83. S. Scheuring, N. Buzhynskyy, S. Jaroslowski, R.P. Goncalves, R.K. Hite, T. Walz, Structural models of the supramolecular organization of AQP0 and connexons in junctional microdomains. *J. Struct. Biol.* **160**, 385–394(2007).
84. M.K. Tsilimbaris, E. Lesniewska, S. Lydataki, C. Le Grimellec, J.P. Goudonnet, I.G. Pallikaris, The use of atomic force microscopy for the observation of corneal epithelium surface. *Invest. Ophthalmol. Vis. Sci.* **41**(3), 680–686(2000).
85. S. Yamamoto, J. Hitomi, M. Shigeno, S. Sawaguchi, H. Abe, T. Ushiki Atomic force microscopic studies of isolated collagen fibrils of the bovine cornea and sclera. *Arch. Histol. Cytol.* **60**(4), 371–378(1997).
86. S. Yamamoto, H. Hashizume, J. Hitomi, M. Shigeno, S. Sawaguchi, H. Abe, T. Ushiki, The subfibrillar arrangement of corneal and scleral collagen fibrils as revealed by scanning electron and atomic force microscopy. *Arch. Histol. Cytol.* **63**(2), 127–135(2000).
87. S. Yamamoto, J. Hitomi, S. Sawaguchi, H. Abe, M. Shigeno, T. Ushiki. Observation of human corneal and scleral collagen fibrils by atomic force microscopy. *Japan J. Ophthalmol.* **46**, 496–501(2002).
88. N.M. Ziebarth, E.P. Wojcikiewicz, F. Manns, V.T. Moy, J.-M. Parel. Atomic force microscopy measurements of lens elasticity in monkey eyes. *Mol. Vis.* **13**, 504–510(2007).
89. N.A. Burnham, O.P. Behrend, F. Oulevey, G. Gremaud, P.J. Gallo, D. Gourdon, E. Dupas, A.J. Kulik, H.M. Pollock, G.A.D. Briggs, How does a tip tap? *Nanotechnology* **8**, 67–75(1997).
90. J.H. Hoh, R. Lal, S.A. John, J.P. Revel, M.F. Arnsdorf, Atomic force microscopy and dissection of gap junctions. *Science* **253**, 1405–1408(1991).
91. J.H. Hoh, G.E. Sosinsky, J.P. Revel, P.K. Hansma, Structure of the extracellular surface of the gap junction by atomic force microscopy. *Biophys. J.* **65**, 149–163(1993).

92. P.K. Hansma, J.P. Cleveland, M. Radmacher, D.A. Walters, P.E. Hillner, M. Bezanilla, M. Fritz, D. Vie, H.G. Hansma, C.B. Prater, J. Massie, L. Fukunaga, J. Gurley, V. Elings, Tapping mode atomic force microscopy in liquids. *Appl. Phys. Lett.* **64**, 1738–1740(1994).
93. E. Nagao, J.A. Dvorak, An integrated approach to the study of living cells by atomic force microscopy. *J. Microsc.-Oxf.* **191**, 8–19(1998).
94. M.A. Poggi, E.D. Gadsby, L.A. Bottomley, W.P. King, E. Oroudjev, H. Hansma, Scanning probe microscopy. *Anal. Chem.* **76**, 3429–3443(2004).
95. B. Cappella, G. Dietler, Force-distance curves by atomic force microscopy. *Surf. Sci. Rep.* **34**, 1(1999).
96. N.J. Tao, S.M. Lindsay, S. Lees, Measuring the microelastic properties of biological-material. *Biophys. J.* **63**, 1165–1169(1992).
97. P. Frederix, T. Akiyama, U. Staufer, C. Gerber, D. Fotiadis, D.J. Muller, A. Engel, Atomic force bio-analytics. *Curr. Opin. Chem. Biol.* **7**, 647(2003).
98. D.C. Lin, E.K. Dimitriadis, F. Horkay, Robust strategies for automated AFM force curve analysis-II: adhesion-influenced indentation of soft, elastic materials. *J. Biomech. Eng.* **129**, 904–912(2007).
99. D.C. Lin, E.K. Dimitriadis, F. Horkay, Robust strategies for automated AFM force curve analysis-I. Non-adhesive indentation of soft, inhomogeneous materials. *J. Biomech. Eng.* **129**, 430–440(2007).
100. F. Obataya, C. Nakamura, S.W. Han, N. Nakamura, J. Miyake, Mechanical sensing of the penetration of various nanoneedles into a living cell using atomic force microscopy. *Biosens. Bioelectron.* **20**, 1652–1655(2005).
101. E.P. Wojcikiewicz, X. Zhang, V.T. Moy, Force and compliance measurements on living cells using atomic force microscopy (AFM). *Biol. Proc. Online* **6**, 1–9(2004).
102. E.P. Wojcikiewicz, X. Zhang, A. Chen, V.T. Moy, Contributions of molecular binding events and cellular compliance to the modulation of leukocyte adhesion. *J. Cell Sci.* **116**, 2531–2539(2003).
103. C.Y. Zhang, Y.W. Zhang, Computational analysis of adhesion force in the indentation of cells using atomic force microscopy. *Phys. Rev.* **77**, 021912(2008).
104. N.M. Ziebarth, E.P. Wojcikiewicz, F. Manns, V.T. Moy, J.M. Parel, Atomic force microscopy measurements of lens elasticity in monkey eyes. *Mol. Vis.* **13**, 504–510(2007).
105. Y.F. Dufrene, P. Hinterdorfer, Recent progress in AFM molecular recognition studies. *Pflugers Arch.* **456**, 237–245(2008).
106. V. Dupres, C. Verbelen, Y.F. Dufrene, Probing molecular recognition sites on biosurfaces using AFM. *Biomaterials* **28**, 2393–2402(2007).
107. C. Verbelen, N. Christiaens, D. Alsteens, V. Dupres, A.R. Baulard, Y.F. Dufrene, Molecular mapping of lipoarabinomannans on mycobacteria. *Langmuir* (2009).
108. M. Gerhard, M.A. Nabil, Novel optical approach to atomic force microscopy. *Appl. Phys. Lett.* **53**, 1045–1047(1988).
109. S. Alexander, L. Hellemans, O. Marti, J. Schneir, V. Elings, P. K. Hansma, M. Longmire, J. Gurley, An atomic-resolution atomic-force microscope implemented using an optical-lever. *J. Appl. Phys.* **65**, 164–167(1989).
110. M.A. Horton, G.T. Charras, G. Ballestrem, P.P. Lehenkari, Integration of atomic force and confocal microscopy. *Single Mol.* **1**, 135–137(2000).

111. J.L. Choy, S.H. Parekh, O. Chaudhuri, A.P. Liu, C. Bustamante, M.J. Footer, J.A. Theriot, D.A. Fletcher, Differential force microscope for long time-scale biophysical measurements. *Rev. Sci. Instrum.* **78**, 043711(2007).
112. V. Heinrich, C. Ounkomol, Force versus axial deflection of pipette-aspirated closed membranes. *Biophys. J.* **93**, 363–372(2007).
113. C. Ounkomol, H. Xie, P.A. Dayton, V. Heinrich, Versatile horizontal force probe for mechanical tests on pipette-held cells, particles, and membrane capsules. *Biophys. J.* **96**, 1218–1231(2009).
114. P.P. Lehenkari, G.T. Charras, A. Nykanen, M.A. Horton, Adapting atomic force microscopy for cell biology. *Ultramicroscopy* **82**, 289–295(2000).
115. P.H. Puech, K. Poole, D. Knebel, D.J. Muller, A new technical approach to quantify cell-cell adhesion forces by AFM. *Ultramicroscopy* (2006).
116. T. Ando, T. Uchihashi, N. Kodera, D. Yamamoto, A. Miyagi, M. Taniguchi, H. Yamashita, High-speed AFM and nano-visualization of biomolecular processes. *Pflugers Arch.* **456**, 211–225(2008).
117. D.J. Muller, A. Engel, J.L. Carrasosa, M. Velez, The bacteriophage phi29 head-tail connector imaged at high resolution with the atomic force microscope in buffer solution. *EMBO J.* **16**, 2547–2553(1997).
118. D.J. Muller, D. Fotiadis, S. Scheuring, S.A. Muller, A. Engel, Electrostatically balanced subnanometer imaging of biological specimens by atomic force microscope. *Biophys. J.* **76**, 1101–1111(1999).
119. D.J. Muller, K.T. Sapra, S. Scheuring, A. Kedrov, P.L. Frederix, D. Fotiadis, A. Engel, Single-molecule studies of membrane proteins. *Curr. Opin. Struct. Biol.* **16**, 489–495(2006).
120. G.T. Charras, M.A. Horton, Determination of cellular strains by combined atomic force microscopy and finite element modeling. *Biophys. J.* **83**, 858–879(2002).
121. M.H. Abdulreda, V.T. Moy, Atomic force microscope studies of the fusion of floating lipid bilayers. *Biophys. J.* **92**, 4369–4378(2007).
122. I. Obataya, C. Nakamura, S. Han, N. Nakamura, J. Miyake, Nanoscale operation of a living cell using an atomic force microscope with a nanoneedle. *NanoLett.* **5**, 27–30(2005).
123. L. Lu, S.J. Oswald, H. Ngu, F.C. Yin, Mechanical properties of actin stress fibers in living cells. *Biophys. J.* **95**, 6060–6071(2008).
124. E.K. Dimitriadis, F. Horkay, J. Maresca, B. Kachar, R.S. Chadwick, Determination of elastic moduli of thin layers of soft material using the atomic force microscope. *Biophys. J.* **82**, 2798–2810(2002).
125. K.D. Costa, A.J. Sim, F.C. Yin, Non-Hertzian approach to analyzing mechanical properties of endothelial cells probed by atomic force microscopy. *J. Biomech. Eng.* **128**, 176–184(2006).
126. K.D. Costa, F.C.P. Yin, Analysis of indentation: implications for measuring mechanical properties with atomic force microscopy. *J. Biomech. Eng.* **121**, 462–471(1999).
127. H. Hertz, On the contact of elastic bodies. In *Hertz's Miscellaneous Papers* (Macmillan, London, 1881) pp. 146–162.
128. M. Radmacher, M. Fritz, C.M. Kacher, J.P. Cleveland, P.K. Hansma, Measuring the viscoelastic properties of human platelets with the atomic force microscope. *Biophys. J.* **70**, 556–567(1996).

129. J.R. Barber, D.A. Billings, An approximate solution for the contact area and elastic compliance of a smooth punch of arbitrary shape. *Int. J. Mech. Sci.* **32**, 991–997(1990).
130. E.M. Darling, S. Zauscher, J.A. Block, F. Guilak, A thin-layer model for viscoelastic, stress-relaxation testing of cells using atomic force microscopy: do cell properties reflect metastatic potential? *Biophys. J.* **92**, 1784–1791(2007).
131. B.A. Smith, B. Tolloczko, J.G. Martin, P. Grutter, Probing the viscoelastic behavior of cultured airway smooth muscle cells with atomic force microscopy: stiffening induced by contractile agonist. *Biophys. J.* **88**, 2994–3007(2005).
132. H.F. Edelhauser, J.L. Ubels, Cornea and Sclera. In *Adler's Physiology of the Eye*, edited by P. Kaufman, A. Alm (Mosby, Inc., St. Louis, MO, 2003).
133. F.E. Fantes, G.O. Waring, Effect of excimer laser radiant exposure on uniformity of ablated corneal surface. *Lasers Surg. Med.* **9**(6), 533–542(1989).
134. T. Møller-Pedersen, H.D. Cavanagh, W.M. Petroll, J.V. Jester, Stromal wound healing explains refractive instability and haze development after photorefractive keratectomy. *Ophthalmology.* **107**, 1235–1245(2000).
135. A. Keirl, C. Christie. *Clinical Optics and Refraction: A Guide for Optometrists, Contact Lens Opticians and Dispensing Opticians.* (Elsevier Health Sciences, New York, 2007).
136. E. Chalupa, H.A. Swarbrick, B.A. Holden, J. Sjöstrand, Severe corneal infections associated with contact lens wear. *Ophthalmology.* **94**(1), 17–22(1987).
137. S. Bhatia, E.P. Goldberg, J.B. Enns, Examination of contact lens surfaces by atomic force microscope (AFM). *CLAO J.* **23**(4), 264–269(1997).
138. J.M. González-Méijome, A. López-Aleman, J.B. Almeida, M.A. Parafita, M.F. Refojo, Microscopic observation of unworn siloxane–hydrogel soft contact lenses by atomic force microscopy. *J. Biomed. Mater. Res. Part B: Appl. Biomater.* **76B**(2), 412–418 (2005).
139. J. Baguet, F. Sommer, T.M. Duc, Imaging surfaces of hydrophilic contact lenses with the atomic force microscope. *Biomaterials.* **14**(4), 279–284(1993).
140. J. Baguet, F. Sommer, V. Claudon-Eyl, T.M. Duc, Characterization of lacrymal component accumulation on worn soft contact lens surfaces by atomic force microscopy. *Biomaterials.* **16**(1), 3–9(1995).
141. E.P. Goldberg, S. Bhatia, J.B. Enns, Hydrogel contact lens–corneal interactions: a new mechanism for deposit formation and corneal injury. *CLAO J.* **23**(4), 243–248(1997).
142. V. Guryca, R. Hobzová, M. Prádný, J. Sirc, J. Michálek. Surface morphology of contact lenses probed with microscopy techniques. *Cont. Lens Anterior Eye.* **30**(4), 215–222(2007).
143. C. Maldonado-Codina, N. Efron, Impact of manufacturing technology and material composition on the surface characteristics of hydrogel contact lenses. *Clin. Exp. Optom.* **88**(6), 396–404(2005).
144. C.E. Rabke, P.L. Valint, D.M. Ammon, Ophthalmic applications of atomic force microscopy. *ICLC* **22**, 32–41(1995).
145. J.H. Teichroeb, J.A. Forrest, V. Ngai, J.W. Martin, L. Jones, J. Medley, Imaging protein deposits on contact lens materials. *Optom. Vis. Sci.* **85**(12), 1151–1164(2008).
146. G.L. Grobe, P.L. Valint, D.M. Ammon, Surface chemical structure for soft contact lenses as a function of polymer processing. *J. Biomed. Mater. Res.* **32**(1), 45–54(1996).

147. R.P. Santos, T.T. Arruda, C.B. Carvalho, V.A. Carneiro, L.Q. Braga, E.H. Teixeira, F.V. Arruda, B.S. Cavada, A. Havt, T.M. de Oliveira, G.A. Bezerra, V.N. Freire, Correlation between *Enterococcus faecalis* biofilms development stage and quantitative surface roughness using atomic force microscopy. *Microsc. Microanal.* **14**(2), 150–158(2008).
148. G.M. Bruinsma, M. Rustema-Abbing, J. de Vries, H.J. Busscher, M.L. van der Linden, J.M. Hooymans, H.C. van der Mei, Multiple surface properties of worn RGP lenses and adhesion of *Pseudomonas aeruginosa*. *Biomaterials* **24**(9), 1663–1670(2003).
149. A. Opdahl, S.H. Kim, T.S. Koffas, C. Marmo, G.A. Somorjai, Surface mechanical properties of pHEMA contact lenses: viscoelastic and adhesive property changes on exposure to controlled humidity. *J. Biomed. Mater. Res. A.* **67**(1), 350–356(2003).
150. S.H. Kim, C. Marmo, G.A. Somorjai, Friction studies of hydrogel contact lenses using AFM: non-crosslinked polymers of low friction at the surface. *Biomaterials* **22**(24), 3285–3294(2001).
151. S.H. Kim, A. Opdahl, C. Marmo, G.A. Somorjai, AFM and SFG studies of pHEMA-based hydrogel contact lens surfaces in saline solution: adhesion, friction, and the presence of non-crosslinked polymer chains at the surface. *Biomaterials* **23**(7), 1657–1666(2002).
152. L.A. Remington, *Clinical Anatomy of the Visual System* (Butterworth-Heinemann, Boston, 1998).
153. R.C. Augusteyn, Growth of the lens: in vitro observations. *Clin. Exp. Optometry* **91**(3), 226–239(2008).
154. S.K. Pandey, J. Thakur, L. Werner, M.E. Wilson, L.P. Werner, A.M. Izak, D.J. Apple, The human crystalline lens, ciliary body, and zonules: their relevance to presbyopia. In *Presbyopia: A Surgical Textbook*, edited by A. Agarwal (Slack Incorporated, Thorofare, NJ, 2002).
155. V.L. Taylor, K.J. Al-Ghoul, C.W. Lane, V.A. Davis, J.R. Kuszak, M.J. Costello, Morphology of the normal human lens. *Invest. Ophthalmol. Visual Sci.* **37**(7), 1396–1410(1996).
156. D.C. Beebe, The lens. In *Adler's Physiology of the Eye*, edited by P. Kaufman, A. Alm (Mosby, Inc., St. Louis, MO, 2003).
157. A. Shiels, S. Bassnett, Mutations in the founder of the MIP gene family underlie cataract development in the mouse. *Nat. Genet.* **12**(2), 212–215(1996).
158. E.F. Fincham. The mechanism of accommodation. *Br. J. Ophthalmol. Monograph Supplement VIII* (1937).
159. J. Kessler, Experiments in refilling the lens. *Arch. Ophthalmol.* **71**, 412–417(1964).
160. J. Kessler, Refilling the rabbit lens. Further experiments. *Arch. Ophthalmol.* **76**(4), 596–598(1966).
161. R.F. Fisher, Elastic constants of the human lens capsule. *J. Physiol.* **201**, 1–19(1969).
162. E.F. Fincham, The changes in the form of the crystalline lens in accommodation. *Trans. Opt. Soc.* **26**, 239–269(1925).
163. B. Gilmartin, The aetiology of presbyopia: a summary of the role of lenticular and extralenticular structures. *Ophthalm. Physiol. Opt.* **15**(5), 431–437(1995).
164. M. Tscherning. Le mecanisme de l'accommodation. *Annals Oculiat.* **131**, 168–179(1904).

165. S. Krag, T. Olsen, T.T. Andreassen, Biomechanical characteristics of the human anterior lens capsule in relation to age. *Invest. Ophthalmol. Visual Sci.* **38**, 357–363(1997).
166. K.R. Heys, S.L. Cram, R.J.W. Truscott, Massive increase in the stiffness of the human lens nucleus with age: the basis for presbyopia? *Mol. Vis.* **10**, 956–963(2004).
167. K.R. Heys, R.J. Truscott, The stiffness of human cataract lenses is a function of both age and the type of cataract. *Exp. Eye Res.* **86**(4), 701–703(2008).
168. H.A. Weeber, G. Eckert, F. Soergel, C.H. Meyer, W. Pechhold, R.G.L. van der Heijde. Dynamic mechanical properties of human lenses. *Exp. Eye Res.* **80**, 425–434(2005).
169. F. Soergel, C. Meyer, G. Eckert, B. Abele, W. Pechhold, Spectral analysis of viscoelasticity of the human lens. *J. Refract. Surg.* **15**, 714–716(1999).
170. H.A. Weeber, G. Eckert, W. Pechhold, R.G.L. van der Heijde, Stiffness gradient in the crystalline lens. *Graefe's Arch. Clin. Exp. Ophthalmol.* **245**, 1357–1366(2007).
171. R.A. Schachar, B.K. Pierscionek, Lens hardness not related to the age-related decline of accommodative amplitude. *Mol. Vis.* **13**, 1010–1011(2007).
172. M. Lombardo, M.P. De Santo, G. Lombardo, R. Barberi, S. Serrao, Analysis of intraocular lens surface properties with atomic force microscopy. *J. Cataract Refract. Surg.* **32**, 1378–1384(2006b).
173. Y. Ohnishi, T. Yoshitomi, T. Sakamoto, K. Fujisawa, T. Ishibashi. Evaluation of cellular adhesions on silicone and poly(methyl methacrylate) intraocular lenses in monkey eyes; an electron microscopic study. *J. Cataract Refract. Surg.* **27**, 2036–2040(2001).
174. C. Cassinelli, M. Morra, A. Pavesio, D. Renier. Evaluation of interfacial properties of hyaluronan coated poly(methylmethacrylate) intraocular lenses. *J. Biomater. Sci. Polym. Edn.* **11**(9), 961–977(2000).
175. D. Bozukova, C. Pagnouille, M.C. De Pauw-Gillet, S. Desbief, R. Lazzaroni, N. Ruth, R. Jrme, C. Jrme. Improved performances of intraocular lenses by poly(ethylene glycol) chemical coatings. *Biomacromolecules.* **8**(8), 2379–2387(2007).
176. M. Dogru, K. Tetsumoto, Y. Tagami, K. Kato, K. Nakamae. Optical and atomic force microscopy of an explanted AcrySof intraocular lens with glistenings. *J. Cataract Refract. Surg.* **26**(4), 571–575(2000).
177. R.K. Sharma, B.E.J. Ehinger. Development and structure of the retina. In *Adler's Physiology of the Eye*, edited by P. Kaufman, A. Alm. (Mosby, Inc., St. Louis, MO, 2003).
178. A. Hendrickson, Organization of the adult primate fovea. In *Macular Degeneration*, edited by P.L. Penfold, J.M. Provis (Springer, Berlin, 2005).
179. P.L. Penfold, J. Wong, D. van Driel, J.M. Provis, M.C. Madigan, Immunology and age-related macular degeneration. In *Macular Degeneration*, edited by P.L. Penfold, J.M. Provis (Springer, Berlin, 2005).
180. G. Wu, *Retina: The Fundamentals* (W.B. Saunders, St. Louis, MO, 1995).
181. W. Halfter, M. Willem, U. Mayer, Basement membrane-dependent survival of retinal ganglion cells. *Invest. Ophthalmol. Visual Sci.* **46**, 1000–1009(2005).
182. R.F. Fisher, The elastic constants of the human lens. *J. Physiol.* **212**, 147–180(1971).

Force-Extension and Force-Clamp AFM Spectroscopies in Investigating Mechanochemical Reactions and Mechanical Properties of Single Biomolecules

Robert Szoszkiewicz

Summary. In every biological system, mechanical forces are present from the molecular scales, to the cellular levels, to the single-molecule level. In this chapter we describe current experimental methods of measuring forces and displacements of biological objects at the molecular length scales. To start with, we quickly review some key experimental techniques like the centroid tracking, FRET, magnetic and optical tweezers, and the atomic force microscopy (AFM) methods. We focus on the AFM methods, and namely on the force-extension (FX) and the force-clamp (FC) AFM-based spectroscopies. These methods are able to inform about conformational changes of the biomolecules subjected to a controlled stretching and/or compression. We discuss the limitations of displacement and force resolution and sensitivity, respectively. We mention some important technicalities in the FX and FC studies like the methods of fingerprinting the stretched biomolecules, and the ways of optimizing the AFM systems. Finally, we discuss the applications of the FX and FC methods in probing mechanical stability of proteins and polysaccharides, as well as using proteins as matrices to study disulfide reductions at the single molecule level. We close by discussing some shortcoming of the presented AFM-based spectroscopies.

Key words: force control, displacement control, nano-scale objects, biomolecules, proteins, atomic force microscopy, force-extension spectroscopy, force-clamp spectroscopy

Abbreviations

aa	Amino acid; a building block of proteins
AFM	Atomic force microscopy
BW	Bandwidth
FC-AFM	Force-clamp AFM spectroscopy
FRET	Forster resonance energy transfer; a local probe fluorescence technique

FX-AFM	Force extension AFM spectroscopy
Ig	Immunoglobulin (usually a small < 100 aa protein)
mechanochemical reactions	broadly defined class of chemical reactions initiated by force. In the context of this chapter, the mechanochemical reactions include (un)folding of single protein molecules, stretching/breaking single bonds (i.e., disulfide bonds), or force initiated conformational changes of larger part(s) of molecules, like chair-boat transitions of the glucopyranose ring in polysaccharides
MT	Magnetic tweezers
NMR	Nuclear magnetic resonance
OT	Optical tweezers
PSPD	Photosensitive photodiode
resolution (of force or displacement)	The minimum change of force or displacement, respectively, which can be detected.
rms	Root mean square
sensitivity (of force or displacement)	The minimum value of force or displacement, respectively, which can be detected.
WLC	The worm-like chain; a model of the polymer elasticity

12.1 Introduction

In every biological system, mechanical forces are inherently present from the macroscopic scales to the cellular levels, and furthermore to the single-molecule level. Mechanical forces are involved in the development and maintenance of bones, blood vessels, muscles, regulation of blood pressure; motility of cells; regulation of cell proliferation, and cell apoptosis; and numerous processes associated with the lifecycle of a protein [1–5].

In this chapter, we concentrate on technical details and list current advances in force-extension (FX) and force-clamp (FC) atomic force microscopy (AFM) spectroscopies in investigating mechanical forces and displacements at the single-molecule level for biomolecules. First, we list current experimental techniques for quantitative measurements of forces and displacements in the single-molecule level. Next, we discuss some limitations and issues that arise when measuring such forces and displacements. Then, we describe pertinent experimental details of FX/FC-AFM techniques. Eventually, we move

toward the particular experimental implementations of the FX- and FC-AFM and discuss their shortcomings.

12.2 Experimental Techniques for Measuring Displacements and Forces at the Single Molecule Level

The single-molecule spectroscopy methods can be divided into motion- and force-sensitive techniques. The most popular motion-sensitive techniques include centroid tracking and Forster (or fluorescence) resonance energy transfer (FRET). The motion-sensitive techniques provide only information about displacements within a data-collection window, but the displacements can be calibrated into forces if effective spring constants between the moving parts are known.

Strictly force-sensitive techniques include magnetic tweezers (MTs), optical traps (OTs), and FX/FC-AFM modes. These techniques obtain calibrated forces from measuring the corresponding displacements. One approach is to use the thermal noise methods essentially measuring Brownian motion of the force probe (e.g., a trapped bead, an AFM cantilever). The probe stiffness is then determined either using an equipartition theorem and/or through the roll-off frequency of the thermal spectrum [3]. Alternatively, known calibrating forces are applied by a reference AFM cantilever or a viscous drag from hydrodynamic flow. All of these approaches have systematic errors that can affect the calibration [6].

Each of the local spectroscopy techniques operate in different spatial, temporal, and force ranges, and may require particular chemistries, as has been recently summarized by Greenleaf et al. [7] and earlier by numerous authors, for example, [8]. Majority of these studies have been done *in vitro*. We briefly discuss these methods in the following text.

12.2.1 Centroid Tracking

Centroid tracking uses a micrometer-sized reporter particle, or a fluorescence dye molecule, which is tagged to the target molecule. The motion of the tagged particle/molecule is tracked by various implementations of optical/fluorescence microscopes. Centroid tracking refers to measuring optical center of mass of the particle/fluorophore or fluorescence from a single fluorophore. Depending on a particular experiment, centroid tracking reports either about rotational dynamics of the flexible linker or about displacement of (and sometimes forces exerted on) a labeled particle. Variety of optical imaging techniques has been used: from open field to confocal imaging. The size, contrast range, and quality of imaging systems determine the spatial resolution of this method, which is typically a fraction of a micrometer. For small and highly efficient fluorophores, long time recording of single molecule fluorescence allows reconstruction of the spatial distribution of the photon

counts. The photon count histogram versus lateral position of a probe is then fitted with an appropriate function, for example, Airy function or a Gaussian. Thus, the center of a histogram is localized with a nanometer resolution provided that a system has high numerical aperture and does not drift over the acquisition time. However, due to the small size of the imaged objects, the dispersive optical transfer functions (through lenses etc.) need to be carefully considered.

Temporal resolution of the centroid tracking is limited either instrumentally by an optical capture rate, or by a relaxation time constant of a particle moving in a medium. When a single particle fluorescence distribution is recorded, the overall measurement time increases up to tens of seconds. A great variety of reporter particles has been used: polymer or glass (fluorescence) beads, gold or other types of nanoparticles, fluorescent actin filaments, or microtubules [9,10]. Larger beads are easier to detect, but they can also perturb the system. In some structural applications, the beads (or fluorophores) need to be attached covalently, which requires significant structural knowledge of the target molecule to minimize steric hindrances and fluorescence quenching. Also, oxygen-scavenging systems are often necessary, since reactive oxygen species can rapidly bleach most of the fluorophores [11]. An efficient dye should have high fluorescence efficiency and a long lifetime. Quantum dots are among the most efficient dyes, but they tend to blink. To avoid disintegration or unwanted changes of the probed molecules, dye absorption/emission wavelengths should be carefully considered.

Centroid tracking has been applied to measure the angular motion, for example, dynamics of rotary motors like F_1 -ATPase and the bacterial flagellar motor employed by some bacteria to generate motion [12–15]. Other studies have investigated dynamics of motor proteins such as kinesin and dynein walking along a microtubule, and several types of myosins walking along actin filaments [16–23].

12.2.2 Fluorescence Resonance Energy Transfer

In FRET, sometimes also called Forster resonant energy transfer, one fluorophore (donor) is excited, and the fluorescence emitted from a coupled fluorophore (acceptor) is measured [24, 25]. FRET pertains to nonradiative donor–acceptor energy transfer via so-called Forster dipole–dipole interactions [26, 27]. The nonradiative energy transfer is extremely sensitive to the donor–acceptor distance and orientation. For freely rotating dyes, the orientational component averages out, but in some situations it needs to be estimated, which is not trivial. For freely rotating dyes, the FRET energy transfer efficiency changes with the six-power of their distance. Thus, FRET is essentially a threshold detector of subnanometer motion probed only within a limited range of several nanometers around an equilibrium fluorophore separation R_0 (typically several nanometer). Many molecular distances and/or

conformational changes are within the FRET range, which makes this technique attractive for many biological applications. However, any estimations of forces associated with molecular dynamics probed by FRET are up-to-now elusive, and we will not elaborate on it.

FRET has been used extensively in a great variety of contexts. These include the effects of calcium binding on calmodulin [28], the folding pathway of cold-shock proteins [29], the migration of Holliday junctions [30], the folding of telomeres [31], transcription initiation [32], the dynamics of DNA, RNA, and ribosomes [33–35].

12.2.3 Magnetic Tweezers

MTs utilize a magnetic field to manipulate the position of a superparamagnetic bead in order to stretch and/or twist the molecules attached to the bead. The bead is usually composed of highly cross-linked polystyrene with evenly distributed magnetic tracers, e.g., iron oxides. Relatively strong and uniform magnetic fields generate and saturate a net magnetic dipole moment in a bead. The particle experiences a magnetic force proportional to the saturated magnetic dipole moment of a bead and to a magnetic field gradient.

Lateral motion of the bead can be probed by centroid tracking, and axial (normal) motions can be tracked by imaging the optical interference fringes from interference between unscattered and bead-scattered light [36]. The MT traps utilize either movable magnets, or permanent magnets, or magnetic coils with soft-iron cores. With respect to optical tweezers (OTs) (next paragraph), the MT setups obviously minimize any sample damage from local heating [37].

MTs can exert calibrated forces between 0.1 pN and ~ 20 pN. The upper force limit depends on a choice of a bead (its material/dimension) and magnetic impurities. Since magnetic beads act as dipoles and have a preferential orientation with respect to the magnetic field, the traps can also apply torques by changing the direction of the magnetic field. To visually detect the angular motion, however, the bead must have a nonspherical, for example, oblique, shape.

MT setups have been primarily used in biology to harness the formation of plectonemes (aka supercoils) in DNA, and to figure out the some dynamics of nucleic acid enzymes, namely the motion of the DNA translocases and topoisomerases [38–46].

12.2.4 Optical Traps

OTs also called “optical gradient traps” were pioneered by Ashkin et al. [47]. In this approach, a polymeric bead (to which a molecule of interest is attached) is trapped in a center of a tightly focused light beam. The other end of the molecule of interest is tethered to a surface or to a second bead. While the detailed theory behind the OT is complicated, there are two main contributions to the optical force exerted on a polymeric sphere: the gradient force

and the scattering force. The gradient force arises due to gradients in the light intensity as the bead develops an electric dipole moment in response to the light's electric field. The scattering force is due to scattering of light by the bead. If the dielectric permittivity of the molecule is greater than the dielectric permittivity of the medium, the gradient force attracts the bead to a focal point of an incident light beam.

To achieve significant gradient forces, tightly focused laser beams need to be produced using high numerical aperture objectives. This, however, can provoke substantial heat dissipation in the sample. Therefore, the wavelengths used for optical trapping are generally in the near infrared – an optically transparent region for most biological specimen [37]. The scavenging systems are often employed to minimize side effects caused by photoreactive species, and namely oxygen. The forces range in OT is ~ 0.1 – 100 pN. The force can be modulated either by adjusting laser intensity, or can be kept constant (“clamped”) and applied to drag the bead. Several distinct OT assemblies have appeared in literature: the surface–dumbbell assay, the micropipette–dumbbell assay, and finally, the dual trap (dumbbell–dumbbell) assembly, which is the most stable.

The time sensitivity of the dual trap can be as low as 0.1 ms, which currently makes OT the most sensitive technique for local force spectroscopy [48, 49]. Furthermore, some hybrid OTs have appeared combining single-molecule fluorescence with single-molecule manipulation [50, 51]. The OT potential well is shallow compared to the magnitude of the nonspecific interaction forces between the two beads (or between the bead and the surface). Thus, the OT techniques often require micrometer-size molecular handles, which complicate probing of the sub-Angstrom scale motions.

The OT setups can apply torques. Circularly polarized light or azimuthally phase modulated beams (like Laguerre-Gaussian beams) and combinations of these introduce an orbital angular momentum component in the light beam [52–54]. One example is the holographic OTs. In such setups, a phase of an incident light beam is modulated (like in a hologram) by initial diffraction on a grating prepared by an orientationally programmable matrix of liquid crystals. Thus, superpositions of arbitrary beams can be instantaneously created, which can produce any type of motion of the trapped bead. Further examples include the use of optical wrenches. Optical wrenches apply torques by either using external alternating magnetic fields acting on optically trapped magnetic particles [55], or using polarized light to orient birefringent, for example, calcite or quartz, particles [56].

A wide variety of biological problems have been investigated by the OT traps. These include mechanisms of motion of kinesin [19, 23, 57–59], myosin [20, 21, 60, 61], and dynein [22, 62], as well as processive nucleic acid enzymes such as exonucleases and helicases [63, 64], DNA translocases [65], and DNA and RNA polymerases [61, 66–70]. Further examples include viral DNA packing [71, 72] and ATPases [73]; elucidating elastic properties of nucleic acids

[74–76]; protein–DNA interactions [77]; folding of nucleic acids [48, 78–80]; and folding of proteins [81, 82].

12.2.5 Single Molecule AFM Force Spectroscopy

The force-spectroscopy mode of an AFM has been developed to measure properties of materials at local scales [83], and very quickly has found its applications in biology. Numerous examples include measuring binding forces of complementary DNA strands [84], forces in receptor–ligand systems [85], forces maintaining sugar rings conformations in polysaccharides [86, 87], and unfolding forces of multidomain proteins [88] as well as the strengths of some covalent bonds [88, 93–105]. In the AFM force spectroscopy, a single biomolecule is clamped between an AFM tip and a substrate. A substrate is usually placed atop an AFM piezo-scanner for precise control of the tip–sample distance (Fig. 12.1). The tip–sample forces are usually monitored by a standard laser beam bounce method using the laser beam, which is reflected on the

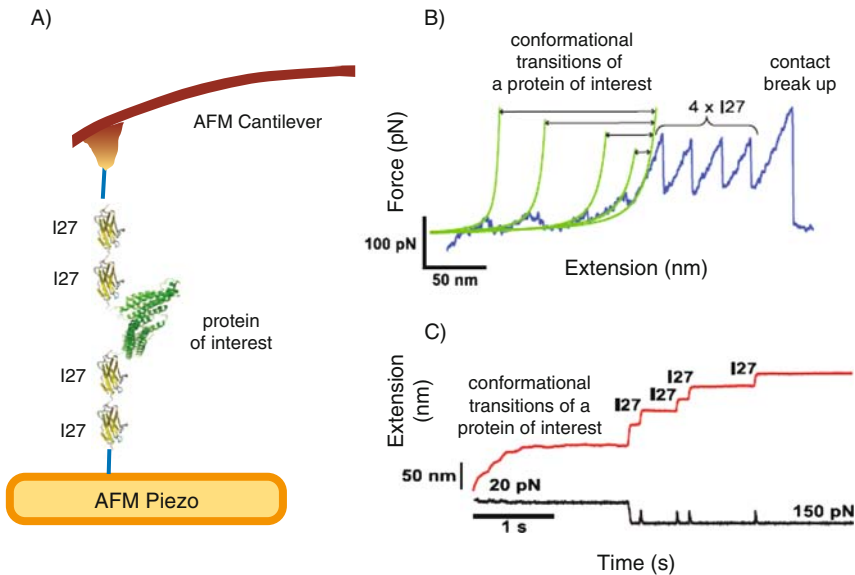


Fig. 12.1. Principles of the single-molecule AFM spectroscopy. (a) The I27₂ – protein of interest – I27₂ construct is clamped between a cantilever and a substrate (placed atop a piezo positioner). (b) A force–extension curve (“saw-toothed pattern”), main curve. Five WLC fits to the conformational changes within a protein of interest are followed by four unfolding I27 peaks and a contact break-up peak. (c) A schematic force-clamp trace. The top and bottom curves represent extension and force, respectively. I27 is the 27th immunoglobulin domain of a human cardiac titin with well-characterized mechanical properties [89]. The I27 handles provide a controlled way of delivering the mechanical force to the protein of interest. Adapted from [90]

back side of a cantilever and captured by a photosensitive photodiode (PSPD). The properties of each AFM component (scanner, PSPD, electronics, and cantilever) define available force and displacement sensitivity as well as resolution and bandwidth (BW) (millisecond to tens of seconds) of the AFM force spectroscopy, which is discussed in the following chapters. Currently, forces can be measured with several piconewton (pN) force sensitivity, that is, sensitivity comparable to the OTs [91,92], and with sub-Angstrom length resolution.

12.3 Displacement and Force as Control Parameters in Small Systems

Depending on the rules guiding the available states, statistical mechanics describe macroscopic systems in terms of various “ensembles”. The ensembles, being a collection of the allowed multistates, are fully specified by temperature, pressure, and chemical potential. Each ensemble yields the same equation of state in the large-volume or thermodynamic limit. For small systems, however, the equation of state and the fluctuation spectra can be determined by the end-to-end distance of a stretched molecule as a function of the external force [106,107]. Alternatively, the distance is a natural reaction coordinate, and the ensemble of states is often described in terms of so-called energy landscape, that is, a collection of Arrhenius-like activation energy barriers versus distance. The apparent energy barrier is the energy barrier for a spontaneous process, that is, without an external force. The energy landscape must encompass a multiatom system with many internal degrees of freedom and its complicated normal mode spectrum. Thus, the energy landscapes are multidimensional.

Any experimentally observed value of force or displacement arises from a projection of the force-modified energy landscape on a pulling coordinate. The simplest and widely used approximation treats a force as a harmonic perturbation linearly tilting the apparent energy landscape. Application of external force can change the heights of energy minima and maxima, and even stabilize unfavorable points on the energy landscape. Thus, selected portion of the energy landscape along a pulling coordinate can be often investigated.

How accurately, however, the forces and displacements can be measured on the single molecule level? We address these issues in the following two sections, but we note that to attain any superior detection of molecular motion and forces, the detector itself must reduce in size as well [50].

12.3.1 Displacement Sensitivity and Resolution

The tiniest possible amount of displacement, or displacement sensitivity, which can be detected for an object with the stiffness, k , is limited by thermal (noise) vibrations of such an object. Using an equipartition theorem, the rms (root mean square) displacement sensitivity in normal direction, Δx_{rms} , is $\Delta x_{\text{rms}} = (\sum_i \langle x^2 \rangle)^{1/2} = (\sum_i \langle (\Delta x)^2 \rangle)^{1/2} = (k_{\text{B}}T/k)^{1/2}$, where k_{B} is the Boltzmann’s

constant, K is the spring constant in normal direction, and T is the absolute temperature. In practical situations, however, the rms displacement sensitivity is greater, because motion is sampled only within a limited spectral range, that is, the bandwidth, BW. We approximate that our object is thermally driven in a harmonic potential, and has a Lorentzian power spectrum. Then, the corresponding Δx_{rms} becomes [7]

$$\Delta x_{\text{rms}} = \left(\frac{4bk_{\text{B}}TBW}{k^2} \right)^{1/2} \quad (12.1)$$

The formula above is only valid for BW below the so-called roll-off frequency (sometimes also referred to as corner or cutoff frequency) ω_0 ($BW \ll \omega_0$). The value of ω_0 is defined as $\omega_0 = k/b$, where “ b ” is the drag (or viscous friction) coefficient. Consequently, the displacement to noise-signal ratio can be enhanced by either (1) reducing the BW or by (2) reducing the drag, or by (3) increasing the stiffness of a molecule. Reduced BW might seriously limit the timescale of detectable processes, but it can be an elegant solution in the case of slow or quasi-static processes. Decreasing the drag coefficient of the system effectively transfers the noise in displacement to higher frequencies, where it averages out in the measurement. If only a force transducer (a bead, or a cantilever) was taken into account, the drag decrease could have been achieved using small physical dimensions or the transducer, and/or a particular material of a bead, and/or using a less viscous solution (if possible).

In reality, however, rheological properties of the whole system, that is, a single molecule force transducer (a bead, or cantilever, or some handles), need to be accounted for in our analysis. Often, a molecule-handle system can be regarded as simply two springs connected in series. Application of force results in the displacement of a molecule and the displacement of a handle (cantilever, bead, etc.). Depending on the experimental setup, the displacement of a handle either adds up or gets subtracted from the displacement of a molecule. If the latter, the displacement sensitivity gets reduced.

The overall stiffness of the molecule-handle in series is determined by the most compliant spring. An example is a stretched polymer (chain), which in the language of the polymer physics is described as a set of hinged, incompressible segments. If the segments are stiff, they are described by a freely jointed chain model. Bendable segments are described by a worm-like chain (WLC) model. Stretchable segments are described by variations of the aforementioned models and (enthalpic) stretching of the molecular bonds. All these models predict monotonic and nonlinear increase of stiffness with force [106]. Thus, small forces produce considerable stretching and correspond to low molecular stiffness and low displacement sensitivity of the molecule in question. On applying greater forces, the molecule becomes stiffer and its displacement sensitivity increases [due to (12.1)], so that such a molecule can be more precisely localized. At the same time, however, the overall displacement of a molecule-handle system is taken up mostly by the handle (or a bead, or a cantilever),

which reduces the displacement resolution of the molecule. This situation is analogous to a classic problem of indenting the stiff sample by a compliant cantilever. Any indentation depths are lost in the calibration, because most of the displacement has been taken up by a cantilever [108].

To get rid of the molecule compliance variations due to force, it is convenient to work in a FC mode, that is, to use constant forces applied to a molecule of interest via either active feedback [109] or passive methods [110]. The feedback response time (comprised the response of any mechanical parts, e.g., an AFM scanner) along with the roll-off frequency determine temporal resolution of the measured displacement, which sometimes conflicts with the timescale for the motions being studied.

12.3.2 Force Sensitivity and Resolution

Forces are usually obtained from displacements, so the force sensitivity, ΔF_{rms} , is determined by (12.1) in the following way:

$$\Delta F_{\text{rms}} = k\Delta x_{\text{rms}} = (4bk_{\text{B}}TBW)^{1/2} \quad (12.2)$$

To decrease the thermal noise and increase the force sensitivity, one ideally needs either (1) to decrease the BW or (2) to decrease the viscous drag. To minimize the viscous drag, one can use the compliant probes with a large roll-off frequency. For the AFM cantilevers, this entails use of compliant levers with a large resonance frequency. Such conditions are satisfied for short and thin cantilevers, since the cantilever's spring constant is proportional to a cube of a ratio of the cantilever's thickness over length, and the cantilever's free resonance frequency scales with a ratio of the spring constant over length. Short AFM cantilevers (or in general small probes) usually also have faster response times. In order to be sensitive to piconewton forces, the cantilevers are also very compliant, for example, with typical spring constants ~ 10 pN/nm.

The force sensitivity (or noise floor) does not explicitly depend on the actual stiffness of a probe. Consequently, a soft force transducer is not a more sensitive detector than a stiff one if “ b ” is same for both: as k decreases, the noise increases exactly as fast as the sensitivity. In reality, the parameter “ b ” depends on the probe stiffness, as previously defined.

The force resolution (usually calibrated from the probe displacement measurements) depends on the overall stiffness of the probe-handle configuration.

12.4 AFM Force Spectroscopy with a Few Piconewton Sensitivity and at a Single Molecule Level

Obtaining good-quality single molecule AFM data takes some practice and some rules/protocols need to be addressed. In this section, we mention common issues associated with both FX and FC-AFM setups.

12.4.1 Fingerprinting the Biomolecules

When an AFM tip indents and later withdraws from a substrate covered with a layer of proteins (or biomolecules), many single molecules can interact with it. Rarely will the native biomolecule of interest be suspended between the substrate and the withdrawing tip. Usually, however, there will be ions, and many other molecules adsorbed to the tip or the surface, inevitably producing a complicated pattern of possible interactions and forces. These “nonspecific” interactions can easily mask forces and displacements pertaining to stretching the biomolecule of interest. Thus, additional fingerprinting is necessary to distinguish between nonspecific interactions and the interactions of interest, and to avoid skewing the obtained statistics. One possibility is to investigate only homo-modular proteins. Homo-modular proteins are multimer proteins with the same domain repeated in tandem. Naturally occurring homo-modular proteins produce repetitive patterns while stretched, for example, multiple immunoglobulin (Ig) and FN-III domains enabled detection of titin and tenascin unfolding [88, 111]. Similarly, small proteins can be expressed as homo-modular proteins [89, 99].

Some biomolecules, like some polysaccharides in particular, have unique fingerprint while being stretched, so there is no need to prepare them as multiple copies in tandem [87]. Also for some modular biomolecules, like green fluorescence protein [97, 112], no homomeric construct is needed if an X-ray or nuclear magnetic resonance (NMR) structures exist. The X-ray or NMR structures are not always available, though. And often a molecule of interest is a large multidomain entity, which is difficult to be cloned and expressed in a form of its multiple copies in tandem. Then, some chimeric constructs can be used that employ (1) the protein of interest and (2) the previously well-characterized tandem-repeat handles (e.g., Ig domains (Fig 12.1), polyubiquitin domains) [90].

Fingerprinting the biomolecules undergoing mechanochemical reactions can potentially be achieved using FX/FC-AFM spectroscopy in combination with micro-Raman, FRET, or various fluorescence techniques able to serve as additional distance and time “rulers.” Efficient coupling schemes between bio-oriented AFM spectroscopy and other techniques are just emerging. One example is “evanescence nanometry”, a technique that can detect the changes in length or position of a single molecule moving along the z -axis of an optical microscopy with nanometer and millisecond time resolution [113]. This approach uses a calibrated evanescence wave to measure the position of a fluorescent particle moving along the z -axis. Such evanescence wave is generated using the total internal reflection fluorescence (TIRF) in a custom built AFM-TIRF instrument.

12.4.2 Optimizing the AFM System

We have treated the issue of displacements and forces sensitivity in a combination of the cantilever-molecule system. However, there are other standard

sources of noise affecting the AFM measurements. These include acoustic couplings, thermal expansions, piezo creep, laser beam intensity variations, and optical interferences between the laser beam reflected from the lever and the sample. Acoustic noise primarily consists of the reflected acoustic noise. Mechanical noises decrease when all the parts are miniaturized, so proper instrument design is important. Acoustic noise and any mechanical noise are minimized by using several stages of mechanical/acoustic isolations including antivibrational tables, active noise cancellation systems, lowering the center of mass of the overall mechanical system, etc.

The noise of the laser beam is reduced by “cleaning” a laser beam from mixed modes via passing it through an optical fiber prior to sending such beam to a cantilever. The laser beam interference effects can be greatly reduced by using low coherence laser beams by introducing small high-frequency modulation of the laser light. The laser signal from a PSPD photodiode must also be divided by its own intensity to minimize any laser beam intensity variations. Tight focus of the incident laser beam on a cantilever minimizes interference. Low-power laser beam reduces cantilever heating. The cantilever heating effects are also diminished when using gold coated cantilevers dissipating heat quickly to the cantilever’s base. A PSPD photodiode (with large area segments and small gaps in between the segments) followed by a high bandwidth and a low noise preamplifier are essential in keeping the AFM photodiode noise below the level of the thermal noise. Compact liquid cells with O-rings reduce cantilever drift and buffer evaporation. Novel liquid cells with only minimal light path in the buffer solution decrease cantilever drift by reducing fluctuations of the refractive index of the buffer [91]. Although cantilever drift due to heating, creep, etc. is difficult to eliminate entirely, it can be partially accounted for by using a second reference cantilever.

The main factors for the overall noise and bandwidth are notably large size of the cantilever and slow feedback actuation time as determined by the combined performance of a piezo, cantilever, and the electronic PID system (approximately milliseconds). Both of these limitations can, however, be improved in the future.

12.5 FX-AFM Probes Mechanical Stability of Proteins and Polysaccharides

12.5.1 Details of the FX Trace

A sample FX trace is presented in Fig. 12.1, where a single biomolecule is stretched between the tip of a cantilever and a flat substrate. In typical setups, the substrate is mounted atop an AFM piezoscanner. In so-called inverse setups, the piezoscanner is located above the cantilever, and the substrate

stands still. A physiological solution of a biomolecule is deposited on a substrate, which is typically a gold coated glass cover slide. During the course of the FX-AFM experiment, the AFM tip taps the substrate surface to pick up and stretch single molecules. During each tap, the following experimental protocol is repeated. First, the AFM tip indents the substrate. If a biomolecule (or a portion of it) gets trapped between the tip and the sample, this point arbitrarily defines its zero extended end-to-end length. Next, the cantilever is retracted from the surface at a constant speed.

The resulting FX curve of a multimodular biomolecule has a characteristic saw-toothed pattern related to sequential stretching and unfolding of the modules. Initially, a whole molecule is being stretched and the detected force increases up to a peak value. At the force peak, the molecule in question looks more like a string with beads, since it is extended to its maximum length at this force. Then, either one of its modules unravels or the tip-sample (or the sample-substrate) contact breaks up. If unfolding happens, the force momentarily drops to the residual force as the cantilever relaxes during accommodating the available extra length. However, the AFM scanner still retracts at a constant speed. Soon, the molecule is stretched again and the described sequence of events begins to repeat itself. The last peak on a saw-toothed pattern marks the ultimate loss of contact.

Flat and freshly cleaned glass or gold coated glass standard microscope cover slides have proved to be the most effective substrates [99]. Low concentration protein solution keeps the probability of picking a single protein low (<1%), but increases the chances of picking only one molecule at a time.

The FX curves are calibrated in a standard way. The raw force (in V) is multiplied by a vertical sensitivity factor (nm/V). The vertical sensitivity factor is an inverse of a slope obtained from the FX curve in the tip-substrate contact region. Next, the force (in nm) is multiplied by a cantilever normal elastic spring constant, k , (pN/nm) to yield forces, F (in pN). The value of k is usually calculated by a thermal method [114–116]. Any calibrated scanner extension (in nm) is obtained by an independent sensing method, and usually a capacitive sensing. Calibrated extension of the biomolecule is obtained by subtracting the cantilever position change (F/k) from the scanner extension. Because of many potential pitfalls in the calibration, extremely precise low-hysteresis and low-creep scanners should be used, as well as preferentially V-shaped cantilevers due to their lower amount of coupling between torsional and normal displacements compared to the I-shaped levers.

12.5.2 What can be Inferred from the FX Trace?

The peak force reached before an unfolding event is a sensitive measure of the mechanical stability of a given module (of the multimer biomolecule). For a multimeric protein composed of several modules weighting ~ 10 kDa each, the typical values of the mean peak force can be as low as tens of piconewton (protein L) and as high as several hundreds of piconewton (I27 domain of

human cardiac titin); and with less than $\sim 10\%$ standard deviation due to stochastic nature of the unfolding events. The events pertaining to picking up several molecules at a time are easy to discriminate. Several molecules stretched in parallel increase the peak force value, while several molecules stretched in series produce uneven extension differences between the force peaks. A few other indicators of a clean FX trace are (1) a minimum force in between the peaks should increase slightly with the number of unfolded modules, since the overall protein length increases along with the entropy of the protein chain; and (2) the zero force baseline after the last peak should match the initial force plateau, which is measured before any tension has been applied to the molecule.

As mentioned earlier, the polymer physics models, like WLC or freely jointed chain, have been widely applied to fit the FX data. These models fit the relationships between the force and extension of the protein using two parameters: the persistence length (basically a measure of molecular stiffness) and the contour length (a measure of the maximum end-to-end distance). Fitting the FX relationship by WLC from zero indentation to the first peak yields the persistence length and contour length of the folded modules. For a multimeric protein composed of ~ 10 kDa modules (less than 100 aminoacid – aa), the typical values are approximately one and several nanometers, respectively. Fitting all the peaks simultaneously gives the mean persistence length of the unfolded module as well as the mean contour lengths of the unfolded modules. For a multidomain protein with ~ 10 kDa modules, the typical values are ~ 0.4 nm and ~ 20 – 30 nm, respectively.

Another approach to FX data analysis with respect to protein unfolding (or a rupture of a single bond, not discussed in this chapter) relies on obtaining the mean unfolding force, $\langle f \rangle$, as a function of the loading rate in a series of experiments [117]. Distributions of $\langle f \rangle$ versus the loading rate (df/dt) relate the FX-experiments to the FC-AFM studies and provide the (un) folding rates. While the mean rupture forces are measured precisely, the loading rates are obtained as tangents to the FX plots in the vicinity of the force peaks. However, the force–distance traces are noisy, and any extension of the molecule (described by the WLC or similar models) depends also on linkers between the tip and the substrate. Furthermore, there are only three or four decades of experimentally available loading rates, and measurements at larger loading rates need to be corrected for a viscous drag [118]. These issues currently limit the applicability of the $\langle f \rangle$ versus (df/dt) histograms.

12.5.3 Applications of FX Force Spectroscopy

There are several excellent and recent reviews on the FX methods in elucidating protein folding [119] and polysaccharides mechanics [120].

It has been showed that same multidomain proteins can have drastically different mechanical stabilities when arranged differently. It has been proven that even single mutations within a protein can significantly alter mechanical

stability of a protein. FX sensitivity to the number of different modules within multidomain proteins has been also investigated [103, 121].

FX experiments have been able to discriminate several (un)folding pathways of some larger proteins, namely, the green fluorescence protein [97, 112]. Also some intermediates in the unfolding pathways of several small (<100 aa) proteins have been resolved [102]. The refolding rates of some proteins have been addressed by initial unfolding and varying the amount of time before the protein was reextended. The zero force extrapolation of the refolding rates obtained from FX data for a small protein (ubiquitin) matched the unfolding rate obtained in bulk chemical denaturation experiments [122].

Another interesting report has revealed an unexpected ability of some unfolded proteins to generate force during refolding, and reported the direct AFM measurement of that force [93].

Ligand-receptor interactions have been studied as well. In particular, ligand-binding effects on the mechanical stability of an enzyme – the dihydrofolate reductase – have been studied by several groups [123, 124].

In another FX study, Schlierf et al. [125] introduced low-frequency lock-in detection that increased the force resolution of conventional AFM force spectroscopy and allowed them to measure protein refolding forces directly.

Low noise, slow retraction speed FX-AFM enabled studies of ligand-dependant close-to-equilibrium folded-unfolded state fluctuations of a simple protein. In particular, folding/unfolding transitions of calmodulin, a ubiquitous calcium sensor in eukaryotic cells have been studied [91]. Such studies open new possibilities in applying the AFM-based techniques into the studies of intrinsically disordered or natively unstructured proteins, which due to very shallow energy landscape operate by frequent switching between folded and unfolded states.

Furthermore, the unique fingerprints of the polysaccharides when subjected to force have been investigated, and several aspects of the force-induced structural transitions in the polysaccharides have been uncovered [87, 126].

12.6 FC-AFM Probes the Details of Protein (Un)folding and Force-Induced Disulfide Reductions in Proteins

12.6.1 Details of the FC Trace

The FX experiments lack the ability to accurately measure force-dependant parameters, and need to account on changes of the molecular stiffness with force. In the FC spectroscopy, a single biomolecule is held at a constant stretching force, and the end-to-end length of the molecule is observed as a function of time (Fig. 12.1c). An unfolding event or a substantial conformational change of the protein can change the length of the stretched molecule, and AFM feedback needs to compensate for such events, marking a step on the extension versus time trace. Changing the applied force results in a

similar feedback action, and the molecule is stretched or contracted by a precise amount of piezo displacement. Thus, once the force is applied to the biomolecule, which can unfold, an initial stretching is followed by a sequence of steps corresponding to unfolding or large conformational changes within such biomolecules.

Same sample preparation and other pertinent experimental details as mentioned for the FX spectroscopy apply to the FC setup. The calibrated molecule end-to-end distance is obtained from the externally calibrated position change of the piezo. Thus, experimental protocols should allow for some time spent at the initial tip-sample contact. Drift is an issue in maintaining the constant force in the FC trace. Piezo drift is automatically accounted for by piezo capacitive sensors. Pure sample drift would change the cantilever position, so it is accounted for by a feedback. A typical protein molecule stretched by tens of piconewtons becomes already stiffer than the cantilever, and the molecule end-to-end length becomes less sensitive to the force alterations. Thus, it is usually the cantilever's drift that limits maintaining the constant force to ~ 100 s even if the system has been equilibrated before the data acquisition.

The FC traces should be corrected for a detachment time (or a histogram of detachment times) [127]. Otherwise, only the outliers, i.e., fast unfolders, are captured. Moreover, only the traces with (1) similar length and (2) maximum possible number of events should be considered. Picking up more short traces (even with the maximum number of events present) biases the data by ending up with a larger sub-ensemble of outliers.

How long should an FC trace be? Let us assume an unfolding process of a multimodular protein with N modules, like the one presented in Fig. 12.1a. The process is stochastic, and has been approximated by treating unfolding events as uncorrelated with one another [128]. Let α be an unfolding rate. Then, the unfolding probability P within a time Δt is $P = \alpha * \Delta t$. Integrated over time P yields the probability of unfolding at a time t as $P(t)$; $P(t) = 1 - \exp(-\alpha * \Delta t)$. The value of $P(t)$ does not depend on N . However, the conditional probability of unfolding k modules out of N up to the time t is given by a binomial distribution to yield $P(N = k, t) = [N! / ((N - k)!(k!))] e^{-\alpha(N-k)t} (1 - e^{-\alpha t})^k$. If we want to get traces with a maximum number of events, then we need to calculate $P(N = N_{\max}, t)$. Typically, $N_{\max} = 8$, so $P(N = N_{\max}, t) = (1 - e^{-\alpha t})^8$. Having waited the time $t = 4/\alpha$ allows to capture $\sim 86\%$ of the data, while $t = 8/\alpha$ yields 99.7% of the data. Thus, at the beginning of each experiment, one should get a few trial FC traces, estimate the value of α , and set the clamp time accordingly. If $\alpha \sim 1 \text{ s}^{-1}$, the clamp time is ~ 10 s, but each experiment has its own considerations.

12.6.2 What can be Inferred from the FC Trace?

Typical FC data are pretty straightforward Fig. 12.1. Nevertheless, the analysis might not be so. There are several techniques available to analyze the single molecule FC data.

First, one can obtain some information about kinetics of the studied processes. Rate constants can be obtained from an average of several FC traces by fitting exponentials [129]. This type of summation and fitting has been originally used extensively in the ion channel literature [130, 131], and it has been applied to the FC-AFM data as well [101]. The averaged and normalized multiple FC curves have been also interpreted in terms of the probability of a given process at a given time, for example, (un)folding probability versus time [132].

While obtaining the kinetic rates, their corresponding errors are apparently very small due to a large number of data points in the FC curve. Thus, to estimate the standard errors of the measured rate constants one should rather use the bootstrapping method [133], or similar type of analysis. In the bootstrapping analysis, a random subset of FC traces is drawn with replacement from the set of collected FC traces. This random subset contains the same total number of traces as the original data set, but some traces may be included multiple times, while others are not included at all. Each subset is averaged and fit by exponentials. This procedure is repeated many times, and generates a distribution of rate constants. The standard deviation of each rate constant distribution is equal to the standard error of the mean for the averaged set of FC recordings [133].

The kinetic rates obtained in the ensemble average analysis can be later plotted as a function of the clamping force. In the case of a relatively simple process, like uncomplicated monomolecular or bimolecular reaction, the rates depend exponentially on the pulling force. Such dependencies are often fitted using Bell's model based on the bimolecular ligand–receptor interactions under force [134]. Bell model is a generalization of an Arrhenius equation and reads as follows:

$$k = A \exp \frac{-E_a + F\Delta x}{k_B T} = k_0 \exp \frac{F\Delta x}{k_B T} \quad , \quad (12.3)$$

where E_a is the apparent activation energy barrier (without force), F is the force, and Δx is the distance to the transition state (in a sense of reaction coordinate); k_0 comprises typical Arrhenius term (in absence of force). The proportionality constant “ A ” depends on the order of the chemical reaction being studied, and it is expressed in s^{-1} for a first order chemical reaction (like initial approximation of protein unfolding), or in $\text{M}^{-1}\text{s}^{-1}$ for a second order (or bimolecular) chemical reaction (like single molecule disulfide reduction). The physical interpretation of “ A ” is related to some kind of an attempt frequency. The Bell model of the reaction rate has worked very well in describing a single molecule disulfide reduction by simple agents [135] and other related processes.

In the case of more complicated processes, or many concurrent reaction pathways present in the system, the plot of the corresponding rates versus force can be non-Arrhenius, or can show some deviations from the Bell model. This can be the case while investigating the enzymatic catalysis [105], where several rivaling reaction pathways are recorded in the FC traces. For several simultaneous processes, the obtained rate constants are described either by

a linear combination of the actual rates or by a more complicated function [136,137]. Then, the overall kinetic model must be postulated in order to fit the rate versus force dependency, and get the actual rates. However, to elucidate the model itself, some more structural information should be known about a system, for example, X-ray or NMR data and the nature of the processes being studied.

To precisely deconvolve various kinetic processes at each force, particularly when a complicated rate versus force dependency is encountered in the FC data, one would like to collect a large number of FC traces and analyze the data using several independent methods. It can be proven [137] that so-called dwell time analysis [138,139] is not only a different way of the FC data treatment than the ensemble averaging, but it can provide a complementary information about the FC data. Dwell time refers to the time elapsed from a beginning of the FC trace to an onset of a given process, or namely, to a step in length in the FC trace. Stochastic in nature, the dwell times must be analyzed by proper probability distribution functions [138]. Pioneered by Sigworth and Sine in the patch clamp studies, the histograms of square root ordinate versus logarithm of the dwell times are convenient to display peaks at each rate constant and provide errors distributed evenly for each bin [140–142]. Thus, dwell times extracted from the FC traces [143] are sorted, binned, and displayed in the logarithmic histograms. Using a large set of FC-AFM data and applying a combination of a multiexponential fitting to averaged FC traces, as well as logarithmic dwell times histograms, Szoszkiewicz et al. [143] have experimentally resolved the two simultaneous pathways for the enzymatic disulfide bond reductions at the single molecule level. Bruijic et al. have applied the dwell time histograms to protein unfolding [144].

Another approach to analyze FC data is to use the maximum likelihood methods (MLM) to fit the dwell-time sequences. The likelihood is equal to the probability of obtaining a particular set of observed dwell times, given the form of the probability distribution, and the parameters. The MLM is free from data binning errors and its estimates are asymptotically unbiased [145]. Recently, Milesco et al. [145] have developed and tested a maximum likelihood algorithm in the context of fitting kinetics of a single-molecule motor. These algorithms can estimate the rate constants between different conformational states of proteins, and can be applied to the FC data. Bruijic et al. have used another implementation of the MLM in analyzing the protein unfolding [136].

12.6.3 Applications of the FC Spectroscopy

From the moment of its introduction [101], the FC spectroscopy has been used to infer many details about the dynamics of the protein folding and refolding, as well as – recently – in measuring kinetics of the disulfide bond reduction in proteins [105,146].

In their seminal paper, Fernandez and Li [101] have investigated engineered unfolding and refolding of the polyubiquitin protein. Ubiquitin was

first unfolded and extended at a high force, and then the stretching force was quenched. The folding trajectories were marked by several distinct stages, which departed from the two-state barrier model of ubiquitin folding obtained from the macroscopic studies. This paper has spurred lots of debate whether multidomain proteins fold cooperatively.

Finally, FC experiments might contain several force pulses, that is, several regimes with different forces applied. In particular, in force quench spectroscopy a given biomolecule (e.g., a protein) is first clamped at a given force (e.g., to unfold it), and then the tensile force is reduced so that the reverse process can be studied. Several distinct stages in protein unfolding have been studied in more detail. In particular, for the proteins that did not refold, the early refolding stage, namely the elastic collapse, has been found to agree with predictions based on a collapse of the entropic chain [147]. Also, a discussion of the hydrophobic effects in the single molecule refolding has been provided. The role of liquids on various stages of folding has been partially investigated as well [148].

Another application of the FC study was to investigate the mechanisms of the single disulfide bond reductions. The disulfide bond has been chemically engineered within the I27 protein. It has been investigated in details both using simple reducing agents, as well as enzymes, and mostly the thioredoxins – ubiquitous disulfide bond reductases [149]. The reduction rate versus force provided lots of interesting details about a possible mechanism of disulfide reduction by thioredoxin, and in general about the transition state of a single bond break-up [150].

Some insight into how mechanical stimulus is translated into a chemical response in cells has been presented recently by a combined OTs – FX/FC-AFM study [90]. In particular, mechanical stretching of single cytoplasmic protein (valin) has been showed to activate binding of other molecules (vinculin). Both valin and vinculin play significant role in cell signaling, adhesion, and migration and, *in vivo*, are localized at the sites of cell–matrix adhesions.

12.7 Some Shortcomings of the FX/FC-AFM Spectroscopies

Clearly, the FX/FC experiments have shattered the established thermodynamic models for protein (un)folding and polysaccharide stretching. However, they have encountered some criticism as well, which we try to address below.

The attachment points of the protein to the substrate and to the AFM tip are not well defined. Proteins are usually labeled at the ends by a His-tag, that is, a series of several charged aa used for a purification purposes. So, preferentially the His-tags serve as attachment points. However, this is not certain. By engineering specific end-group, the proteins can be forced to bind covalently to the substrate, and to the AFM tip. However, an elegant and simple solution is to use the recombinant chimeras between a proteins

of interest flanked by several units of small, previously characterized tandem-repeat handles (Fig 12.1).

The FX/FC data might be corrupted by the surface proximity, interactions with the surface, and knotting. Typically studied biomolecules, once adsorbed on the surface are usually pretty compact (\sim up to tens of nm) [151]. Despite short protein sizes, knotting cannot be entirely excluded due to current limited evidence. If occurred, however, knotting would lead to weird FX/FC traces, which will be rejected anyway. As for the surface interactions, any significant changes in the protein end-to-end length occur only after the molecule has been picked up from the surface and elastically stretched over a distance of tens of nanometers. Application of only several piconewton of forces stretches the multimodular protein to 30–60% of its unfolded contour length. Thus, the unfolding processes should not depend on the water structure at the surface [152], nor on the elastoadhesive properties of the contacting asperities between the AFM tip, the protein, and the substrate. In certain circumstances, however, care must be taken to minimize potentially complex long-range electrostatic interactions between the tip and the substrate [104, 153].

Typical pulling speeds available in the FX experiments are between tens to hundreds of nm/s. At these pulling speeds, most of the proteins (un)fold via trajectories far from their thermodynamic equilibrium. Thus, their (un)folding rate constants should differ from the equilibrium values. Furthermore, experimental pulling rates are four to six orders of magnitude slower than rates used for studying the molecular motion *in silico*, that is, by current computational methods like the steered molecular dynamics [146]. However, there are several novel “fluctuations theorems” providing an important bridge between the equilibrium and nonequilibrium statistical mechanics [107]. Noteworthy, opposite to mechanical unfolding, molecules undergoing thermal or chemical denaturation in bulk explore a wide range of unfolded, molten globule structures without a well-defined reaction coordinate. As a consequence, the single-molecule unfolding times are expected to be longer than in the bulk, which creates certain advantage to study these processes in singulo.

Because of limited BW of the FX/FC experiments (currently approximately from a millisecond to a hundred seconds), certain bulk processes cannot be studied. This is, however, a current limitation of most of the single molecule techniques. The FX/FC spectrometer with largely increased bandwidth is a current instrumental challenge. Such an instrument will certainly uncover interesting issues in the stretched biomolecules.

References

1. J.A. Kenniston, R.T. Sauer, Signaling degradation. *Nat. Struct. Mol. Biol.* **11**, 800–802 (2004).
2. A.W. Orr, B.P. Helmke, B.R. Blackman, M.A. Schwartz, Mechanisms of mechanotransduction. *Dev. Cell.* **10**, 11–20 (2006).

3. K. Svoboda, S.M. Block, Biological applications of optical forces. *Annu. Rev. Biophys. Biom.* **23**, 247–285 (1994).
4. C.R. Hickenboth, J.S. Moore, S.R. White, N.R. Sottos, J. Baudry, S.R. Wilson, Biasing reaction pathways with mechanical force. *Nature* **446**, 423–427 (2007).
5. C. Bustamante, Y.R. Chemla, N.R. Forde, D. Izhaky, Mechanical processes in biochemistry. *Annu. Rev. Biochem.* **73**, 705–748 (2004).
6. Berg- K. Sorensen, H. Flyvbjerg, Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.* **75**, 594–612 (2004).
7. W.J. Greenleaf, M.T. Woodside, S.M. Block, High-resolution, single-molecule measurements of biomolecular motion. *Annu. Rev. Biophys. Biom.* **36**, 171–190 (2007).
8. C. Bustamante, J.C. Macosko, G.J.L. Wuite, Grabbing the cat by the tail: Manipulating molecules one by one. *Nat. Rev. Mol. Cell. Bio.* **1**, 130–136 (2000).
9. J. Gelles, B.J. Schnapp, M.P. Sheetz, Tracking kinesin-driven movements with nanometre-scale precision. *Nature* **331**, 450–453 (1988).
10. M.P. Sheetz, S. Turney, H. Qian, E.L. Elson, Nanometer-level analysis demonstrates that lipid flow does not drive membrane glycoprotein movements. *Nature* **340**, 284–288 (1989).
11. A. Yildiz, P.R. Selvin, Fluorescence imaging with one nanometre accuracy: Application to molecular motors. *Accounts. Chem. Res.* **38**, 574–582 (2005).
12. A.D. Rowe, M.C. Leake, H. Morgan, R.M. Berry, Rapid rotation of micron and submicron dielectric particles measured using optical tweezers. *J. Mod. Optic.* **50**, 1539–1554 (2003).
13. Y. Sowa, A.D. Rowe, M.C. Leake, T. Yakushi, M. Homma, A. Ishijima, R.M. Berry, Direct observation of steps in rotation of the bacterial flagellar motor. *Nature* **437**, 916–919 (2005).
14. H. Noji, R. Yasuda, M. Yoshida, K. Kinoshita, Direct observation of the rotation of F-1-ATPase. *Nature* **386**, 299–302 (1997).
15. R. Yasuda, H. Noji, K. Kinoshita, M. Yoshida, F-1-ATPase is a highly efficient molecular motor that rotates with discrete 120 degree steps. *Cell* **93**, 1117–124 (1998).
16. A. Yildiz, J.N. Forkey, S.A. McKinney, T. Ha, Y.E. Goldman, P.R. Selvin, Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061–2065 (2003).
17. A. Yildiz, M. Tomishige, R.D. Vale, P.R. Selvin, Kinesin walks hand-over-hand. *Science* **303**, 676–678 (2004).
18. R.S. Rock, S.E. Rice, A.L. Wells, T.J. Purcell, J.A. Spudich, H.L. Sweeney, Myosin VI is a processive motor with a large step size. *Proc. Natl. Acad. Sci. USA* **98**, 13655–13659 (2001).
19. K. Svoboda, C.F. Schmidt, B.J. Schnapp, S.M. Block, Direct observation of kinesin stepping by optical trapping interferometry. *Nature* **365**, 721–727 (1993).
20. D. Altman, H.L. Sweeney, J.A. Spudich, The mechanism of myosin VI translocation and its load-induced anchoring. *Cell* **116**, 737–749 (2004).
21. J.T. Finer, R.M. Simmons, J.A. Spudich, Single myosin molecule mechanics - piconewton forces and nanometer steps. *Nature* **368**, 113–119 (1994).
22. R. Mallik, B.C. Carter, S.A. Lex, S.J. King, S.P. Gross, Cytoplasmic dynein functions as a gear in response to load. *Nature* **427**, 649–652 (2004).

23. K. Svoboda, S.M. Block, Force and velocity measured for single kinesin molecules. *Cell* **77**, 773–784 (1994).
24. L. Stryer, Fluorescence energy-transfer as a spectroscopic ruler. *Annu. Rev. Biochem.* **47**, 819–846 (1978).
25. T. Ha, T. Enderle, D.F. Ogletree, D.S. Chemla, P.R. Selvin, S. Weiss, Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. USA* **93**, 6264–6268 (1996).
26. T. Forster, Energiewanderung Und Fluoreszenz. *Naturwissenschaften* **33**, 166–175 (1946).
27. T. Forster, *Zwischenmolekulare Energiewanderung Und Fluoreszenz. *Ann. Phys-Berlin.* **2**, 55–75 (1948).
28. S. Brasselet, E.J.G. Peterman, A. Miyawaki, W.E. Moerner, Single-molecule fluorescence resonant energy transfer in calcium concentration dependent cameleon. *J. Phys. Chem. B* **104**, 3676–3682 (2000).
29. B. Schuler, E.A. Lipman, W.A. Eaton, Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* **419**, 743–747 (2002).
30. S.A. McKinney, A.D.J. Freeman, D.M.J. Lilley, T.J. Ha, Observing spontaneous branch migration of Holliday junctions one step at a time. *Proc. Natl. Acad. Sci. USA* **102**, 5715–5720 (2005).
31. J.Y. Lee, B. Okumus, D.S. Kim, T.J. Ha, Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci. USA* **102**, 18938–18943 (2005).
32. A.N. Kapanidis, E. Margeat, S.O. Ho, E. Kortkhonjia, S. Weiss, R.H. Ebright, Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**, 1144–1147 (2006).
33. E. Tan, T.J. Wilson, M.K. Nahas, R.M. Clegg, D.M.J. Lilley, T. Ha, A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc. Natl. Acad. Sci. USA* **100**, 9308–9313 (2003).
34. X.W. Zhuang, Single-molecule RNA science. *Annu. Rev. Biophys. Biom.* **34**, 399–414 (2005).
35. X.W. Zhuang, H. Kim, M.J.B. Pereira, H.P. Babcock, N.G. Walter, S. Chu, Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**, 1473–1476 (2002).
36. C. Gosse, V. Croquette, Magnetic tweezers: Micromanipulation and force measurement at the molecular level. *Biophys. J.* **82**, 3314–3329 (2002).
37. K.C. Neuman, E.H. Chadd, G.F. Liou, K. Bergman, S.M. Block, Characterization of photodamage to *Escherichia coli* in optical traps. *Biophys. J.* **77**, 2856–2863 (1999).
38. J. Gore, Z. Bryant, M. Nollmann, M.U. Le, N.R. Cozzarelli, C. Bustamante, DNA overwinds when stretched. *Nature* **442**, 836–839 (2006).
39. S.B. Smith, L. Finzi, C. Bustamante, Direct mechanical measurements of the elasticity of single dna-molecules by using magnetic beads. *Science* **258**, 1122–1126 (1992).
40. T.R. Strick, J.F. Allemand, D. Bensimon, A. Bensimon, V. Croquette, The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835–1837 (1996).
41. T.R. Strick, J.F. Allemand, D. Bensimon, V. Croquette, Stress-induced structural transitions in DNA and proteins. *Annu. Rev. Biophys. Biom.* **29**, 523–543 (2000).

42. T.R. Strick, V. Croquette, D. Bensimon, Single-molecule analysis of DNA uncoiling by a type II topoisomerase. *Nature* **404**, 901–904 (2000).
43. S. Bigot, O.A. Saleh, F. Cornet, J.F. Allemand, F.X. Barre, Oriented loading of FtsK on KOPS. *Nat. Struct. Mol. Biol.* **13**, 1026–1028 (2006).
44. O.A. Saleh, C. Perals, F.X. Barre, J.F. Allemand, Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment. *Embo. J.* **23**, 2430–2439 (2004).
45. D.A. Koster, V. Croquette, C. Dekker, S. Shuman, N.H. Dekker, Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. *Nature* **434**, 671–674 (2005).
46. G. Charvin, J.F. Allemand, T.R. Strick, D. Bensimon, V. Croquette, Twisting DNA: single molecule studies. *Contemp. Phys.* **45**, 383–403 (2004).
47. A. Ashkin, J.M. Dziedzic, J.E. Bjorkholm, S. Chu, Observation of a single-beam gradient force optical trap for dielectric particles. *Opt. Lett.* **11**, 288–290 (1986).
48. W.J. Greenleaf, K.L. Frieda, D.A.N. Foster, M.T. Woodside, S.M. Block, Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**, 630–633 (2008).
49. J.D. Wen, L. Lancaster, C. Hodges, A.C. Zeri, S.H. Yoshimura, H.F. Noller, C. Bustamante, I. Tinoco, Following translation by single ribosomes one codon at a time. *Nature* **452**, 598–U2 (2008).
50. C. Bustamante, In singulo biochemistry: When less is more. *Annu. Rev. Biochem.* **77**, 45–50 (2008).
51. J.R. Moffitt, Y.R. Chemla, S.B. Smith, C. Bustamante, Recent advances in optical tweezers. *Annu. Rev. Biochem.* **77**, 205–228 (2008).
52. D.G. Grier, A revolution in optical manipulation. *Nature* **424**, 810–816 (2003).
53. L. Paterson, M.P. MacDonald, J. Arlt, W. Sibbett, P.E. Bryant, K. Dholakia, Controlled rotation of optically trapped microscopic particles. *Science* **292**, 912–914 (2001).
54. V. Bingelyte, J. Leach, J. Courtial, M.J. Padgett, Optically controlled three-dimensional rotation of microscopic objects. *Appl. Phys. Lett.* **82**, 829–831 (2003).
55. F.S. Pavone, G. Romano, M. Capitanio, L. Sacconi, A novel 3D magneto-optical manipulator for single bio-molecule applications. *Biophys. J.* **82**, 41a-a (2002).
56. M.E.J. Friese, T.A. Nieminen, N.R. Heckenberg, H. Rubinsztein-Dunlop, Optical alignment and spinning of laser-trapped microscopic particles. *Nature* **394**, 348–350 (1998).
57. A.N. Fehr, C.L. Asbury, S.M. Block, Kinesin steps do not alternate in size. *Biophys. J.* **94**, L20–L2 (2008).
58. K. Kaseda, H. Higuchi, K. Hirose, Alternate fast and slow stepping of a heterodimeric kinesin molecule. *Nat. Cell. Biol.* **5**, 1079–1082 (2003).
59. M.J. Schnitzer, S.M. Block, Kinesin hydrolyses one ATP per 8-nm step. *Nature* **388**, 386–390 (1997).
60. J.E. Molloy, J.E. Burns, J. Kendrickjones, R.T. Tregear, D.C.S. White, Movement and force produced by a single myosin head. *Nature* **378**, 209–212 (1995).
61. K.C. Neuman, E.A. Abbondanzieri, R. Landick, J. Gelles, S.M. Block, Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* **115**, 437–447 (2003).

62. S. Toba, T.M. Watanabe, L. Yamaguchi-Okimoto, Y.Y. Toyoshima, H. Higuchi, Overlapping hand-over-hand mechanism of single molecular motility of cytoplasmic dynein. *Proc. Natl. Acad. Sci. USA* **103**, 5741–5745 (2006).
63. S. Dumont, W. Cheng, V. Serebrov, R.K. Beran, I. Tinoco, A.M. Pyle, C. Bustamante, RNA translocation and unwinding mechanism of HCVNS3 helicase and its coordination by ATP. *Nature* **439**, 105–108 (2006).
64. T.T. Perkins, R.V. Dalal, P.G. Mitsis, S.M. Block, Sequence-dependent pausing of single lambda exonuclease molecules. *Science* **301**, 1914–1918 (2003).
65. J.L. Ptacin, M. Nollmann, C. Bustamante, N.R. Cozzarelli, Identification of the FtsK sequence-recognition domain. *Nat. Struct. Mol. Biol.* **13**, 1023–1025 (2006).
66. G.J.L. Wuite, S.B. Smith, M. Young, D. Keller, C. Bustamante, Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* **404**, 103–106 (2000).
67. E.A. Abbondanzieri, W.J. Greenleaf, J.W. Shaevitz, R. Landick, S.M. Block, Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460–465 (2005).
68. M.H. Larson, W.J. Greenleaf, R. Landick, S.M. Block, Applied force reveals mechanistic and energetic details of transcription termination. *Cell* **132**, 971–982 (2008).
69. K. Adelman, La A. Porta, T.J. Santangelo, J.T. Lis, J.W. Roberts, M.D. Wang, Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proc. Natl. Acad. Sci. USA* **99**, 13538–13543 (2002).
70. R.V. Dalal, M.H. Larson, K.C. Neuman, J. Gelles, R. Landick, S.M. Block, Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. *Mol. Cell.* **23**, 231–239 (2006).
71. D.E. Smith, S.J. Tans, S.B. Smith, S. Grimes, D.L. Anderson, C. Bustamante, The bacteriophage phi 29 portal motor can package DNA against a large internal force. *Nature* **413**, 748–752 (2001).
72. Y.R. Chemla, K. Aathavan, J. Michaelis, S. Grimes, P.J. Jardine, D.L. Anderson, C. Bustamante, Mechanism of force generation of a viral DNA packaging motor. *Cell* **122**, 683–692 (2005).
73. J.R. Moffitt, Y.R. Chemla, K. Aathavan, S. Grimes, P.J. Jardine, D.L. Anderson, C. Bustamante, Intersubunit coordination in a homomeric ring ATPase. *Nature* **457**, 446–U2 (2009).
74. Y. Seol, G.M. Skinner, K. Visscher, Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* **93**, (2004).
75. Y. Seol, G.M. Skinner, K. Visscher, A. Buhot, A. Halperin, Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Lett.* **98**, (2007).
76. S.B. Smith, Y.J. Cui, C. Bustamante, Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271**, 795–799 (1996).
77. S.J. Koch, A. Shundrovsky, B.C. Jantzen, M.D. Wang, Probing protein-DNA interactions by unzipping a single DNA double helix. *Biophys. J.* **83**, 1098–1105 (2002).
78. J. Liphardt, B. Onoa, S.B. Smith, I. Tinoco, C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**, 733–737 (2001).

79. B. Onoa, S. Dumont, J. Liphardt, S.B. Smith, I. Tinoco, C. Bustamante, Identifying kinetic barriers to mechanical unfolding of the T-thermophila ribozyme. *Science* **299**, 1892–1895 (2003).
80. M.T. Woodside, P.C. Anthony, W.M. Behnke-Parks, K. Larizadeh, D. Herschlag, S.M. Block, Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**, 1001–1004 (2006).
81. C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Direct observation of the three-state folding of a single protein molecule. *Science* **309**, 2057–2060 (2005).
82. M.S.Z. Kellermayer, S.B. Smith, H.L. Granzier, C. Bustamante, Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science* **276**, 1112–1116 (1997).
83. N.A. Burnham, R.J. Colton, Measuring the nanomechanical properties and surface forces of materials using an atomic force microscope. *J. Vac. Sci. Technol. A* **7**, 2906–2913 (1989).
84. G.U. Lee, L.A. Chrisey, R.J. Colton, Direct measurement of the forces between complementary strands of DNA. *Science* **266**, 771–773 (1994).
85. E.L. Florin, V.T. Moy, H.E. Gaub, Adhesion forces between individual ligand-receptor pairs. *Science* **264**, 415–417 (1994).
86. M. Rief, F. Oesterhelt, B. Heymann, H.E. Gaub, Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science* **275**, 1295–1297 (1997).
87. P.E. Marszalek, H.B. Li, J.M. Fernandez, Fingerprinting polysaccharides with single-molecule atomic force microscopy. *Nat. Biotechnol.* **19**, 258–262 (2001).
88. M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109–1112 (1997).
89. Carrion- M. Vazquez, A.F. Oberhauser, T.E. Fisher, P.E. Marszalek, H.B. Li, J.M. Fernandez, Mechanical design of proteins-studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys. Mol. Bio.* **74**, 63–91 (2000).
90. A. del Rio, R. Perez-Jimenez, R.C. Liu, P. Roca-Cusachs, J.M. Fernandez, M.P. Sheetz, Stretching single talin rod molecules activates vinculin binding. *Science* **323**, 638–641 (2009).
91. J.P. Junker, F. Ziegler, M. Rief, Ligand-Dependent Equilibrium Fluctuations of Single Calmodulin Molecules. *Science* **323**, 633–637 (2009).
92. M. Schlierf, M. Rief, Surprising simplicity in the single-molecule folding mechanics of proteins. *Angew. Chem. Int. Edit.* **48**, 820–822 (2009).
93. G. Lee, K. Abdi, Y. Jiang, P. Michaely, V. Bennett, P.E. Marszalek, Nanospring behaviour of ankyrin repeats. *Nature* **440**, 246–249 (2006).
94. M. Rief, J. Pascual, M. Saraste, H.E. Gaub, Single molecule force spectroscopy of spectrin repeats: Low unfolding forces in helix bundles. *J. Mol. Biol.* **286**, 553–561 (1999).
95. R.B. Best, B. Li, A. Steward, V. Daggett, J. Clarke, Can non-mechanical proteins withstand force? Stretching barnase by atomic force microscopy and molecular dynamics simulation. *Biophys. J.* **81**, 2344–2356 (2001).
96. H. Dietz, F. Berkemeier, M. Bertz, M. Rief, Anisotropic deformation response of single protein molecules. *Proc. Natl. Acad. Sci. USA* **103**, 12724–12728 (2006).

97. H. Dietz, M. Rief, Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc. Natl. Acad. Sci. USA* **101**, 16192–16197 (2004).
98. P.M. Williams, S.B. Fowler, R.B. Best, J.L. Toca-Herrera, K.A. Scott, A. Steward, J. Clarke, Hidden complexity in the mechanical properties of titin. *Nature* **422**, 446–449 (2003).
99. Carrion- M. Vazquez, A.F. Oberhauser, S.B. Fowler, P.E. Marszalek, S.E. Broedel, J. Clarke, J.M. Fernandez, Mechanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Acad. Sci. USA* **96**, 3694–3699 (1999).
100. H.B. Li, W.A. Linke, A.F. Oberhauser, M. Carrion-Vazquez, J.G. Kerkvliet, H. Lu, P.E. Marszalek, J.M. Fernandez, Reverse engineering of the giant muscle protein titin. *Nature* **418**, 998–1002 (2002).
101. J.M. Fernandez, H.B. Li, Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science* **303**, 1674–1678 (2004).
102. P.E. Marszalek, H. Lu, H.B. Li, M. Carrion-Vazquez, A.F. Oberhauser, K. Schulten, J.M. Fernandez, Mechanical unfolding intermediates in titin modules. *Nature* **402**, 100–103 (1999).
103. Carrion-M. Vazquez, H.B. Li, H. Lu, P.E. Marszalek, A.F. Oberhauser, J.M. Fernandez, The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Biol.* **10**, 738–743 (2003).
104. J. Zlatanova, S.M. Lindsay, S.H. Leuba, Single molecule force spectroscopy in biology using the atomic force microscope. *Prog. Biophys. Mol. Biol.* **74**, 37–61 (2000).
105. A.P. Wiita, Perez- R. Jimenez, K.A. Walther, F. Grater, B.J. Berne, A. Holmgren, J.M. Sanchez-Ruiz, J.M. Fernandez, Probing the chemistry of thioredoxin catalysis with force. *Nature* **450**, 124 (2007).
106. C. Bustamante, S.B. Smith, J. Liphardt, D. Smith, Single-molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.* **10**, 279–285 (2000).
107. J. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C. Bustamante, Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science* **296**, 1832–1835 (2002).
108. N.A. Burnham, G. Gremaud, A.J. Kulik, P.J. Gallo, F. Oulevey, Materials' properties measurements: Choosing the optimal scanning probe microscope configuration. *J. Vac. Sci. Technol. B* **14**, 1308–1312 (1996).
109. R.M. Simmons, J.T. Finer, S. Chu, J.A. Spudich, Quantitative measurements of force and displacement using an optical trap. *Biophys. J.* **70**, 1813–1822 (1996).
110. W.J. Greenleaf, M.T. Woodside, E.A. Abbondanzieri, S.M. Block, Passive all-optical force clamp for high-resolution laser trapping. *Phys. Rev. Lett.* **95**, (2005).
111. A.F. Oberhauser, P.E. Marszalek, H.P. Erickson, J.M. Fernandez, The molecular elasticity of the extracellular matrix protein tenascin. *Nature* **393**, 181–185 (1998).
112. R. Perez-Jimenez, S. Garcia-Manyes, S.R.K. Ainavarapu, J.M. Fernandez, Mechanical unfolding pathways of the enhanced yellow fluorescent protein revealed by single molecule force spectroscopy. *J. Biol. Chem.* **281**, 40010–40014 (2006).
113. A. Sarkar, R.B. Robertson, J.M. Fernandez, Simultaneous atomic force microscope and fluorescence measurements of protein unfolding using a calibrated evanescent wave. *Proc. Natl. Acad. Sci. USA* **101**, 12882–12886 (2004).

114. E.L. Florin, M. Rief, H. Lehmann, M. Ludwig, C. Dornmair, V.T. Moy, H.E. Gaub, Sensing specific molecular-interactions with the atomic-force microscope. *Biosens. Bioelectron.* **10**, 895–901 (1995).
115. H.J. Butt, M. Jaschke, Calculation of thermal noise in atomic-force microscopy. *Nanotechnology* **6**, 1–7 (1995).
116. N.A. Burnham, X. Chen, C.S. Hodges, G.A. Matei, E.J. Thoreson, C.J. Roberts, M.C. Davies, S.J.B. Tendler, Comparison of calibration methods for atomic-force microscopy cantilevers. *Nanotechnology* **14**, 1–6 (2003).
117. O.K. Dudko, G. Hummer, A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* **96**, (2006).
118. H. Janovjak, J. Struckmeier, D.J. Muller, Hydrodynamic effects in fast AFM single-molecule force measurements. *Eur. Biophys. J. Biophys.* **34**, 91–96 (2005).
119. A. Borgia, P.M. Williams, J. Clarke, Single-molecule studies of protein folding. *Annu. Rev. Biochem.* **77**, 101–125 (2008).
120. M. Rabbi, P.E. Marszalek, Single-molecule techniques: a laboratory manual, ed. by P.R. Selvin, T. Ha (CSH, 2008).
121. H.B. Li, Carrion- M. Vazquez, A.F. Oberhauser, P.E. Marszalek, J.M. Fernandez, Point mutations alter the mechanical stability of immunoglobulin modules. *Nat. Struct. Biol.* **7**, 1117–1120 (2000).
122. H.B. Li, A.F. Oberhauser, S.B. Fowler, J. Clarke, J.M. Fernandez, Atomic force microscopy reveals the mechanical design of a modular protein. *Proc. Natl. Acad. Sci. USA* **97**, 6527–6531 (2000).
123. R.K. Aivarapu, L.W. Li, J.M. Fernandez, Fingerprinting DHFR in single-molecule AFM studies. *Biophys. J.* **91**, 2009–2010 (2006).
124. M. Rief, J.P. Junker, M. Schlierf, K. Hell, W. Neupert, Fingerprinting DHFR in single-molecule AFM studies – Response to the comment by Aivarapu et al. *Biophys. J.* **91**, 2011–2012 (2006).
125. M. Schlierf, F. Berkemeier, M. Rief, Direct observation of active protein folding using lock-in force spectroscopy. *Biophys. J.* **93**, 3989–3998 (2007).
126. P.E. Marszalek, A.F. Oberhauser, Y.P. Pang, J.M. Fernandez, Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring. *Nature* **396**, 661–664 (1998).
127. S. Garcia-Manyes, J. Brujic, C.L. Badilla, J.M. Fernandez, Force-clamp spectroscopy of single-protein monomers reveals the individual unfolding and folding pathways of I27 and ubiquitin. *Biophys. J.* **93**, 2436–2446 (2007).
128. J. Brujic, J.W. Fernandez, Response to comment on “Force-clamp spectroscopy monitors the folding trajectory of a single protein”. *Science* **308**, (2005).
129. A.P. Wiita, S.R.K. Aivarapu, H.H. Huang, J.M. Fernandez, Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. *Proc. Natl. Acad. Sci. USA* **103**, 7222–7227 (2006).
130. F.J. Sigworth, E. Neher, Single Na⁺ channel currents observed in cultured rat muscle cells. *Nature* **287**, 447–449 (1980).
131. R.W. Aldrich, D.P. Corey, C.F. Stevens, A reinterpretation of mammalian sodium channel gating based on single channel recording. *Nature* **306**, 436–441 (1983).
132. M. Schlierf, H.B. Li, J.M. Fernandez, The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7299–7304 (2004).

133. B. Efron, *The Jackknife, the Bootstrap, and Other Resampling Plans* (S.I.A.M., Philadelphia, PA, 1982).
134. G.I. Bell, Models for the specific adhesion of cells to cells. *Science* **200**, 618–627 (1978).
135. S.R.K. Ainarapu, A.P. Wiita, L. Dougan, E. Uggerud, J.M. Fernandez, Single-molecule force spectroscopy measurements of bond elongation during a bimolecular reaction. *J. Am. Chem. Soc.* **130**, 6479–6487 (2008).
136. J. Brujic, R.I. Hermans, K.A. Walther, J.M. Fernandez, Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin. *Nature Phys.* **2**, 282–286 (2006).
137. X.C. Xue, L.C. Gong, F. Liu, Z.C. Ou-Yang, Two-pathway four-state kinetic model of thioredoxin-catalyzed reduction of single forced disulfide bonds. *Phys. Rev. E* **77**, (2008).
138. D.R. Cox, H.D. Miller, *The theory of stochastic processes* (Methuen, London, 1965).
139. O.P. Hamill, A. Marty, E. Neher, B. Sakmann, F.J. Sigworth, Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflugers Archiv-European Journal of Physiology* **391**, 85–100 (1981).
140. F.J. Sigworth, S.M. Sine, Data transformations for improved display and fitting of single-channel dwell time histograms. *Biophys. J.* **52**, 1047–1054 (1987).
141. A.A. Carter, R.E. Oswald, Linear prediction and single-channel recording. *J. Neurosci. Methods.* **60**, 69–78 (1995).
142. R.E. Wachtel, Use of Bmdp statistical package to generate maximum-likelihood estimates for single channel data. *J. Neurosci. Methods* **25**, 121–128 (1988).
143. R. Szoszkiewicz, S.R.K. Ainarapu, A.P. Wiita, R. Perez-Jimenez, J.M. Sanchez-Ruiz, J.M. Fernandez, Dwell time analysis of a single-molecule mechanochemical reaction. *Langmuir* **24**, 1356–1364 (2008).
144. J. Brujic, R.I.Z. Hermans, Garcia- S. Manyes, K.A. Walther, J.M. Fernandez, Dwell-time distribution analysis of polyprotein unfolding using force-clamp spectroscopy. *Biophys. J.* **92**, 2896–2903 (2007).
145. L.S. Milescu, A. Yildiz, P.R. Selvin, F. Sachs, Maximum likelihood estimation of molecular motor kinetics from staircase dwell-time sequences. *Biophys. J.* **91**, 1156–1168 (2006).
146. L. Dougan, J. Brujic, J.M. Fernandez, *Single molecule dynamics in life science*, ed. by T. Yanagida, Y. Ishii (Wiley, GmbH, 2009) pp 297–322.
147. K.A. Walther, F. Grater, L. Dougan, C.L. Badilla, B.J. Berne, J.M. Fernandez, Signatures of hydrophobic collapse in extended proteins captured with force spectroscopy. *Proc. Natl. Acad. Sci. USA* **104**, 7916–7921 (2007).
148. L. Dougan, G. Feng, H. Lu, J.M. Fernandez, Solvent molecules bridge the mechanical unfolding transition state of a protein. *Proc. Natl. Acad. Sci. USA* **105**, 3185–3190 (2008).
149. A. Holmgren, Thioredoxin. *Annu. Rev. Biochem.* **54**, 237–271 (1985).
150. M. Grandbois, M. Beyer, M. Rief, Clausen- H. Schaumann, H.E. Gaub, How strong is a covalent bond? *Science* **283**, 1727–1730 (1999).
151. E. Casero, L. Vazquez, J. Martin-Benito, M.A. Morcillo, E. Lorenzo, M. Pariente, Immobilization of metallothionein on gold/mica surfaces: Relationship between surface morphology and protein-substrate interaction. *Langmuir* **18**, 5909–5920 (2002).

152. T.D. Li, J.P. Gao, R. Szoszkiewicz, U. Landman, E. Riedo, Structured and viscous water in subnanometer gaps. *Phys. Rev. B* **75**, 115415 (2007).
153. R. Szoszkiewicz, A.J. Kulik, G. Gremaud, M. Lekka, Probing local water contents of in vitro protein films by ultrasonic force microscopy. *Appl. Phys. Lett.* **86**, (2005).

Multilevel Experimental and Modelling Techniques for Bioartificial Scaffolds and Matrices

F. Consolo, F. Mastrangelo, G. Ciardelli, F.M. Montecvecchi, U. Morbiducci, M. Sassi, and C. Bignardi

Summary. Tissue engineering (TE) is the application of principles and methods of engineering and life sciences towards the fundamental understanding of structure–function relationships in normal and pathological mammalian tissues and the development of biological substitutes to restore, maintain or improve tissue function.

One key component to TE is using three-dimensional porous scaffolds to guide cells during the regeneration process. These scaffolds are intended to provide cells with an environment that promotes cell attachment, proliferation, and differentiation. After sufficient tissue regeneration using *in vitro* culturing methods, the scaffold/tissue structure is implanted into the patient, where the scaffold will degrade away, thereby leaving only regenerated tissue; on a different approach, non-cellularised scaffolds are inserted into the patient to elicit *in vivo* cell recruitment, growth and tissue regeneration. Tissue-engineered scaffolds need to meet both the biological goals of tissue formation and the stresses and loading conditions present in the human body. For this reason, any design approach must ensure that the mechanical properties of the resulting scaffold structure are compatible and optimally match the requirements from the environment, that, respectively, are the cell adhesion transmembrane protein, the cytoskeleton structure, the cell population. The need to design scaffold structures, the need for precision control during their fabrication and for determining the metrological indices and the need to characterise their structural behaviour at different scales have led to numerous experimental and computational challenges. In particular, there is a need for modelling and test tissue at multiple scales to gain insight into issues such as drug delivery, drug interaction, gene expression and cellular–environment interactions. The analysis of the tissue constructs at different scales includes a macro-scale model where the macro-scale tissue construct is characterised, a multi-cellular model where a sufficiently large multi-cellular representative element volume is selected to represent a microstructure of the tissue construct and a single cell model wherein the microstructures of the cell like the nucleus and the cytoplasm have been incorporated. A multi-scale approach is already being applied to bridge nano- and micro-scales as well as micro- and macro-scales within various research areas in TE.

In this chapter, a review of the experimental and modelling techniques used for the evaluation, at different scales, of the mechanical and morphological properties of bioartificial scaffolds and matrices, such as compression testing, nanoindentation,

AFM technique, Dynamical Mechanical Analysis (DMA), micro-CT, microMR, Asymptotic Homogenisation Theory, Finite Element Analysis (FEA), Rule-of-Mixtures, is proposed.

Key words: Tissue Engineering, Scaffold, Computational Aided Tissue Engineering, Transport Phenomena, CFD, micro-CT.

13.1 Scaffolds for Tissue-Engineering Applications

Over time, experimental focus of bioengineers and researchers involved in the investigation of strategies for the regeneration of pathological injured organs has shifted progressively away from the efforts of one stage organ development towards the improvement of single tissue components with the capacity of in vitro growth and in vivo integration. Along with this process, components of the extracellular matrix (ECM) of the tissues have gained increasing attention as crucial elements in directing the development and maintaining the proper characteristics of three-dimensional (3D) cellular struts [1].

In fabricating engineered tissue-like constructs, the design and manufacture of a polymeric (natural or synthetic) structure (scaffold) to provide cells with an environment serving as a substrate for the growth of the cells, that will give rise to the formation of the new tissue, is a primary issue [2]. Following this approach leads, in substance, to the goal of bioartificially mimicking, substituting and replacing, by means of scaffolds, the native ECM architecture and functions. For this reason, in designing scaffolds, the first step should be the effort in faithfully reproducing the complex ECM nano-architecture, which is no longer regarded as a mere carrier material or tissue conductive mechanical element, but, as mentioned before, as an important tissue component with bioinductive functions able to promote a positive cell response to the in vitro culture protocol, piloting mechanical, physics and chemical signals to the cells.

Scaffolds are open systems formed by structural elements such as fibres or membranes, characterised by a network of interconnected pores with dimensions in the range of the micrometers, in analogy with the ECM morphology (Fig. 13.1) [3]. The ideal characteristics required for such scaffolds are the following:

- (1) Suitable surface chemical features allowing cell adhesion, proliferation and differentiation [4];
- (2) A 3D porous structure permitting and guiding cell growth and allowing an adequate flow of nutrients and removal of wastes as well the formation of a new-vascularization network (angiogenesis);
- (3) Biocompatibility minimising the host's immune response of the seeded cells and of the donor native tissue and guaranteeing the absence of toxic effects [5];

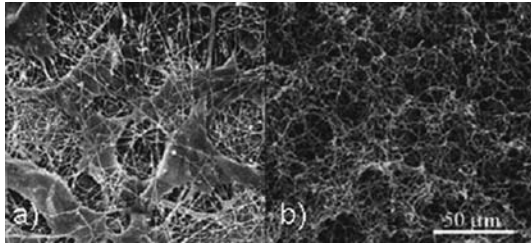


Fig. 13.1. SEM micrograph of (a) cells embedded in ECM fibrous matrix and (b) biomimetic scaffolds attempting to emulate the native ECM morphology [3]

- (4) Bioresorbability, that is, the close co-ordination of degradation and new tissue formation rate [6];
- (5) Appropriate mechanical properties allowing for structural support and mechano-transduction of the forces generated by the cells during tissue formation and functionalisation.

Currently, there have been only a few successes within TE in human clinical applications, but all of them have been achieved using scaffold-mediated techniques. Success have been reached in growing human skin [7], non-load bearing cartilage [8], and human urinary bladders [9]. These milestones reinforce the possibility of growing organs or tissues such as bone, liver, heart valves, myocardial tissue and blood vessels, for clinical applications, within the next few decades. In order to reach the previous mentioned requirements for the scaffold, a comprehensive knowledge of the morphology, topology and functions of the ECM is needed, as well as an in-depth analysis and comprehension of the cellular mechanics and, at last, of the mechanisms involved and responsible for ‘active’ cell–substrate interactions. Thanks to the continuous technological progress, it has been possible to point out, for instance, how the permeability of biological tissues, and consequently of tissue-engineered scaffolds, plays a primary role in the nutrient and waste transport within the cellular shells, by means of diffusion or transport through capillary networks formed throughout the scaffold via angiogenesis [10]. The influence of fluid flow has been studied in a variety of tissues [11–14]. Furthermore, scaffold and tissue permeability also influence cyclic changes in biophysical stimuli due to the fluid flow through the structure during mechanical loading. Specifically, permeability affects the magnitude of pressure and fluid shear stresses within the construct or tissue and they have been identified as potential stimuli for cellular differentiation or functional adaptation [15–17]. Construct permeability has also been shown to influence the degradation rate of biodegradable scaffolds for TE [18]. The permeability of a scaffold is determined by a combination of micro- and nano-structural factors including porosity, pore size, geometry and distribution, pore interconnectivity, fenestration size, and the orientation of pores with respect to flow direction [19]. Additionally, construct chemical composition can influence permeability. Different biological compounds

such as glycosaminoglycans and collagen [20] have been shown to affect the permeability of many different tissues. Significant considerations can also be drawn referring to mechano-transduction mechanisms, which are primarily mediated by the mutual interaction between the substrate/ECM and the cells [21]. Mechano-transduction can be roughly divided into two categories: (1) direct, through which conformational changes in the cellular cytoskeleton transmit information on the extracellular environment (shape, flow, strain, etc.) to the nucleus as mechanical signals and (2) indirect, when information to the nucleus of the cell is relied through signalling cascades involving chemical messaging. As known, mechano-transduction plays key roles in blood pressure regulation, vascular response to fluid shear stress, remodelling of bone, maintenance of muscle etc. [22]. It is likely that cell growth, migration and gene expression are influenced by mechano-transductive events in most, if not all, cell types [23].

All the previous considerations imply a further essential remark, which is the following: all the concerned issues call for the analysis and investigation of multiple and different scales phenomena. A multiple scales approach is necessary, moving from the macro- to the nano-scale at increasing levels of detail, in order to establish any scaffold design approach, to plan activities and strategies to develop systems for scaffolds production, as well as for modelling and testing bioartificial engineered or native tissues functioning. Referring to the analysis of the tissue constructs at different scales, it includes a macro-scale model where the macroscopic tissue construct is characterised, a micro-scale model where a sufficiently large multi-cellular representative volume is selected to represent a micro-structure of the tissue construct and a nano-scale model wherein the nano-structures of a single cell, like the cytoskeleton, the nucleus or the cytoplasm, has been incorporated [7]. As an example, let us consider the case of the myocardial tissue in which the myocardial ECM (micro-scale) allows to collect the physical forces that are developed by the single cells (nano-scale) at each excitation–contraction cycle, in order to generate the macroscopic contraction force of the heart (macro-scale) [24, 25]. Furthermore, moving through a more detailed analysis, in order to determine, for instance, the cardiomyocytes mechanics, a multi-scale approach could be again adopted: the cell excitation–contraction represents the macro-scale model, the sub-micron level deformation, such as localised distortion of the cell membrane or conformational modifications of the cytoskeleton, corresponds to the micro-scale model and the cascade of biochemical reactions for mechano-transduction signalling and regulatory pathways, wired together in a complex network of intermolecular interactions, is the nano-scale model.

Regarding the scaffolds, the need for precision control during fabrication, together with the need for the evaluation of the effect of each specific scaffold parameter affecting tissue regeneration, has lead to the development of several multi-scale–based methods for scaffold modelling. Multi-scale modelling consists of a systematic approach to obtain data from the natural tissue and to incorporate that data into a scaffold. In the next sections, a more detailed

analysis of the multi-scale approach in modelling scaffolds and biological tissue mechanics will be provided.

13.2 Multi-Scale Computer-Aided Approach in Designing and Modelling Scaffold for Tissue Regeneration

Matching the proper scaffold requirements implies a design methodology through which the single scaffold parameters affecting the tissue regeneration could be selectively analysed, aiming at controlling such parameters during fabrication stage, allowing for mimicking the native ECM structure and functions. This issue is nowadays established on multi-scale-based scaffold modelling techniques which have evolved in the development of a new field of computer-aided design and modelling named computer-aided-TE (CATE) [26]. CATE methodologies are, in a general survey, based on the utilisation and integration of: (1) imaging technologies such as non-invasive computed tomography (CT) or magnetic resonance imaging (MRI) for anatomical modelling and 3D reconstruction, (2) computer-aided design (CAD), (3) computer-aided manufacturing (CAM), and (4) rapid prototyping (RP) together with solid freeform fabrication (SFF) technology in TE. Another important branch of CATE includes numerical/computational methods for scaffold modelling: computational fluid dynamics (CFD) technique, allowing for numerical modelling and analysis of the micro-fluidics and transport phenomena within the interconnected pores of the scaffolds, and structural analysis and modelling, allowing for the investigation of the mechanics of the scaffold and/or the growing tissues, will be discussed in the next sections.

As a basis for such a process, it appears evident the need for as well accurate methods for native tissue visualisation and modelling. CATE has been successfully applied in computer-aided tissue anatomical modelling, 3D anatomy visualisation and 3D reconstruction, as well as for the design of computer-assisted devices supporting surgical procedures for tissue constructs implantation [26]. The objective of this section is to furnish an overview of the most recently and innovative CATE techniques, with particular attention on multi-scale-based approaches, and of the most promising and successful obtained results.

13.2.1 CATE: Computer-Aided Anatomical Tissue Representation, CT and MRI Techniques

Based on the final objective, anatomical modelling techniques can be classified in two primary areas: (1) development of artificial replacement for tissues where the characterisation of natural tissue behaviour is needful, in

order to reproduce the natural environment realistically; and (2) tissue modelling related to the diagnostic area through the use of artificial materials, mathematical approach and continuum formulations.

In anatomical modelling, two non-invasive imaging techniques are successfully used: computer tomography (CT) and magnetic resonance imaging (MRI). CT is a powerful non-destructive evaluation technique resulting in the acquisition of 2D X-rays cross-sectional images of an object: digital processing is then used to generate 3D models of the object from a series of 2D X-ray images, taken around a single axis of rotation. Technically, in CT, an X-ray beam penetrates the object and the transmitted beam intensity is measured by an array of detectors passing the collected data to a dedicated PC unit capturing and processing images to give the final 3D reconstructed image. Characteristics of the internal structure of an object such as dimensions, shape, internal defects, and density are readily available from CT images: the test component is placed on a turntable stage that is between a radiation source and an imaging system. The imaging system provides a shadowgraph of an object, with the 3D structure compressed onto a 2D plane. The density data along one horizontal line of the image are then uncompressed and stretched out over an area. From this information, when the test component is rotated and similar data for the same linear slice are collected and overlaid, an image of the cross-sectional density of the component begins to develop. A number of slices through the object can be reconstructed to provide a 3D view of internal and external structural details. Hence, CT can furnish closely spaced axial slices of patient anatomy that, when rejoined in the appropriate manner, fully describe a volume of a tissue. Each CT slice image is computed of tiny picture elements (pixels). Each pixel, in turn, is actually a small volume element (voxel) of patient tissue sampled by the CT scanner. Of the existing methods for generating an anatomical model of a physical part, CT can non-destructively dimension with steady accuracy internal as well as external surfaces [27]. However, CT scanning, because of the dose of radiation administered (approximately 30–40 mGy), cannot be extensively used on patients. In the last years, the advent of MRI has become increasingly popular for its ability to show subtle differences in soft tissue anatomy without the harmful effects of ionising radiation present in CT. MRI scanning is a non-invasive alternative that projects a 3D image of the soft tissues together with bone. MRI is an imaging technique used primarily in medical settings to produce high-quality images of the inside of the human body. MRI is based on the principles of nuclear magnetic resonance (NMR), a spectroscopic technique used by scientists to obtain microscopic chemical and physical information about molecules. MRI started out as a tomographic imaging technique and has advanced beyond a tomographic imaging technique to a volume imaging technique [28]. The basics of the functioning of the MRI are herein briefly described: the human body is mainly composed of water molecule each containing two hydrogen nuclei or protons. When inside the powerful magnetic field of the MRI scanner, the protons align with the direction of the applied

field. A second radiofrequency electromagnetic field is then briefly turned on causing the protons to absorb some of its energy. When this field is turned off, the protons release this energy at a radiofrequency which can be detected by a scanner. The position of protons in the body can be determined by applying additional magnetic fields during the scan which allows an image of the body to be built up. These are created by turning gradients coils on and off. Diseased tissues can be detected, using MRI, because the protons of tissues with different properties and characteristics return to their equilibrium state at different rates. By changing the parameters on the scanner, this effect is used to create contrast between different types of body tissue. Furthermore, contrast agents may be injected intravenously to enhance image quality. While CT provides good spatial resolution (the ability to distinguish two structures at an arbitrarily small distance from each other as separate), MRI provides comparable resolution with far better contrast resolution (the ability to distinguish the differences between two arbitrarily similar but not identical tissues). MRI provides much greater contrast between the different soft tissues of the body than CT. Disadvantage of using MRI is the length of time that a patient is required to be exposed and remain motionless during scanning. Table 13.1 summarises a comparison of the basic imaging characteristics commonly used in producing 3D reconstruction CT and MRI [26].

Micro-computed tomography (micro-CT) is a ‘miniaturised’ version of computerised axial tomography: by means of micro-CT, a voxel resolution of the order of a few micrometers can be achieved. Up to now, the use of micro-CT has been successfully used in different branches of science for the study of porous or cavity-containing objects such as metallic foams, electronics, stones, wood and composite polymers. In particular, micro-CT has been used to quantify complex geometries in 3D in great details: in bone biology, for instance, a great body of literature is concerned with the measurement of characteristics of the trabecular network [29]. In the 1980s, this technique has been extensively applied in the study of the 3D micro-architectural changes occurring in the various types of osteoporosis and the method has also provided interesting results in the survey of anti-osteoporotic treatments which can preserve bone architecture. More recently, it has been applied in the study of cancer diseases as well as for the study of properties and characteristic of both natural and new-manufactured biomaterials. By means of micro-CT imaging, images with voxel sizes less than $10\ \mu\text{m}$ are potentially achievable [30], making

Table 13.1. Characteristics comparison of CT and MRI [26]

Characteristics	CT	MRI
Matrix size (pixels)	512×512	256×256
Voxel size (mm)	$0.5 \times 0.5 \times 2.0$	$0.5 \times 0.5 \times 1.5$
Density resolution	4,096 levels (12 bit)	128 levels (16 bit)
Signal-to-noise ratio	High	Moderate

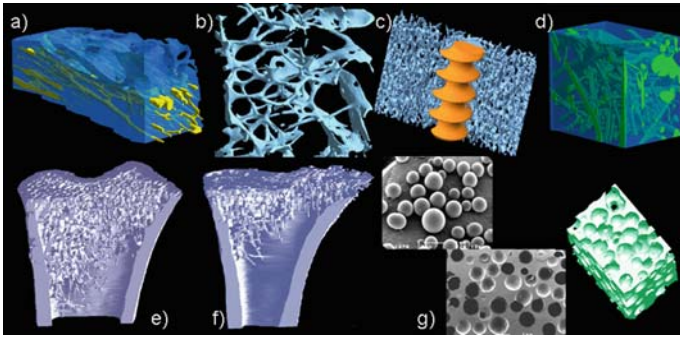


Fig. 13.2. 3D micro-CT reconstruction of (a) the cortical bone (*blue*) with a reconstruction of the Haversian canals (*yellow*) to illustrate the complex branching of these structures; (b) transiliac bone biopsy from a patient with glucocorticosteroid-induced osteoporosis; (c) a titanium surgical screw implanted in trabecular bone; (d) overimposition of the porosity (*green*) through synthetic biomaterial (*blue*) composed of poly 2-hydroxymethacrylate with an interconnected porosity; (e) tibial metaphysis of a control Wistar male rat and (f) tibial metaphysis of an orchidectomised (ORX) rat 16 week after surgery, highlighting the importance reduction in bone mass and the marked conversion of trabeculae into pillars; (g) SEM micrograph of polystyrene microbeads (*left*), porous poly 2-hydroxymethacrylate with porosity created by 1,200 nm polystyrene microbeads (*centre*) and (*right*) the same porous block analysed by micro-CT [http://www.med.univ-angers.fr/discipline/lab_histo/page_microCT.htm]

it superior to other techniques such as ultrasound ($30\ \mu\text{m}$) and magnetic resonance imaging ($100\ \mu\text{m}$), and mostly suitable for imaging in the field of TE concerning scaffold modelling and designing. Figure 13.2 displays the ability of this technique in micro-anatomical visualisation of the micro-architecture of porous materials, and its potency as toolkit for 3D reconstruction.

13.2.2 CATE: From Computer-Aided Anatomic 3D Reconstruction to Scaffolds Modelling and Design

By 3D reconstruction, volumes and surface areas of objects may be determined. 3D reconstruction of a volumetric data set is usually accomplished by extracting a region of interest (ROI), bounding surfaces to make a closed structure, defining the edges, and reconstructing the surface from ROI to ROI throughout the image set [31]. The process of the reconstruction of 3D anatomic models from CT/MRI data is described in Fig. 13.3. In the roadmap shown in this figure, the CT/MRI images are integrated using 2D segmentation and 3D region growth. This approach leads to the generation of 3D anatomical modelling, used for contour-based generation and 3D shaded surface representation and, finally, for the CAD-based models and manufacturing procedures. Figure 13.4 shows an example of 3D image reconstruction

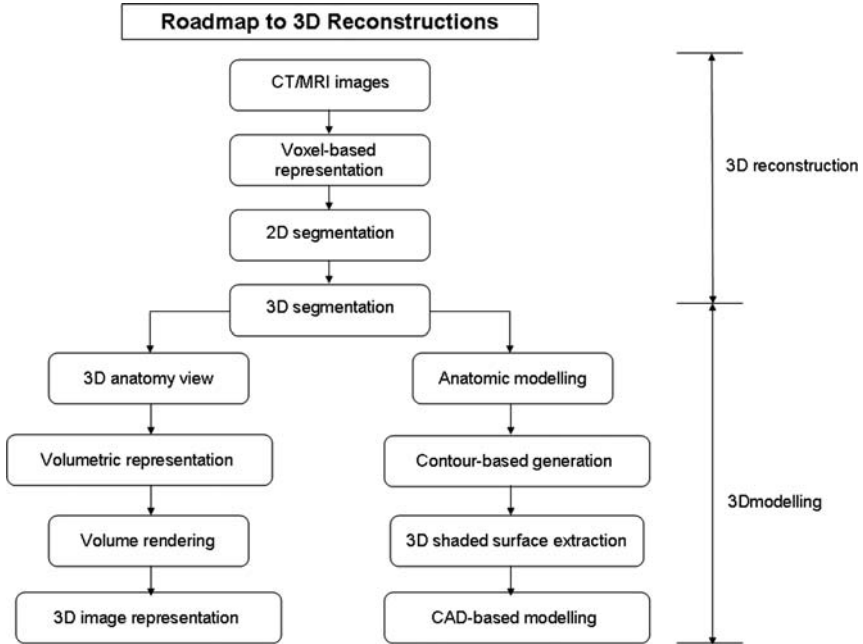


Fig. 13.3. From CT/MRI to 3D reconstruction (re-arranged from [26])

by means of CT/MRI-based techniques highlighting the potency of patient-specific provided images. Although imaging tools such as micro-CT/MRI are able to produce accurate 3D tissue descriptions, the voxel-based anatomical representation cannot be effectively used in several biomechanical fields of interest. Despite 3D volumetric models reproduce a realistic 3D anatomical appearance, they do not contain geometric topological relation. In general, activities in anatomical modelling design, analysis and simulation need to be carried out in a vector-based modelling environment, such as using CAD system and CAD-based solid modelling. Modern CAD systems use the so-called boundary representation (B-REP), in which a solid object is defined by the surfaces which bound it. These surfaces are mathematically described using special polynomial functions such as non-uniform rational B-spline (NURBS) functions. The use of NURBS to develop computer models can be applied and highly recommended in the fabrication of implants [32–34].

Unfortunately, the direct conversion of the CT data set of a human bone into its NURBS-based solid model is not simple. In the last few years, some commercial programs were presented as solutions to this conversion problem, for example, Surgi-CAD by Integraph ISS, USA; Med-Link, by Dynamic Computer Resources, USA; and Mimics and MedCAD, by Materialise, Belgium. One of the methods to fabricate physical models of anatomical prototypes and implants is to use CAD/CAM interfaced numerically controlled machine

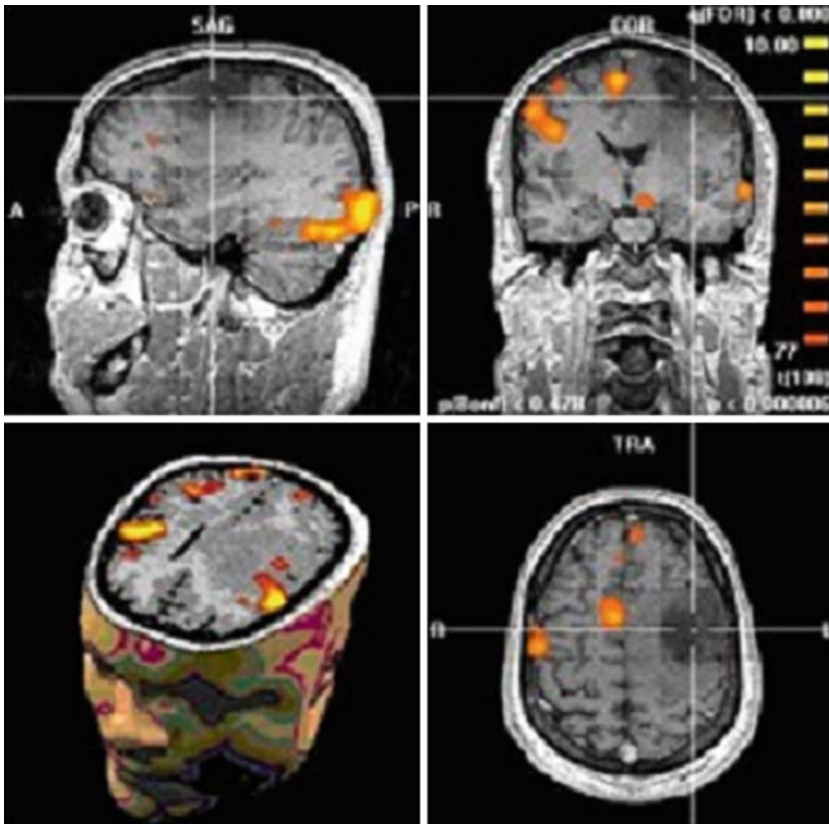


Fig. 13.4. A 3D-reconstructed CT/MRI-based image of the human brain and craniomaxillofacial anatomy [<http://www.unmc.edu>]

tools. The computer numerical control (CNC) machines fabricate prototypes by carving away material from the outside of a solid block or sheet of foam, plastic, wax or metal [35–39]. The limit of this method is the ability to create intricate structures, especially when there is a high degree of internal complexity (e.g., to prototype human craniomaxillofacial anatomy).

In the late 1980s, the introduction of Rapid Prototyping (RP) technologies offered new possibilities for medical modelling [40–43]. RP approach is based on the principle of layered manufacturing to create a model layer by layer. This approach lends itself readily to the free-form sculpture present in human anatomy. RP approach produces extremely detailed models that serve as excellent templates for the creation of custom implants. A precise physical model can offer an accurate prediction of implant size and type, and provide ‘hands-on’ surgical planning and rehearsal [44]. This leads to the opportunity of (1) direct production of casting molds for custom CAD designed to match patient data surgical implants, and (2) evaluation of prosthesis fit. Available

Table 13.2. RP processes used in anatomical-based scaffold modelling

RP technique	Description
• Stereolithography [45]	• Creates models by tracing a lower power ultraviolet laser across a vat filled with resin
• Selective laser sintering [46]	• Creates models out of a heat fusible powder by tracing a modulated laser beam across a bin covered with the powder
• Fused deposited modelling [47, 48]	• Creates models out of heating thermo-plastic material, extruded through a nozzle positioned over a computer controlled x-y table
• Laminated object manufacturing [49]	• Creates models out of heat-activated, adhesive coated paper, by tracing a focused laser beam to cut a profile on sheets positioned on a computer controlled x-y table
• Multiphase jet solidification [50]	• Creates metal or ceramic models out of various low-viscosity materials in powder or pellet form, by extruding the build material through a jet in liquid form
• 3D printing [51]	• Creates models by spraying liquid binder through ink-jet printer nozzles on to a layer of metallic or ceramic precursor powder

RP processes commonly used in prototyping anatomical modelling, as well as in scaffold manufacturing, permitting the development by the use of different material categories, are summarised in Table 13.2.

Once the 3D patient-specific reconstructed image of the tissue/organ has been achieved, the next step is the scaffold modelling and design stage. Traditional manufacturing methods (fibre bonding, blowing agents, solvent casting, gas foaming, particle leaching, phase separation, etc.) do not produce scaffolds with consistent, pre-established microscopic (i.e., large surface for cell adhesion and a volume permitting cell growth, proliferation and easy relative movement) and macroscopic structures (mechanical resistance) [52–54]. The use of these techniques is limited to producing scaffolds with simple geometry which may not satisfy the geometric requirements [55]. However, with the development of RP technologies based on CAD and CAM, these goals now appear easier to achieve. The processing of accurate images of the internal parts of the body (tissues and organs) provided by MRI and CT, combined with RP and Solid Freeform Fabrication (SFF) methodologies, allows to pre-define in the computer both the microscopic and the macroscopic shape and morphology of scaffolds, obtaining a final product with the desired proper

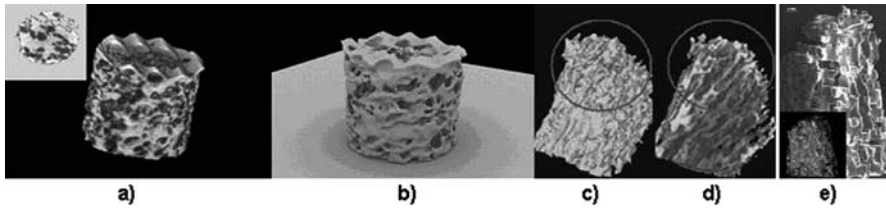


Fig. 13.5. (a) 3D micro-CT reconstruction and (b) CAD virtual reconstruction of porous HA; (c) internal sectional views of CAD reconstruction of (d) human bone micro-CT, (e) SEM micrograph of the SLP resin scaffold [45]

characteristics [26]. Once digitalised, the CT and MRI images can be analyzed by the CAD/CAM systems that guide the RP systems with ad hoc numerical control software [56–59]. Combination of conventional methods together with RP techniques and imaging applications has been as well tested for producing scaffolds [60]. In 2005, Quadrani et al. [45] experimented an application of RP techniques for bone-like scaffold construction: starting from a micro-CT based 3D reconstruction of a porous hydroxyapatite (HA) bulk, CAD processing to generate an iso-surface has been applied and finally the bone-like scaffold has been manufactured in acrylic resin by means of a stereolithography (SLP) with the aim of providing seeded cells within a 3D-tissue-like environment (Fig. 13.5). The integrated approach applied by Quadrani et al. [45] allowed to realise a synthetic scaffold with efficient bone biomimetic functions (Fig. 13.5), proving the effectiveness of the applied methodology in scaffold designing. Furthermore, morphological details of the engineered scaffold have been extracted in order to verify the properties matching of the scaffold with respect to the 3D micro-CT reconstructed HA structure.

The design of customised scaffolds through the use of RP techniques has been also performed in 2005 by Naing et al. [61], which developed a prototype CAD system called the computer aided system for tissue scaffolds (CASTS). Based on a library of open cellular polyhedral units that can be assembled into uniform cellular matrices of various shapes, the algorithm of the system is able to automatically generate a structure suitable for the specific intended application. Patient's data are collected by imaging techniques and through the use of image processing the surface profile of the tissue segment can be created. The scaffold model can be then generated (by choosing an appropriate unit cell from the parametric library and making use of the sizing routines), and, finally, the scaffold is customised using the patient's tissue profile by means of SLP. The prototype allows the user to properly alter the different morphological parameters, such as the shape and spatial arrangement of the polyhedral units constituting the model, so to reach a patient-specific engineered constructs. In this way, not only the consistency in terms of structural form can be achieved, but also porosity and pore sizes can be varied so to cater to different types of seeded cells. In fact, in bone tissue regeneration, for example, the

pore size of the scaffold dictates the type of cell which will penetrate and proliferate within the structure [62]: in bone tissue scaffolds, two types of pores can be defined: global or macropores (at least 100–300 μm in diameter), which are incorporated into the design, and interconnected local pores (minimum 10–12 μm diameter), which are created as a result of processing techniques such as porogen leaching [63]. Among the various structural parameters characterising a scaffold, pore size and porosity are the primary factors for in vivo ingrowth and reorganisation of tissue, as they directly affect the physiological diffusion of nutrients and waste products of cells metabolism. Increased porosity and interconnectivity induce tissue ingrowth [64,65] and interconnectivity ensures that the eventual tissue will be complete and integrated to the surrounding tissue. However, Bhatia and Chen [66] demonstrated that highly porous scaffolds do not always encourage cell proliferation, differentiation and reorganisation. In scaffolds with uncontrolled micro-architecture, for instance, pore sizes, can result in non-uniform cell density within the scaffold: occlusion may also occur at smaller pore sizes, thus preventing cellular penetration and matrix formation [64]. Hence, the development of RP techniques in scaffolds fabrication, helping controlling micro-architecture, is required in order to achieve the desired outcome. Following this design philosophy by means of the CASTS system, Naing et al. [61] designed defined patient-specific scaffolds with different architecture and properties (Fig. 13.6).

A further example of the importance of the micro-structure of healthy human bone, which is correlated with its anatomic function [67–71], has been provided by Klein et al. [72], who analysed hydroxyapatite (HA) samples

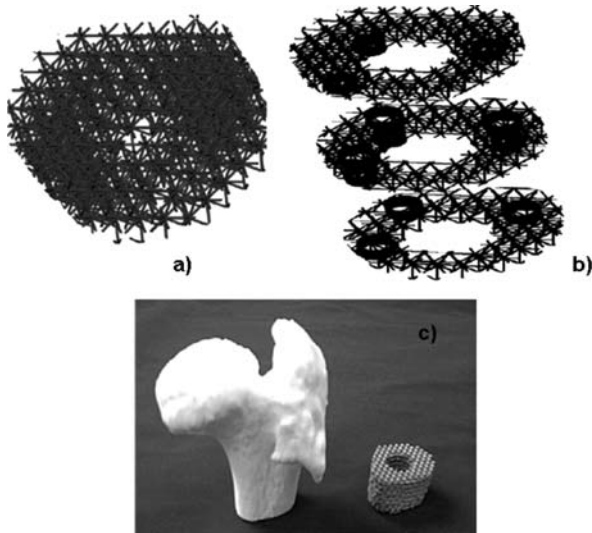


Fig. 13.6. Femur scaffold model (a) before and (b) after slicing (c) a femur segment (left) and the SLP fabricated customised scaffold (right) [61]

Table 13.3. Investigated bone substitute materials [72]

Chemical composition	Origin	Commercial name	Granulation size (μm)
β -TCP	synthetic	Cerasorb classic,	50–150, 150–500, 500–100, 1,000–2,000
		Cerasorb M,	150–500, 500–1,000, 1,000–2,000
		ChronOs	700–1,400
Biological apatite	Phycogenic bovine	Algipore,	500–1,000, 1,000–2,000
		Bio-Oss	250–1,000, 1,000–2,000
Biphasic β -TCP, HA	synthetic	Strauman Bone Ceramic	400–700, 500–1,000
Silicate	synthetic	NovaBone-O/M,	90–710
		PerioGlass	90–710
Biphasic SiO_2 , HA	synthetic	NanoBone	$600 \times 2,000$, $1,000 \times 2,000$

of biological origin (Algipores and Bio-Osses), of synthetic origin (Cerasorbs classic, Cerasorbs M, ChronOss), pure bioactive glass (Nova-Bones, Perio-Glass), and composed by biphasic materials (Straumanns BoneCeramic and NanoBones), with different granulation size (Table 13.3).

By high-resolution micro-CT, reconstruction and modelling of cancellous human bone sample has been achieved (Fig. 13.7). On the basis of the previously collected data, characterisation of the porosity of the commercial scaffolds has been carried out: dimensions (thickness) and distribution of solid particles, interconnecting pores and micro-cavities, have been extracted by means of an algorithm based on the direct distance transformation method [73], starting from micro-CT reconstructed models of the scaffolds [72]. An example of the achieved results is shown in Fig. 13.8. The generation and interpretation of micro-CT-based aggregate pore models provide further insight into the expected osteoconduction and osteoinduction dynamics and therefore might serve as a basis for further modifications of scaffold size and geometry as well as for further systematic investigations of the *in vivo/in situ* behaviour of bone substitute materials. A further application in scaffold modelling by micro-CT imaging methods has been reported by Rich et al. [74] who tried to pre-determine the porosity of biodegradable scaffolds manufactured by a novel selective leaching method for the porogenisation of an ϵ -caprolactone-based scaffold (Fig. 13.9). In detail, using a well-defined multi-scale approach, the amount of the porosity following the leaching method was estimated and modelled by micro-CT analysis: a micro-CT system was used to image and quantify the 3D micro-structural morphology of the samples, at 16- μm voxel resolution, as shown in Fig. 13.10. Micro-CT images identify the different pore structures obtained depending on the form of the porogen used. Interconnectivity of the scaffold pore network, estimated with cubic thickness model in micro-CT,

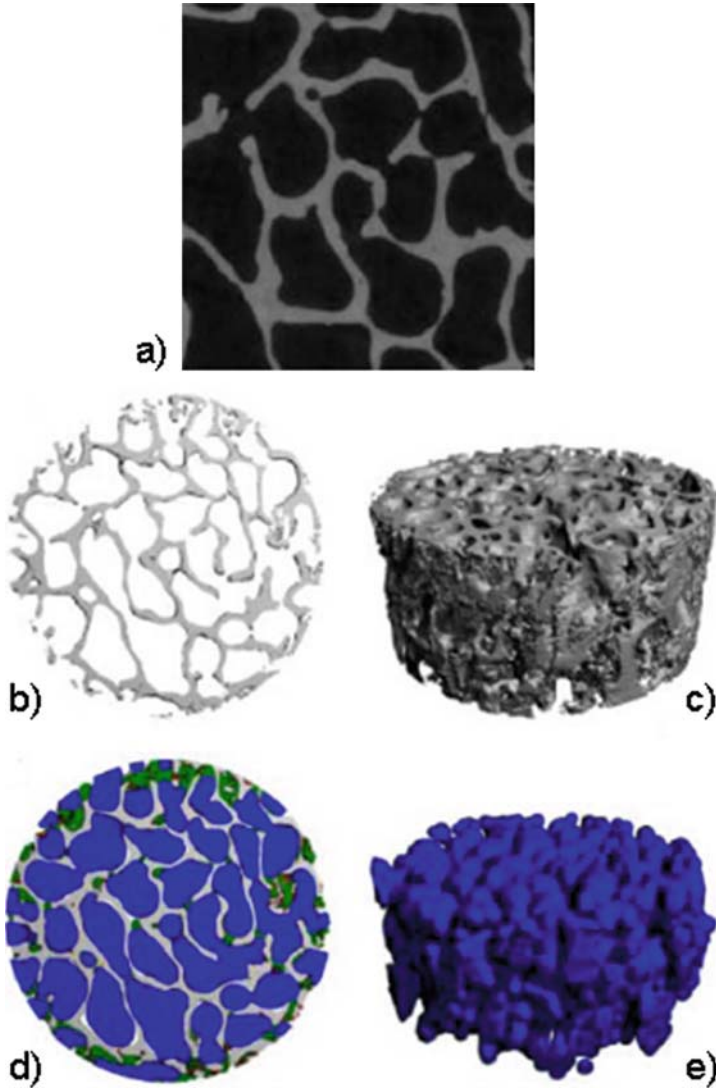


Fig. 13.7. (a) Detail of a cross-section obtained from a cancellous human bone sample; (b) and (c) 2D and 3D reconstructions of the cancellous human bone sample; (d) and (e) colour-encoded pore distribution (pores $< 60\ \mu\text{m}$: *red*, pores $60\text{--}250\ \mu\text{m}$: *green*, pores $>250\ \mu\text{m}$: *blue*) [72]

confirm that the pores penetrated the device. In this case, micro-CT imaging has been successfully applied in an attempt to optimise the requirements for the design of the scaffold.

Imaging techniques have been also applied in the effort of analysing and characterising the microstructure of the manufactured scaffold during the cell

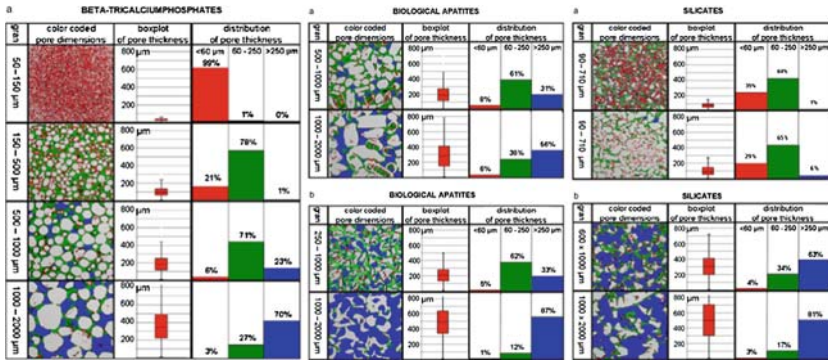


Fig. 13.8. Characterization of porosity of different commercial materials for bone substituting. The described experimental methodology offers quantifiable information about 3D pore configurations [72]

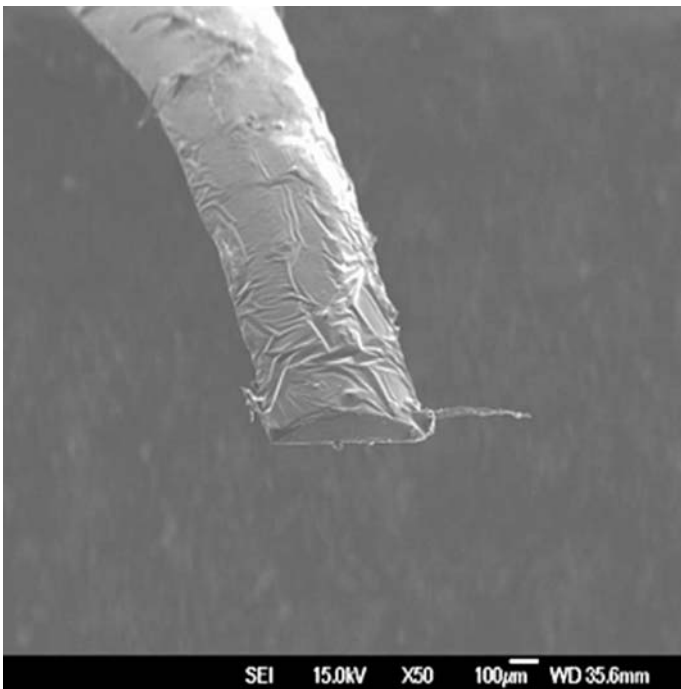


Fig. 13.9. SEM micrographs of porogen fibre [74]

growth and new tissue formation within the scaffold. In 2007, Jones et al. [75] used 3D micro-CT imaging for the assessment and the quantification of the 3D structure of a porous scaffold for bone ingrowth, aiming at visualising the regenerative processes within the scaffold structure. The study was performed on porous bioceramic-based scaffolds, which have great potential

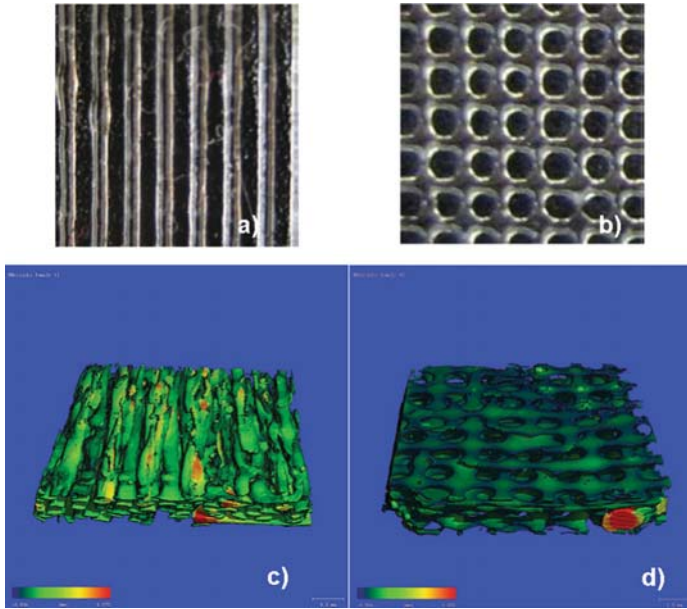


Fig. 13.10. 3D reconstructed visualizations of (a) scaffold phase, (b) scaffold with bone phase added (due to cellular ingrowth), and 3D network structure of the (c) initial scaffold phase and (d) after morphological rearrangement (due to cellular ingrowth) [74]

in providing solutions for bone replacement. Even though micro-CT technique cannot provide all the data like a standard histological analysis, this technique has substantial potential advantages in identifying scaffold characteristics and bone ingrowth parameters in 3D [76, 77]. Indeed, micro-CT imaging could suffer from limitations such as insufficient resolution and inaccurate phase identification within the scaffold volume. This is due to the fact that a quantitative characterisation of the initiation, onset and continuing growth of mineralised tissue within porous specimens (exhibiting pore sizes of 50–300 μm) requires imaging with resolutions of the order of 10 μm : only imaging at this resolution allows to accurately map pore size and structure, to quantify bone ingrowth phase fractions within individual pores and to study transport properties within the porous structure. Moreover, the complex process of bone remodelling inside a tissue-engineered construct, made up of scaffold material, host bone, mineralised bone and soft tissue, makes the partitioning of the tomogram into discrete phases non-trivial. In the attempt to settle the previous issues associated with the use of micro-CT, Jones et al. [75] implemented a highly sophisticated three-phase segmentation approach to develop a practical phase separation of the tissue-engineered construct with minimal artefacts. In particular, the developed imaging technique has resulted in analysing and quantifying the following scaffold properties: (1) porosity,

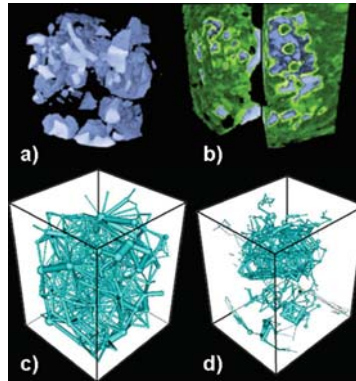


Fig. 13.11. Photograph of (a) paralleled fibre porogens and (b) porogen mesh; micro-CT image of (c) the pore structure within the matrix porogenised with paralleled fibres and (d) the pore structure within the matrix porogenised with porogen mesh [75]

(2) bone ingrowth volume, (3) pore volume and (4) fraction of pore occupied by bone, furnishing a detailed description of the changes in scaffold morphology over time during the regenerative process of the bone within the scaffold (Fig. 13.11). This example illustrates the potential of the micro-CT-based imaging techniques combined with advanced segmentation algorithms, in bone ingrowth assessment. Referring to Fig. 13.11, the 3D visualisation shows the significant extent of the ingrowth through the centre of the scaffold. Finally, micro-CT-based imaging and modelling techniques appear to be extremely helpful tools to be applied not only in computer-aided scaffold designing but also in a non-destructive evaluation of the modification in scaffold morphology and characteristics during *in vivo/in vitro* new tissue formation as well.

In vivo micro-CT imaging has been applied also in tissue-engineered vascular conduits applications [78]. Tubular autologous vascular conduits in which cells can be seeded *in vitro* or directly *in vivo* can be used as vascular grafts exhibiting, theoretically, many of the properties of an ideal vessel including antithrombogenicity, durability and even the potential for new tissue growth [79]. The development of a tissue-engineered vascular conduit with growth potential would have a significant impact on the treatment of vascular defects in particular on patients in the paediatric population and could potentially expand the boundaries of current surgical practice. Clinical trials evaluating the use of tissue-engineered autologous venous conduits for surgical reconstruction of congenital cardiac anomalies are ongoing [80]. Although preliminary results from these studies are quite promising, there is a paucity of data pertaining to their long-term function. As these data become available, design deficiencies will be identified and primary modes of graft failure will be determined. The opportunity of a non-destructive *in vivo* evaluation can be used to redesign grafts in an attempt to improve graft function. In animal

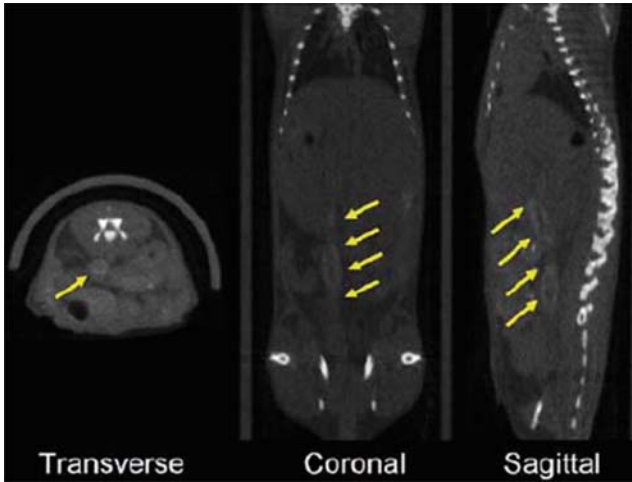


Fig. 13.12. Micro-CT images of tissue-engineered aortic interposition grafts in a mouse model on postoperative day 35. The arrows in the coronal and sagittal sections demonstrate the marked thickening of the wall of the interposed graft. In the sagittal section, the luminal narrowing that can be seen between the proximal and distal ends of the graft is demonstrated [78]

models developed by Goyal et al. [78], small-calibre ovine decellularised arteries (0.5–1.0-mm diameter) have been used as scaffold for neovascularisation of end-to-end aortic, end-to-end inferior vena cava, and side-to-side aortocaval segments by means of a 1-mm diameter aortic interposition graft. Graft function has been serially monitored using micro-CT, as shown in Fig. 13.12. The benefits of the vascular graft interposition has been first analysed by means of the computer-aided anatomical tissue representation and imaging methods (micro-CT imaging) and afterwards validated by histological analysis that confirmed the observed luminal patency: evidence of cellular ingrowth into the TE scaffold was noted, as was evidence of non-occlusive wall thrombus and neointimal hyperplasia, sustaining the data acquired with the micro-CT. The study of Goyal et al. [78] demonstrates the feasibility of developing imaging-based validation tools of *in vivo* models to provide a cost-effective, non-destructive and rapid screen/analysis method for vascular TE applications. This assumption has been corroborated by Bolland et al. [81], who used contrast-enhanced micro-CT for the *in vivo* evaluation of neovascularisation in tissue-engineered bone constructs.

Moreover, a key issue in fabricating clinically translatable regenerative tissue is the ability to generate a functional micro-vascular network within engineered constructs to provide oxygen (O_2) and nutrients that facilitate growth, differentiation and tissue functionality [82,83]. Techniques employed to study the micro-vascular architecture and angiogenesis include histology, immunohistochemistry [84] *in vitro* gels [85], animal perfusion [86], vascular

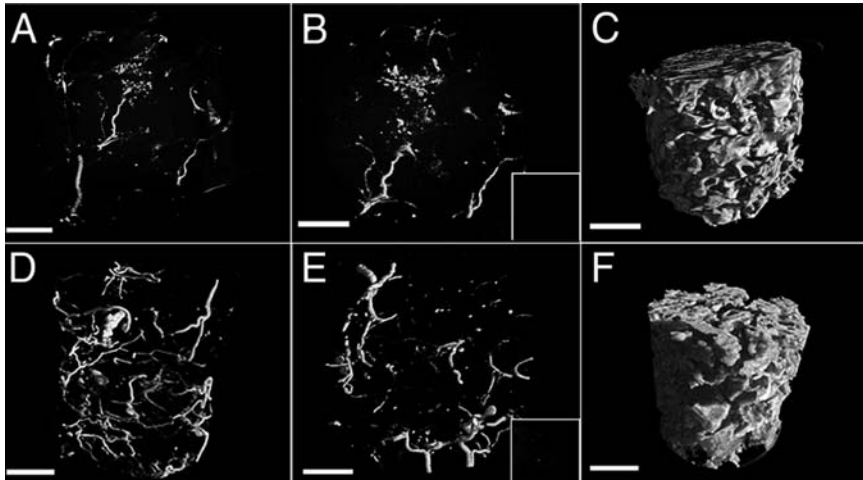


Fig. 13.13. Micro-CT visualization demonstrating new vessel formation (**A**) outside and (**B**) inside the allograft compared with (**D**) outside and (**E**) inside the PLA scaffold; micro-CT 3D reconstruction of the (**C**) allograft and (**F**) PLA scaffold [82]

casts [87]. However, each of these techniques holds limitations in aiding the understanding of angiogenesis: to adequately understand angiogenesis, an extremely sensitive, quantifiable assay of 3D tissue-engineered structures is required to measure vessels number, diameter and vascular network density. Micro-CT has been successfully used for *in vivo* evaluation of angiogenesis [88–93]. Based on this, the detection and quantification of neovascularisation in tissue-engineered constructs by micro-CT appear to be feasible (Fig. 13.13).

An additional example regarding multi-scale scaffold modelling based on imaging methods is provided by Cioffi et al. [94] who setup a computational model, based on micro-CT imaging, for the numerical simulation of the fluid-induced mechanical stimulation in scaffold for articular cartilage replacement where chondrocytes are cultured. In chondrocyte-seeded constructs, ECM synthesis is particularly promoted by hydrodynamic flow [95], which results in cell membrane stretching, caused by fluid-induced shear, and in enhanced convective transport of nutrients to the cells and of catabolites away from the cells. The shear stress imposed to chondrocytes in such systems will depend not only on the culture medium flow rate through the constructs but also on the scaffold 3D architecture [2]: it appears evident the need of experimental set-up able to monitor shear stress values imposed to cells inside a realistic 3D scaffold model, during new tissue formation. For such analysis a multi-scale approach is required. Analytical and computational methods can be a key tool for the design of TE experiments with controlled fluid–dynamic conditions (further details about multi-scale fluid dynamics modelling of tissue-engineered scaffold will be provided in the next section). The optimal flow conditions (typically imposed by the use of perfusion bioreactors) should not be determined

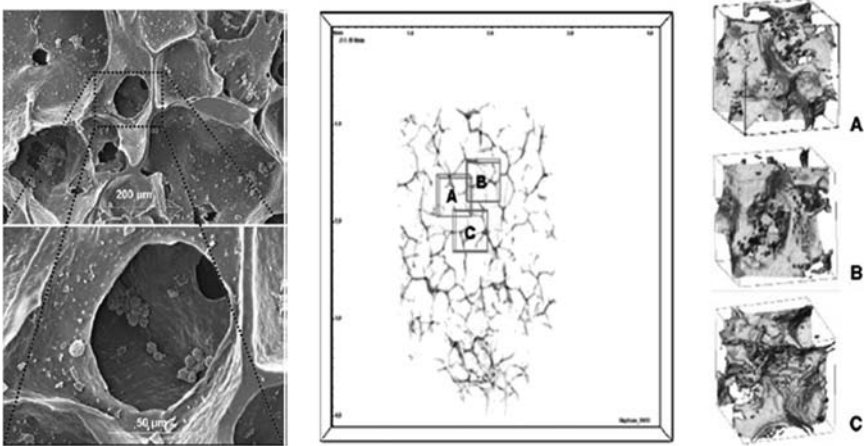


Fig. 13.14. SEM micrograph (*left panel*) of the biodegradable Degrapol scaffold seeded with chondrocytes and (*right panel*) the method employed to build geometrical models of the scaffold microstructure: three areas A, B and C were selected in a representative micro-CT image of the scaffold cross-section and three solid models (A, B and C) were reconstructed from incremental sections in the selected areas [94]

by a trial-and-error approach, but rather should be predicted by simulation methods [96]. A quantitative description of the fluid-dynamic domain in a micro-structured scaffold can be obtained by computational fluid-dynamics (CFD) simulations which allows for predicting locally at the scaffold internal walls the fluid-induced shear. In the study by Cioffi et al. [94], the characterisation of the local fluid dynamics to which chondrocytes are exposed inside constructs obtained using a porous scaffold made of a polyurethane foam (Degrapol), cultured in a direct-perfusion bioreactor system, has been achieved by the combination of micro-CT imaging and CFD simulations. In Fig. 13.14, the 3D micro-CT-derived geometrical model for CFD simulation is shown, reproducing the micro-structure of the scaffold morphology. Once the model has been constructed [94], the flow field within the porous scaffold and the shear stress conditions experienced by seeded chondrocytes have been numerically evaluated by means of a computational software-based on the finite-volume technique (Fig. 13.15). The study by Cioffi et al. [94] confirms that computational modelling can be successfully used in combination with micro-imaging techniques, to quantify, on different scale levels, the shear stress artificially applied to cells in 3D engineered cellular bioreactors: these predicted data provide insight into the potential mechanical interaction between the cells and their surrounding environment, and may be of practical use to investigate mechanisms of mechanical signal transduction. Furthermore, computational models may provide new scaffold design criteria, such as indications on the microstructure that could optimise transport phenomena to and from the cells.

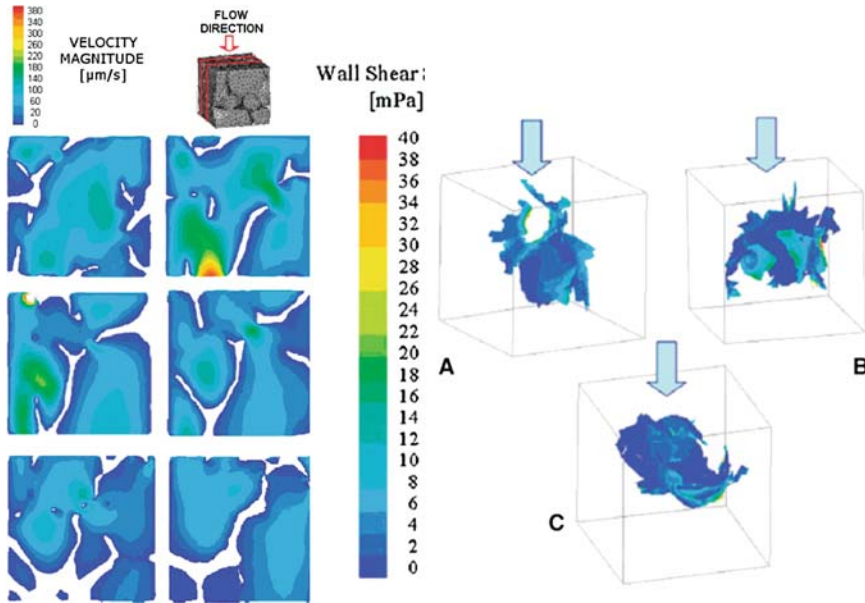


Fig. 13.15. Numerically quantification of the velocity magnitude field, visualised on six cross-sections (*left panel*), and of the wall shear stress magnitude distribution within inner regions of the geometrical model (*right panel*) [94]

The determination of optimal design requirements of bioreactors for cellular growing has been investigated by Porter et al. in 2007 [97], who analysed, by means of non-invasive imaging-based analysis, the 3D construct mineralisation within a bioreactor perfusing bone grafts for the challenging treatment of bone defects. Perfusion of osteogenic medium increases the amount of mineralised matrix produced by cells seeded within 3D scaffolds in a flow rate-dependent manner [98]. Moreover, in addition to improving mass transport, fluid flow delivers shear stresses that are well known to modulate cell function [99–102]. Therefore, perfusion bioreactors offer the potential to develop larger, more mineralised constructs for implantation and can also serve as physiologically relevant 3D model systems for studying cell–scaffold interactions *in vitro* as well as to further test novel TE strategies prior to work in animal preclinical models. Although the beneficial effects of perfusion on cell-mediated mineralisation have been demonstrated in several studies, the scaling up to clinically relevant dimensions of the mineralised constructs is still a limiting factor. In order to study the feasibility of different sizes constructs, Porter et al. [97] applied micro-CT imaging together with quantifying algorithms for monitoring and assessing mechanisms of construct mineralisation within constructs varying in thickness (Fig. 13.16). Number, size and distribution of mineralised particles formed within constructs have been quantified. The analysis of multiple concentric volumes inside each construct indicates that a great

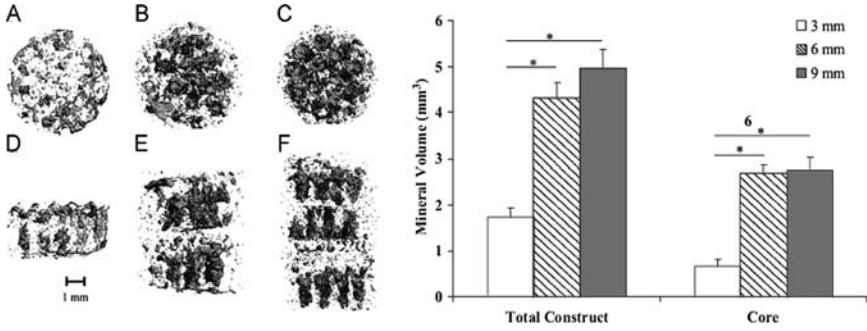


Fig. 13.16. Representative micro-CT images of mineral deposits within stacked PCL + Collagen scaffolds (*left panel*), and mineralised matrix volume for perused constructs of three different stacked thicknesses (*right panel*) after 5 weeks of perfusion; images A–C show the distribution of mineral looking at the constructs from the top, images D–F depict mineral deposits throughout the full thickness of the constructs. A, D = 3-mm-thickness scaffold, B, E = 6-mm-thickness scaffold, C, F = 9-mm-thickness scaffold [97]

proportion of the mineral volume is localised within the interior of the perused constructs. Interestingly, intermediate-sized thick constructs were found to have the highest core mineral volume fraction and the largest mineralised particles. The ability to quantify mineralised matrix formation non-invasively within 3D constructs would benefit efforts to optimise bioreactor conditions for scaling-up constructs to clinically relevant dimensions.

The all above reported studies reveal the ability of the micro-CT-based imaging methods, together with RP techniques, to obtain reliable models of the micro-architecture of various porous scaffolds typologies. The result is a provided precise and controlled design and manufacture procedure of such scaffolds. Furthermore, as seen, modelling of the scaffold morphology allows for highly detailed analysis of different properties and characteristics parameters of the scaffold that primarily influence the new tissue formation within the engineered constructs, together with a non-destructive *in vivo* monitoring of the progressive cellular growth, at different stages of formation of the new tissue. The final outcome of this multi-scale-based methodology is the optimisation and customisation of the properties of the scaffold affecting the cellular growth protocol. Data can be easily extracted and subsequently applied for the development of numerical, computational and experimental setup, modelling the overall behaviour of the engineered constructs. This leads to an even better effective assessment of the techniques required to select or design scaffolds with suitable properties, right at the initial stages of the research and development; at the same time, a more in-depth knowledge and a comprehensive overview of the mechanisms involved and responsible for ‘active’ cell–substrate interactions can be obtained.

In conclusion, the evaluation of scaffold architecture parameters is necessary as it facilitates the improvement of the scaffold design process and its eventual biomedical applications. Hence, an appropriate technique needs to be selected in analysing the scaffold and an understanding of the potentials and concerns of the technique is crucial. In this optics, micro-CT is a rather new technique but it has demonstrated various key advantages: it is non-destructive, hence samples remain intact for further assays so it appears to be largely suitable in multi-scale modelling/design/manufacturing ECM-like scaffold for TE applications.

13.2.3 CATE: FEM and CFD-Based Scaffolds Modelling and Design Methods

A further branch of CATE methodologies, extensively used in scaffold modelling for TE applications, includes numerical and computational methods for solving the fluid and solid mechanics features involved in scaffold design.

Computational fluid dynamics (CFD) modelling is a technique based on the development of numerical methods and algorithms allowing to solve and analyse problems in which fluid mechanics aspects are involved. As a first result, flow characteristics, as a function of the medium flow rate in perfusion bioreactors, within the porous scaffold can be investigated. CFD analysis, together with multiple scales modelling can be a powerful method for investigating how the scaffold micro/nano-structure affects the flow field [103, 104]. By means of CFD-based modelling, various features related to the flow field characteristics can be analysed such as: (1) O₂ and nutrients supply to the viable cells (transport phenomena), (2) cellular O₂ consumption and catabolites production rates (cellular metabolic activities), and (3) drag forces and shear stresses (fluid-induced mechanical stress) experienced by cells; in addition, the optimised 3D scaffold geometry and micro-architecture affecting all these features can be obtained, for a reliable design of the scaffolds. The evaluation of the previous parameters is extremely difficult to be achieved by means of experimental measure techniques.

Cells growing in an *in vivo* organ have several characteristics such as (1) 3D shape for cell attachment and spatial organisation, (2) high density and (3) a homogeneous and high-level supply of O₂ and nutrients for cell sustainment, which are difficult to reproduce during *in vitro* cells growing in culture dishes, as well accounting partly for the differences between *in vivo* and *in vitro* cell behaviour [105]. The first condition can be realised by using a 3D scaffold fabrication process [63, 106, 107] and/or using highly porous material with an optimised surface/volume ratio [108]. Moreover, 3D fabrication processes also allow for better control of internal geometry of the scaffold, in an attempt to mimic the native ECM morphology [109, 110]. The second condition is of importance because cell density can regulate growth, differentiation [111, 112], metabolic activity, function and morphology of cells [113].

Furthermore, cells phenotypes as of liver, kidney, skeletal muscles and others are anchorage-dependent cells, forming multi-clustered structures in order to perform correctly their specific functions: to achieve this condition in vitro high cell density culture, within a 3D organised scaffold, is required. The third condition can be fulfilled by using more 'complex' cultivation systems where perfusion of medium through the microfluidic channels between the porosity of the scaffold is possible. Many scaffolds or reactors designed for high cell density culture have been reported [108, 114–117]. Despite of these attempts, nowadays, scaffolds combining volume in the micro-scale range with the possibility of immobilising a high cell density ($>10^7$ cells ml^{-1}), together with a controlled internal geometry, are not yet available. Maximising O_2 and nutrient delivery to the viable cells is critical for optimal in vitro culture conditions: insufficient oxygenation is, indeed, the main limiting factor for cell sustainment in in vitro perfusion bioreactors, resulting in the most common cause of performance decrease for high-density cells cultivation [118]. Moreover, the main challenge for successful cell culture at high density under perfusion is the trade-off between a high supply of medium and a sufficiently low shear stress [119, 120].

To date, several studies have proved the usefulness of CFD-based modelling in analysing fluid dynamics and transport of chemical species in scaffold and bioreactor models [96, 103, 104, 118, 121–128]. Briefly, the solution of a CFD model is fundamentally based on the numerical solution of the Navier-Stokes equations, which are differential equations describing the motion of a fluid phase. The most fundamental consideration in CFD is the discretisation of a continuous fluid phase: one method is to discretise the spatial domain into small 'cells' (uniform or non-uniform) to design a volume mesh or grid, and then apply a suitable algorithm to solve the equations of motion for the discretised fluid domain. For this purpose, the Finite Volume Method (FVM) is the standard approach used most often in commercial software and research codes treating fluid phase problems. High-resolution discretisation schemes allow for modelling a variety of geometries, even if spatial discontinuities are present in the fluid domain. In many instances, other equations are solved simultaneously with the Navier-Stokes equations: these can include those describing species concentration (mass transfer), chemical reactions, heat transfer, etc. More advanced computational codes provide as well for the simulation of more complex cases involving multi-phase flows (e.g., liquid/gas, solid/gas, liquid/solid), non-Newtonian fluids (such as blood), or chemically reacting flows.

In developing a CFD-based model, the following general procedure can be performed: the first step (i.e., the geometrical model design) includes (1) the definition of the geometry (physical bounds) of the problem, (2) the partition of the volume occupied by the fluid into discrete cells (the mesh), (3) the definition of the physical problem (e.g., combining the equations of motions with those for enthalpy or species conservation calculations), (4) definition of the boundary and, for transient flows, of the initial conditions (this involves specifying the fluid behaviour and properties at the boundaries of the problem

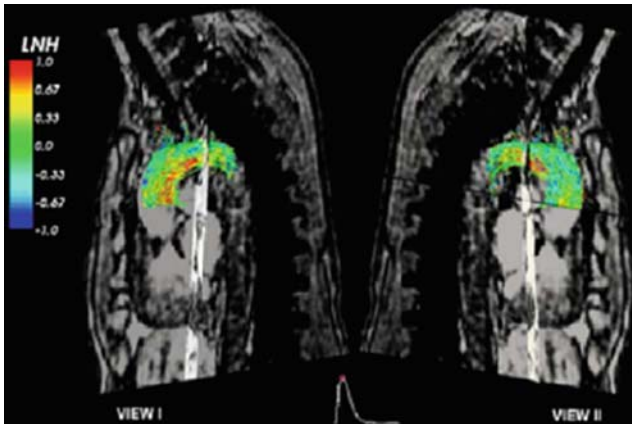


Fig. 13.17. The evolving flow patterns in aorta, obtained by means of a Lagrangian analysis, on the particle set emitted at peak systole, fused to the modulus image of the aorta; two different views put in evidence prevalence of helical content opposite in sign towards the inner and the outer wall [135]

and at the beginning of the numerical simulation); the next step is the running of the numerical simulation (equations, depending on the previous settings, are solved iteratively as a steady-state or transient) and finally a postprocessor is used for the analysis and visualisation of the resulting solution. As for the solution algorithms, discretisation in space produces a system of ordinary differential equations for unsteady problems and algebraic equations for steady problems. Implicit or semi-implicit methods are generally used to integrate the ordinary differential equations, producing a system of (usually) nonlinear algebraic equations [129–134]. As a case study of the enormous potential of CFD-based methods applied to the regenerative medicine field, we mention the study by Morbiducci et al. [135], who applied CFD algorithms to time-resolved 3D cine phase contrast (PC) MRI for *in vivo* quantifying the helical blood flow in the human aorta, thus addressing the challenging issue of a quantitative description from the *in vivo* available information (Fig. 13.17).

Regarding our concern in scaffold design, by applying approaches similar to the one presented above, significant clues for the design of the culture environment can be retrieved. Recently, Cantini et al. [136] used CFD modelling to optimise the microenvironment inside scaffolds for hematopoietic stem cell (HSCs) culture in a perfusion bioreactor, and the fluid dynamic conditions so to avoid non-adherent stem cells being dragged away while ensuring adequate nutrient supply during perfusion. Given their potential, HSCs have been used routinely in clinical practice for more than three decades, particularly in the form of bone marrow transplantation [137–140].

Recent studies have proved that HSCs might be used to treat several pathologic conditions [141]. Several types of bioreactors have been designed for the expansion of HSCs: perfusion chambers, stirred/rotating, hollow fibre,

and packed bed reactors [142–146]. The approach that seems to be most promising is the mimicry of the *in vivo* physiological microenvironment for HSCs (i.e., the bone marrow) [142, 147]. To achieve this goal, it is necessary to design 3D porous substrates, which have to be structured in a proper way for the dynamic co-culture of two different cellular populations: HSCs and bone marrow stromal cells (BMSCs). A reliable design of these systems would require an accurate evaluation of local variables, such as drag forces and shear stresses acting on the cells or nutrient concentration inside the scaffold pores. CFD technique validly allows for modelling all these issues. Furthermore, the CFD-based modelling permits these studies to be carried out at two different scale levels: at a microscopic level for the analysis of fluid dynamics inside the idealised or realistic 3D micro-architecture of the scaffolds used for cell culture [103, 122, 124] and at a macroscopic level for the analysis of the global structure of the culture chamber of a bioreactor [121, 126, 148, 149]. The CFD geometrical model developed by Cantini et al. [136], for the fluid dynamic microenvironment characterisation, is based on the simplification of the geometry of synthetic micro-porous scaffolds obtained using the solvent casting/particulate leaching technique. The experimental procedure of particulate leaching consists of manufacturing the template of a porous scaffold via packing and interconnecting solid particles (porogen), casting and evaporating a polymer solution over them, and finally removing the particles by leaching. The technique allows scaffold morphology to be controlled easily by varying porogen size, shape and interconnection. To analyse the flow through the void porous microstructure, the geometrical model developed consists of interconnected spheres representing the pores where culture medium flows and HSCs are homed, whereas BMSCs are assumed to be spread on the surface of the pores and therefore to coincide with the walls of the model (Fig. 13.18). Fully developed flow inside the scaffold was modelled using periodic inlet and outlet boundaries, because of the periodically repetitive nature of the physical

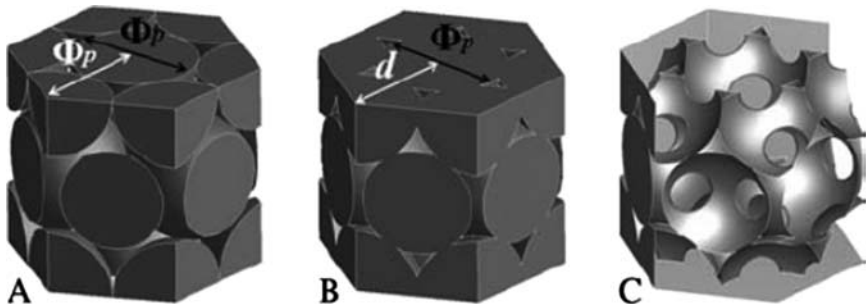


Fig. 13.18. Hexagonal close packing of spheres: (A) noninterconnected spheres ($d = \Phi_p$); (B) interconnected spheres ($d < \Phi_p$); (C) walls of the pores of the basic unit modeling a porous scaffold. Φ_p is the diameter of the spheres (pores) and d the distance between the center of two interconnected spheres (pores) [136]

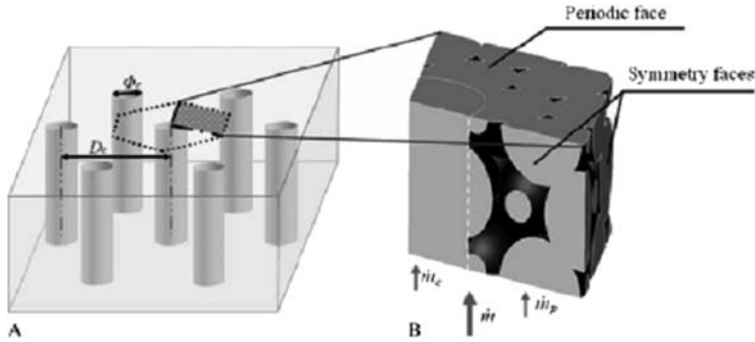


Fig. 13.19. (A) Schematic representation of a channel-provided scaffold and (B) geometric model of a periodic unit, with indication of boundary conditions and mass flow rate distribution between channel and porous medium; Φ_c is the diameter of the channels and D_c the distance between them [136]

geometry of interest and of the expected pattern of the flow solution. Another hypothesis tested in this work is whether the addition of longitudinal micro-channels to a homogeneous porous scaffold is able to improve its in-depth perfusion by providing artificial capillaries for the supply of nutrients during culture. To maintain the periodicity of the model, micro-channels are arranged according to a hexagonal distribution (Fig. 13.19). For O_2 transport and consumption, a different model was used, because a periodic description does not fit these phenomena straightforwardly. A non-periodic model was built, consisting of superimposed periodic units, to allow for the full height of a 3-mm-thick scaffold. The model accounted for calculation of shear stress on the walls of the reference case and of O_2 partial pressure (pO_2). In calculating pO_2 , different geometries were considered and compared with those of a solely porous model (without channels) of a 3-mm-thick scaffold (Fig. 13.20). The modification of the characteristics concerning the porous medium (diameter of the pores and porosity) mainly accounts for evaluating the influence of the scaffold design on the distribution of shear stresses on pore walls. The results of O_2 transport computations show the effectiveness of the channels in improving O_2 delivery, which has been quantified in a 34% higher pO_2 at the outlet of a 3-mm-thick scaffold. In conclusion, the work by Cantini et al. [136] provides the basis for a CFD-driven design of a dynamic culture system with micro-structured scaffold, acknowledging the increasingly important role of computational methods in defining optimal flow conditions for perfusion bioreactors. These models may result in a valuable tool to appreciate how a controlled variation of the geometry affects the scaffold fluid–dynamic environment; moreover, they provide clues to define the value of basic operation parameters (e.g., perfusion flow rate).

A combined macro/micro-scale computational approach to quantify O_2 transport and flow-mediated shear stress to human chondrocytes cultured in

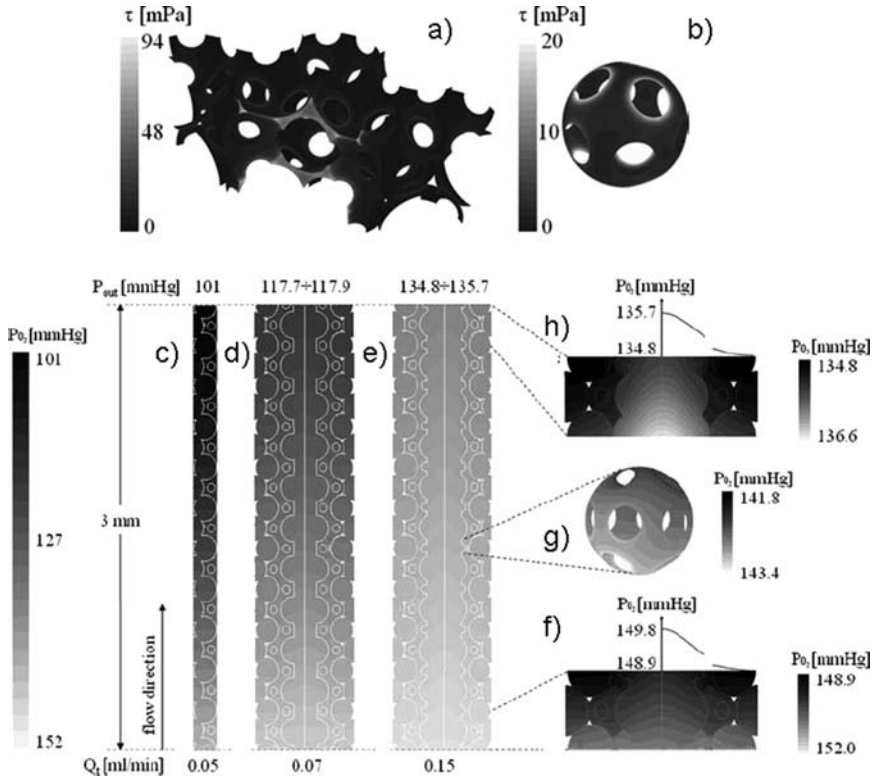


Fig. 13.20. Top: (a) shear stress map on the walls of the reference geometry and (b) detail of a single pore inside the porous bulk of the scaffold; bottom: pO_2 contour maps for (c) a solely porous scaffold model, (d) the channel-provided geometry, and (e) for the reference geometry; pO_2 maps are detailed for (f) the first and (h) the last unit composing the transport mode; (g) the detail of a pore next to the channel is shown [136]

3D scaffolds in a perfusion bioreactor system has been proposed by Cioffi et al. [150]. A macro-scale model has been first developed to assess the influence of the bioreactor design and to identify the proper boundary conditions for the micro-scale model. The micro-scale model, based on a micro-CT reconstruction of a foam scaffold, has been developed to assess the influence of the scaffold micro-architecture on local shear stress and O_2 levels within the scaffold pores. Furthermore, perfusion culture experiments were performed in the modelled cell-scaffold bioreactor system to derive specific O_2 consumption rates for the simulations.

An important application where CFD modelling could play a fundamental role is the engineering of vascular grafts. Generating grafts with sufficient and consistent quality for clinical applications is still, nowadays, highly challenging, particularly since the engineered grafts will generally have to be a few

millimetres to several centimetres in size. Engineering large 3D constructs with conventional cell and tissue culture techniques typically results in a necrotic central region with only a dense layer of cells at the construct periphery, so not suitable for clinical applicability. Engineered tissues with more homogeneous distributions of cells and matrix have been reported when cell–scaffold constructs were cultured in bioreactors that perfuse the culture medium directly through the pores of the scaffold [151,152]. However, the quality of tissues generated under perfusion can be dependent on the flow rate applied [98,123,153] likely the result of different magnitudes of flow-mediated shear stress applied to the cells and/or changes in mass transport phenomena within the construct. For example, increased magnitudes of shear have been correlated with increased mineralised-matrix production by bone marrow cells [100] and with decreased cartilaginous matrix synthesis by chondrocytes [123]. In particular, regarding O₂ supply, quantifying O₂ levels within perfused constructs is necessary to predict whether specific flow rates can supply sufficient O₂ levels to maintain cell viability and function. Additionally, since the differentiation stage of many cell types can be modulated by applying different ranges of O₂ during culture [109], it may be important to identify conditions which can supply cells–scaffold constructs with a defined range of O₂, as a potential strategy to modulate cell phenotype and the quality of the engineered tissue. In such applications, characterising and quantifying the velocity profiles, mechanical stresses and transport of specific biochemical species, in the local micro-environment of the cells, are of primary importance.

Macro-scale computational models have been developed to characterise the flow profiles within bioreactors, to quantify the shear stress on the outer surface of cultured constructs [125,154], and to estimate the O₂ supply inside scaffolds [126,155]. In an attempt to better characterise the local hydrodynamic environment seen by the cells (i.e., within the scaffold pores), CFD models have been developed based on idealised scaffold structures with well-defined pore architectures [103,156,157].

Despite of these results, for a comprehensive analysis of these phenomena, models able to potentially account for the dynamics of cell proliferation are required. In this context, more experimental studies are needed in order to extract the biological basis for the model algorithms: a challenge remains the implementation of a comprehensive model to simulate the transient development of an engineered tissue under perfusion. In order to address the previous issue, the numerical/computational modelling of the mechanics within the engineered constructs can be applied combined together with the previous methods. Faced with the task of understanding a complex system as cellular and biological phenomena implicate, it is indeed very useful to split it into various sub-components, to extract from them the most essential features and to use them to create a simplified but realistic representation of the system, or a ‘model’ of the system. Modelling is widely used in biomechanics, joining both physical models, that use real constructions, and mathematical models, that use conceptual representations [158]. In such a way, a multi-scale model

is obtained, allowing one to observe more closely the behaviour of the system and to make predictions regarding its performance under altered input conditions and different system parameters.

The finite elements methods (FEM) allows for numerically solving problems in which the mechanics of the scaffold and/or the growing tissues is involved. In TE, the computational-based study of biomechanics ranges from the inner workings of a cell to the mechanical properties of soft tissue, bones and materials used in scaffold manufacturing, accounting for the development of improved treatments for a wide array of pathologies as well. FEM, which allows the numerical solution of the equations of continuum mechanics applied to complex systems, is naturally appealing, since it allows the flexible, repeatable and quantitative analysis of multi-factorial scenarios, such as those that characterise physiological function as well for pathological and post-operative conditions [159]. FEM allows entire designs to be constructed, refined and optimised before the design is manufactured. Moreover, the introduction of FEM in scaffold modeling has substantially decreased the time to take products from concept to the realisation. In summary, benefits of FEM include (1) increased accuracy, (2) enhanced design and better insight into critical design parameters, (3) virtual prototyping, (4) fewer hardware prototypes, (5) a faster and less expensive design cycle, (6) increased productivity and (7) increased revenue [160]. According to the 1983 review of Huiskes and Chao [161], the first application of finite element analysis (in orthopaedics applications) was in 1972 [162]. Since then, finite element models have been increasingly used for three main purposes [163,164]: (1) design and pre-clinical analysis of prostheses, (2) to obtain fundamental biomechanical knowledge about musculoskeletal structures and (3) to investigate time-dependent cellular adaptation processes in tissues growing. FEM-based models, and structural ones in particular, are becoming more and more widely adopted, thanks to the continual increase in computing performance. Technically, to create a finite element representation of a structure, it is first conceptually divided into simple parts called elements. Considering a single element of a complex structure, the forces and displacements at the nodes are related by the ‘stiffness matrix’ for the element. Each element has nodes which join with the nodes on adjacent elements to re-create the total structure. The stiffness term for a node is then the addition of all the stiffness terms from the elements joined at that node. In this way, the stiffness matrix for the whole structure can be obtained by re-assembly of the individual elements, called the global stiffness matrix [165]. In developing a FEM-based model, the previous described general procedure for CFD models can be again adopted, which includes (1) the definition of the geometry (physical bounds) of the problem, (2) the partition of the global geometry into discrete elements (the mesh), (3) the definition of the physical problem, (4) the definition of the boundary and of the initial conditions. As for the solution algorithms, discretisation in space produces a system of ordinary differential equations and implicit or semi-implicit methods are generally used to integrate the ordinary differential equations, producing a system

of (usually) nonlinear algebraic equations. Again, in FEM modeling, micro-CT/MRI imaging techniques can be applied in order to extract a realistic representation of the structure to be modeled.

Modeling of the biologic tissue implies a further complexity: tissues are highly heterogeneous materials containing several different solid and liquid constituents. This presents a computational problem because, when a discontinuous material is modelled as a continuum, a length scale over which the material properties are averaged to smooth out the holes needs to be established. It is only at that length scale and no smaller ones that the theoretically predicted results of the continuum models apply [166]. Often, however, results at a length scale smaller than this are needed (e.g., cellular level) and a solution at this level can be impractical from a computational point of view. In the attempt to resolve this facet, the so-called homogenisation approach can be used: the aim of homogenisation is to represent a heterogeneous medium by a homogenous continuum having the same macroscopic behaviour. It can be successfully used to calculate and numerically predict micro-level strains based on macro-level strains calculated in a global finite element model [167, 168]. Turning back to the need for modelling the dynamics of cell proliferation within scaffolds, a computational analysis, based on mechano-regulation algorithms, has been carried out for analysing engineered bone subjected to axial loading, by Byrne et al. [169]. Previous research has shown that bone regeneration during fracture healing and osteochondral defect repair can be simulated using mechano-regulation algorithms based on computing strain and/or fluid flow in the regenerating tissue. Scaffolds for bone TE are subject to many interlinked and often opposing biological and structural requirements. The conflicting nature of the requirements was described by, among others, Karageorgiou et al. [110], who reported that higher porosities result in greater bone ingrowth *in vivo* but that the resulting lowering of mechanical stiffness and strength sets an upper functional limit for porosity. Thus, it would seem that the porosity of the scaffold must lie within a critical range small enough to maintain the mechanical integrity of the scaffold and large enough to provide optimal bioactivity [170]. Another design parameter that can be controlled, at least to some degree, is the rate of dissolution of a bioresorbable scaffold. Resorption will increase the porosity and, as a consequence, will reduce the strength and stiffness that increases the load transferred to the regenerate. Loading on the regenerate has been theorised and experimentally shown to directly influence the cellular and tissue differentiation patterns [171].

A variety of biodegradable scaffolds are used for the regeneration of bone. In particular, calcium phosphate ceramics have been shown to interact strongly and specifically with bone cells [172]. Generally, a scaffold manufactured using these materials, resorb within a few months depending on the site of implantation. It is thought that calcium phosphate biomaterials resorb *in vivo* via chemical dissolution and cell-mediated degradation [173, 174]. Since the Young's modulus, porosity and dissolution rate have been shown to influence the rate and degree at which bone formation occurs within the scaffold,

it should be possible to determine values for these parameters that maximise bone regeneration. Towards this effort, computational analysis holds great promise in enhancing TE.

Numerical simulations of bone growth into scaffolds based on mechano-biological models have been proposed by Sanz-Herrera et al. [175] who used a 2D finite element model to analyse bone growth within a scaffold implanted in the femoral condyle of a rabbit. Their results were corroborated against the experimental results of Pothuau et al. [176]. Another approach proposed by Adachi et al. [177] models bone regeneration in a unit cell of a 3D scaffold microstructure: this study compares the effect of geometrical parameters on bone regeneration while undergoing scaffold degradation. The previous studies use bone remodelling theories to model bone formation into the scaffolds.

Research on the relationship between mechanical forces and tissue phenotype began with Pauwels [178], who proposed that the shear stress and the hydrostatic stress regulated the type of soft tissue formed within a fracture callus. Analyses correlating mechanical stimuli with tissue differentiation were developed further by Carter et al. [179], who proposed that distortional strain history and hydrostatic pressure history regulate tissue differentiation. By quantifying the thresholds between bone, cartilage and fibrous tissue in mechano-regulation, Claes et al. [180] successfully simulated fracture healing. Using a constitutive model of the tissue as poroelastic, Prendergast et al. [17] proposed the stimuli on mesenchymal stem cell differentiation to be the ones causing cell distortion: this mechano-regulation model successfully predicted tissue differentiation observed experimentally around implants [181]. Geris et al. [182] have also used Prendergast model to analyse tissue differentiation inside a bone chamber placed in a rabbit tibia. In that study, mechano-regulation by strain and fluid flow correlated best with experimentally observed tissue formation patterns in the chamber.

However, most of these models used a diffusion equation to describe the proliferation of cells occurring within the regenerating tissue, while a random walk model gives a more general approach, allowing not only for the simultaneous dispersal of several cell populations, but also for explicit modelling of cell proliferation [183–185].

The hypothesis of the study by Byrne et al. [169] was that the extent of bone regeneration in a scaffold can be maximised by appropriate selection of porosity, Young's modulus, and dissolution rate, and that these parameters will be dependent of the magnitude of the local loading. To model the dispersal of the various cell populations in 3D, a 'lattice' is created within each finite element of the granulation tissue (Fig. 13.21). Each lattice point is considered a region of space for both the cell and the extracellular matrix. The number of lattice points is determined by the dimensions of the element. Therefore, the length, height and width of each finite element are divided by the average diameter of a cell (taken here as $25\ \mu\text{m}$) to give the number of rows, columns and the depth of the lattice. Both cell proliferation and cell migration are based on a stochastic process consisting of a sequence of discrete steps of

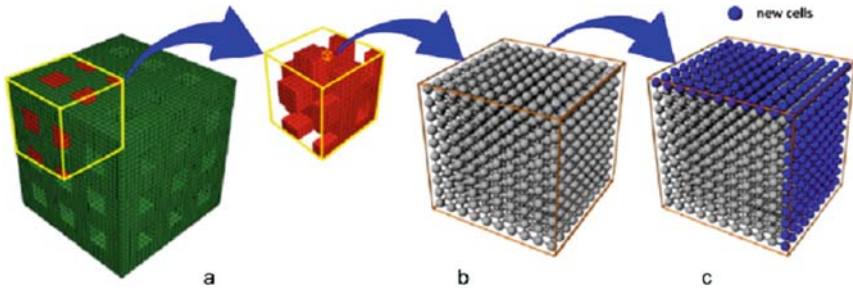


Fig. 13.21. (a) Finite element model of a 50% porous scaffold with regular porosity (green). Only one-eighth needs to be modelled because of symmetry (yellow box). The cavity is initially occupied by granulation tissue (in red). (b) Lattice generated for each granulation element to model cellular activity; (c) increase lattice points to account for dissolution of scaffold material [169]

fixed lengths. To model proliferation, a cell is initially presumed (in 3D) to be surrounded by six possible locations that a daughter cell can occupy. First, a new position is randomly selected from the surrounding locations (including its original position). In turn, one of the remaining neighbouring positions is then chosen for the daughter cell to occupy. In the event that the chosen location is already occupied another position is chosen again at random. This process continues until either the simulation ends or all lattice positions are occupied. To model migration a new position is chosen at random from the surrounding locations (including its current position). Recognising that migration is a more rapid process, a new location for a migrating cell is chosen n times per iteration of the proliferation process. According to the mechano-regulation algorithm for tissue differentiation [17], the tissue differentiation is regulated by magnitudes of the shear strain and of the relative fluid/solid velocity. Shear strain and fluid velocity are calculated from a biphasic poroelastic analysis and, next, differentiation of the mesenchymal cell population is simulated. Depending on the value of the tissue differentiation parameter, a tissue phenotype is predicted for each lattice point throughout the model: high stimulus levels promote the differentiation of mesenchymal cells into fibroblasts, intermediate levels stimulate the differentiation into chondrocytes, low levels of these stimuli promote the differentiation into osteoblasts and very low values promote resorption.

As for scaffold mechanics modelling, a 3D FEM-based model of a regular structured bone scaffold was used, giving geometry similar to printed scaffolds [186]. The scaffolds porosity, mechanical properties, dissolution rate and applied loading conditions change, as a function of the evaluated differentiation of different cellular phenotypes [187–192]. During modelling dissolution, the volume of granulation tissue will increase and the scaffold volume will decrease. The cells within each lattice differentiate based on the stimuli calculated by the mechano-regulation algorithm. As it is likely that several tissues

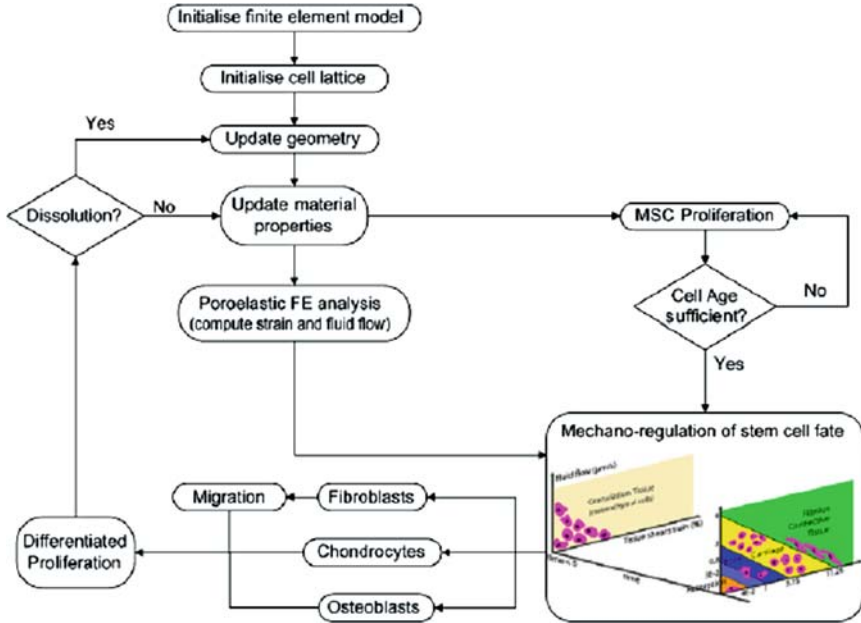


Fig. 13.22. Computational algorithm used to model cellular migration, proliferation and tissue differentiation [169]

can coexist within one finite element, the mechanical properties are calculated using the rule of mixtures, accounting for both the number and phenotype of cells within each element: therefore, the material properties will change gradually towards the phenotype determined by the stimulus. Furthermore, based on Richardson et al. [193] who observed an exponential increase in stiffness in differentiating tissue, a rate equation is used to better describe the evolution of the Young’s modulus of the regenerating tissue [194]. The simulation is performed as described in the diagram in Fig. 13.22.

Starting from an initial condition with mesenchymal stem cells randomly distributed throughout the granulation tissue, the cellular activity within the regenerating tissue is predicted, with phenotype patterns appearing and disappearing during regeneration (Fig. 13.23). The most significant finding of the computational model proposed by Byrne et al. [169] is the possibility to pre-determine optimised morphology and micro-architecture of the scaffold for bone tissue regeneration, based on a realistic simulation of the in situ regeneration. Moreover, concerning modelling of scaffolds for bone tissue regeneration, a computational multi-scale approach has been proposed by Sanz-Herrera et al. [195], elucidating the effect of some scaffold parameters on bone tissue regeneration by means of a mathematically based approach. In particular, factors such as scaffold stiffness, porosity, resorption kinetics, pore size and pre-seeding are analysed in a specific bone tissue application.

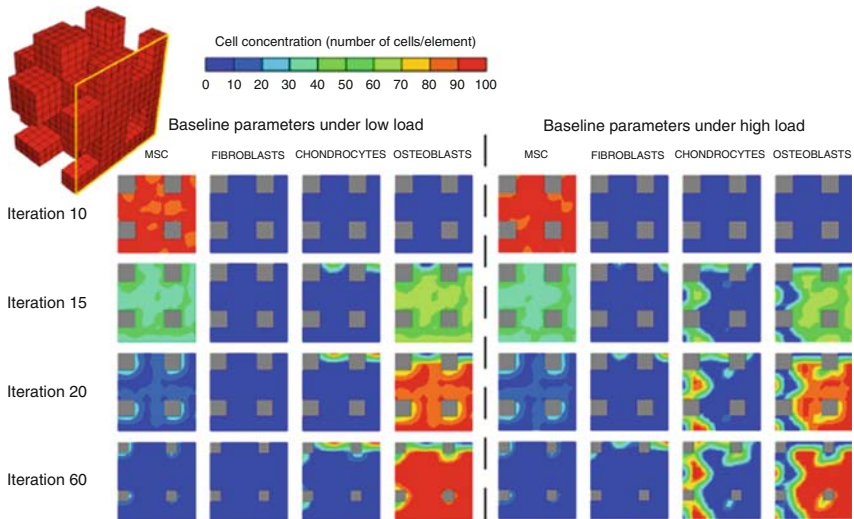


Fig. 13.23. Predicted cell distribution in the granulation tissue: porosity increasing over time due to dissolution of the scaffold biomaterial [169]

The model predicts the *in vivo* rate of bone formation within the scaffold, the scaffold degradation and the interaction between the implanted scaffold and the surrounding bone. Tissue regeneration using scaffolds *in vivo* inherently works on two well-differentiated spatio-temporal scales: the tissue level (macro-scale) and the pore scaffold level (micro-scale). As discussed above, at the microscopic or pore scaffold scale, homogenisation techniques together with CAD tools allow macroscopic mechanical properties to be designed by controlling, in particular, porosity and pore size [63,177,196,197]. At the same scale, simulation of bone regeneration in a unit cell of the scaffold microstructure is conceivable. In the study by Sanz-Herrera et al. [195], the macroscopic mechanical, diffusive and flow properties at the tissue level are derived by means of the asymptotic homogenisation theory [198–200], while at the microscopic scale, bone tissue regeneration at the scaffold micro-surface is simulated using bone growth models based on a bone remodelling theory [201], together with the implementation of degradation models for the scaffold structure. Scaffold degradation and bone growth are mathematically modelled by the voxel finite element method (voxel FEM) [177,202]. With respect to a reference scaffold [198], the effect of changing the following variables (1) Young’s modulus (stiffness) of the bulk biomaterial, (2) resorption kinetics of the bulk biomaterial, (3) overall porosity, (4) mean pore size and (5) the effect of the pre-seeding have been analysed. Two scales of analysis are here considered, that is, the tissue (macroscopic) and pore (microscopic) scales are shown in Fig. 13.24. Performing a multi-scale analysis, the interaction between the two scales has been implemented and, finally, both the models of bone growth

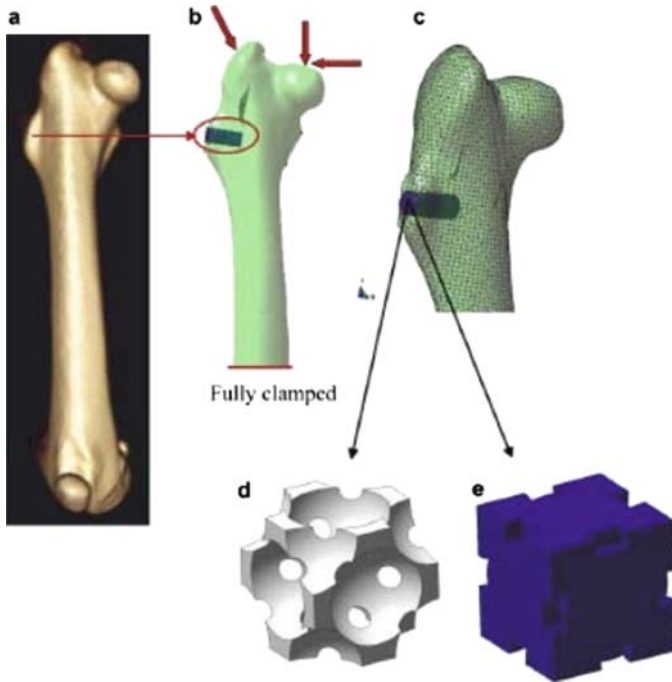


Fig. 13.24. Multi-scale approach for in silico simulations taken from an experimental model: (a) CT of the rabbit femur and (b) CAD macroscopic model (tissue level) of the femur and scaffold implantation, showing boundary loads and prescribed displacements. (c) A detailed FE mesh of that zone. (d) Unit cells of the idealised scaffold microstructure: (d) pore level for the solid scaffold and (e) the fluid domain [195]

and scaffold resorption at the microscopic scale have been introduced. The micro–macro coupling is required since the different ways the bone and scaffold are modelled: the bone organ domain is treated only macroscopically through a fully heterogeneous model, while the scaffold macroscopic properties are directly and explicitly related to its underlying microstructure by using homogenisation techniques [203–205]. Moreover, the process of healing and vasculogenesis is modelled as the diffusion or invasion of cells (considered in the sense of cell populations) and formation of blood vessels within the scaffold core. This is simulated, as a first approach, through Fick’s law in the macroscopic domain. Degradation of the polymeric biomaterial is a mechanism assumed to be driven by bulk erosion due to hydrolysis [206]. Finally, bone growth at the surface of the microstructure of the scaffold occurs as a consequence of bone deposition by osteoblasts. The cascade of events of bone formation onto a bioceramic scaffold surface is thought as follows (adapted from Ohgushi and Caplan [203]):

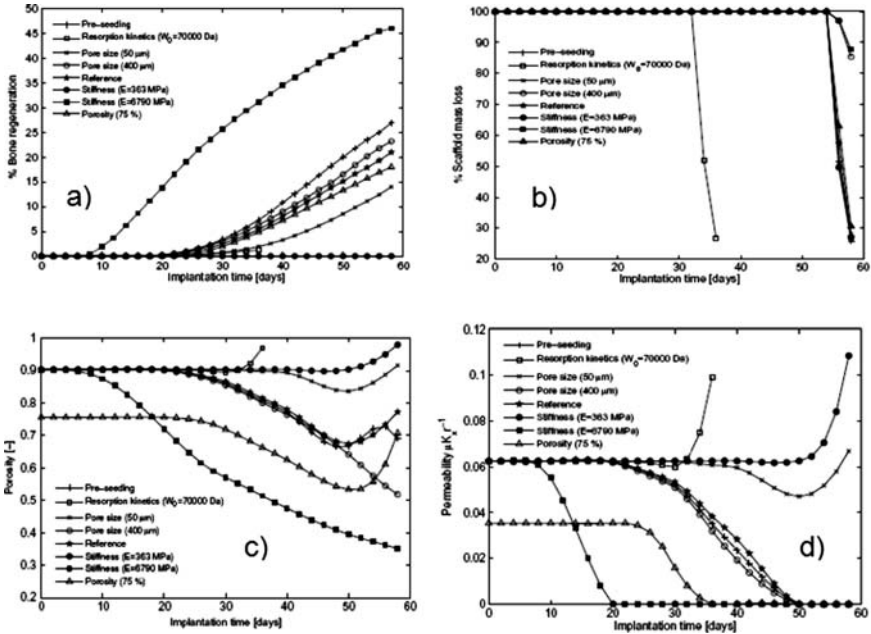


Fig. 13.25. Top: global percentage of (a) bone regeneration and (b) scaffold mass loss for the different analysed cases bottom: (c) unit cell scaffold porosity along implantation time at the mid-point of the scaffold midsection and (d) unit cell scaffold dimensionless permeability Kx (x -direction) with implantation time at the mid-point of the scaffold midsection [195]

- The insertion of a bone substitute produces an immunological response around the scaffold surface and a hematoma is formed. Osteoprogenitor cells then attach to the scaffold surface, and MSCs and osteoblastic cells from the hematoma migrate to scaffold surface;
- These cells attach to the scaffold surface to form a continuous cell layer on the surface of the ceramic within the pores;
- Cell differentiation of these osteoprogenitor cells to form osteoblasts results in a cell layer growth onto the scaffold surface;
- The osteoblasts fabricate immature primary bone onto the surface;
- The primary bone becomes mature and forms fully mineralised bone.

The obtained results are reported in Fig. 13.25. The introduced multi-scale approach allows for modelling almost the most important phenomena involved in bone regeneration using scaffolds, including macroscopic cell migration and mechanics, together with, at the microscopic scale, the determination of the mechanical and migration properties, as well as representing bone formation and scaffold resorption in a microstructure which is statistically representative of the scaffold.

In order to study the possibility of the bioartificial maintenance or restoration of the mechanical function of the intervertebral disc, Yao et al. [205] proposed a FEM-based study of the initial mechanical performance of the disc after implantation of an hybrid engineered construct. The study appears to be extremely valid since the model was developed taking into account the material nonlinearities and to impose different and complex loading conditions. Another significant example of multi-scale modelling, addressed to the realisation of 3D tissue-like constructs for cartilage repair and transplant, is the one proposed by Hussein et al. [206]. The work was aimed at numerically investigating the functioning of a recently designed flow bed bioreactor for cartilage constructs perfusion during *in vitro* growth, in order to optimise the culture protocol. Basically, for junction tissues (i.e., cartilage), the nutrition level and loading stresses are fundamental circumstances for tissue to carry out optimal functioning *in vitro* and *in vivo*. Several bioreactors were built to investigate such effects on the tissue-engineered cartilage–cell morphology. A two-phase flow bed reactor was proposed which showed some substantial advantages over the others, and the flow characteristics within the reactor has been simulated in order to evaluate the effect of fluid mechanics on those conditions both numerically and experimentally. The simulations were made using the multi-phase lattice Boltzmann method (MLBM). The results extracted from this work, as a function of the different parameters set for the numerical simulations, helped to develop a better understanding of the integral phenomena in the bioreactor.

All the previous studies, showing integrated multi-scale together with CFD/FEM-based modelling methodology for the extraction of experimental data, demonstrate the feasibility and the helpfulness of CATE methodologies applied in scaffold design procedures. As a result, cellular metabolism in different perfusion conditions, fluid-induced mechanical stress experienced by the cells during *in vitro* growth, parameters optimising the culture protocol in dedicated bioreactors, the material properties required for a tissue-engineered scaffold, the (pre-determined) morphology and micro-architecture of the scaffold for tissue regeneration, the dynamics of cellular proliferation within scaffolds together with the modification of the topological and structural properties of the scaffold during new tissue formation, are some of the issues that can be analysed by means of these techniques. Most of these tasks are particularly difficult and also time-consuming to be experimentally obtained.

In conclusion, the methodologies described here may be applicable to design and manufacture regularly micro-structured scaffolds in order to produce, by 3D micro-fabrication, engineered tissues or implantable artificial organs [204, 207–210].

13.3 Understanding the Cell and Tissue Mechanics: A Multi-Scale Approach

In the previous sections, the importance of attempting to design and produce scaffold mimicking the native ECM micro-architecture and behaving as a bioinductive substrate for *in vitro* cellular growth and tissue formation and functionalisation has been pointed out. A further relevant topic for the successful realisation of 3D engineered tissue constructs for consistent therapeutic applicability in regenerative medicine is the comprehension of the physiological stimuli experienced *in vivo* by the cells, which is of basic relevance for providing to the cultured cells a substrate and an environment ensuring cell growth and appropriate differentiation to form a new tissue. A detailed analysis of such mechanisms permits, first, to meet the processes involved and responsible for the ‘active’ cell–substrate interactions, which contribute to the global tissue/organ activities and, second, to bioartificially reproduce them.

Since living tissue are composed by various elements, hierarchically assembled and organised, a multi-scale approach results again suitable and extremely helpful in investigating such phenomena.

Due to the nature and function of biological systems, data that pertain to the cellular scale must be explored, since an effective tissue reconstruction can only occur if proper cellular activities are ensured. The major challenge of TE is, indeed, directing the cells to establish the physiological structure and function of the tissue being replaced across different hierarchical scales. As an example of a multi-scale approach in analysing cellular mechanics, let us consider the development of a 3D assembly of myocardial tissue constructs [211]: to engineer the myocardium, biophysical regulation of the cells needs to recapitulate multiple signals present in the native heart. This process implies the analysis of various cellular physiological aspects and mechanisms which must congregate, through different scales (molecular–cellular–tissue), for the new tissue development. The authors assessed that the formation of the typical cardiac cell gap junctions, responsible for the synchronous contraction of the myocardial tissue, can be realised by inducing electric stimulation to the cultured cells reproducing those of the native heart, which allows for concurrent and progressive development of conductive and contractile properties of cardiac constructs cultured *in vitro* (Fig. 13.26). The culture protocol aimed at the functionalisation of the new tissue is temporally divided into two distinguished phases, that can be identified, in a multi-scale point of view, as related to the molecular and micro-cellular level and to the tissue level. Referring to Fig. 13.26, during phase 1 (pre-culture without electrical stimulation), cells accumulate and assemble conductive and contractile proteins lost or disorganised during isolation from heart tissue; during this stage, electrical stimulation has an inhibitory role. During phase 2, constructs are cultured with the application of electrical stimulation. Electrical stimulation, induced at the proper new tissue formation stage, enhances the development of ultra-structural and contractile properties of the individual cells (arrow to the right) and increases

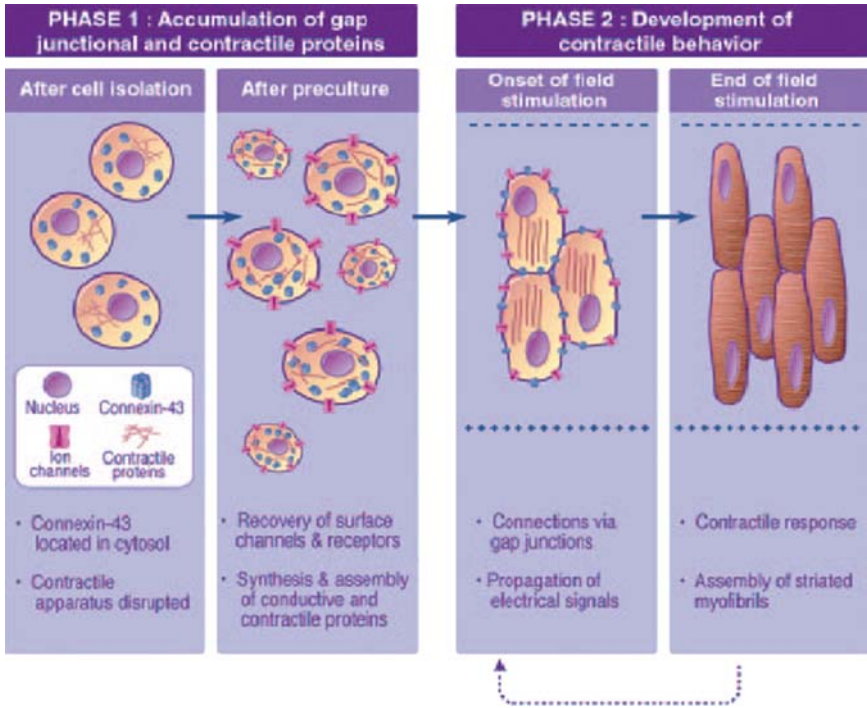


Fig. 13.26. Concurrent and progressive development of conductive and contractile properties of cardiac constructs cultured in vitro [211]

the number of functionally coupled cells engaged in synchronous contractions of the constructs (arrow to the left). We can conclude that, in any TE applications, the proper setup for cultivating cells can be suitably designed only if it is based on a previous understanding of the multi-level physiological phenomena affecting the cellular response to the applied stimuli.

The main goal of cell biomechanics remains the attempt to clarify the relationship between the principal biological functions (such as differentiation, gene expression, migration) and the state of stress and strain of living adherent cells. These parameters depend on the cell’s mechanical behaviour, essentially relying on the behaviour of its cytoskeleton (CSK) [21, 23, 212–224]. Hence, numerous studies aimed at analysing the CSK structure and propose mechanical structural models, providing a better understanding of the phenomena involved, have been developed [225–227]. Based on the previous assumptions, in the present section, a brief summary of the techniques and methodologies successfully applied for the study and the modelling of the biomechanics of living cells is given. Once the physiological processes involved in mechanotransduction, affecting cell growth, migration and gene expression, have been pointed out, in engineering a tissue, modelling methodologies can be adopted in order to pre-determine and establish the cellular response to

the scaffold design and culture protocol. Improved biological and mechanical functionality of tissue-engineered constructs, required for clinical application, can only be achieved by comprehensive multidisciplinary research [228]. Computational and numerical modelling is, to date, a framework for obtaining an integrated understanding of multiple key processes. Such an integrated perspective of these key aspects will be critical to open up new directions in TE research, as significant progress can be made by combining existing computational and experimental methods.

First, mechanical cells–substrate interactions must be analysed in designing scaffolds. Mechanical interactions between different components, that is, cells, water and scaffold material during tissue growth can determine whether cells will form aggregates to develop in new tissue formation or disperse throughout the scaffold [229]. A main organising principle is that cells generate internal stresses in their cytoskeleton, which in 2D *in vitro* culture are transferred to the substrate via focal adhesion complexes: while on a too flexible 2D substrate cells adopt a circular shape, on a rigid surface they are able to spread out. Furthermore, 3D tissue constructs can deform substantially under the influence of cellular force generation, which emphasises the importance of the stiffness of the scaffold material as a design parameter [230]

FEM modelling demonstrated that the regions in which proliferation was concentrated corresponded to mechanical stress patterns generated within the cell layer, which were validated later by measurements using a micromechanical force sensor array. An improved understanding of these mechanisms can therefore be exploited to direct cell migration and promote ingrowth by designing the scaffold architecture with respect to fibre size, orientation and surface topology [231–233]. In 2008, Nair [234] developed a multi-scale 3D model for characterising mechanical properties of cell encapsulated in scaffolds composed by alginate discs. A 3D multi-scale FEM-based model has been developed to analyse the stresses and deformations of the cell when the bio-fabricated tissue construct is subjected to macro-level compressive loads. Specifically, this methodology characterises the macro-scale structural behaviours of the scaffold, quantifies 3D stresses and deformations in the micro-scale level comprising multiple cells, and studies cell damage at a cellular level wherein individual cell components like the nucleus and cytoplasm are modelled. First, the scaffold (macro-scale model) mechanical response to mechanical stimuli has been estimated (Fig. 13.27a). Once established the mechanical response of the scaffold material, a 3D multi-cell model has been incorporated, to determine the micro-stress and deformation fields and the effects of geometrical heterogeneity in the cell's microenvironment. The distributions of the cells within the microstructures were determined by a uniform random distribution probability function. The cells were modelled as single inclusions: the three different material phases are described by the three different material models, namely, the Ogden hyperelastic polynomial for capturing the inherent non-linearity of alginate, and the Neo-hookean polynomial function for describing the cytoplasm and the nucleus [235–237]. Mapped

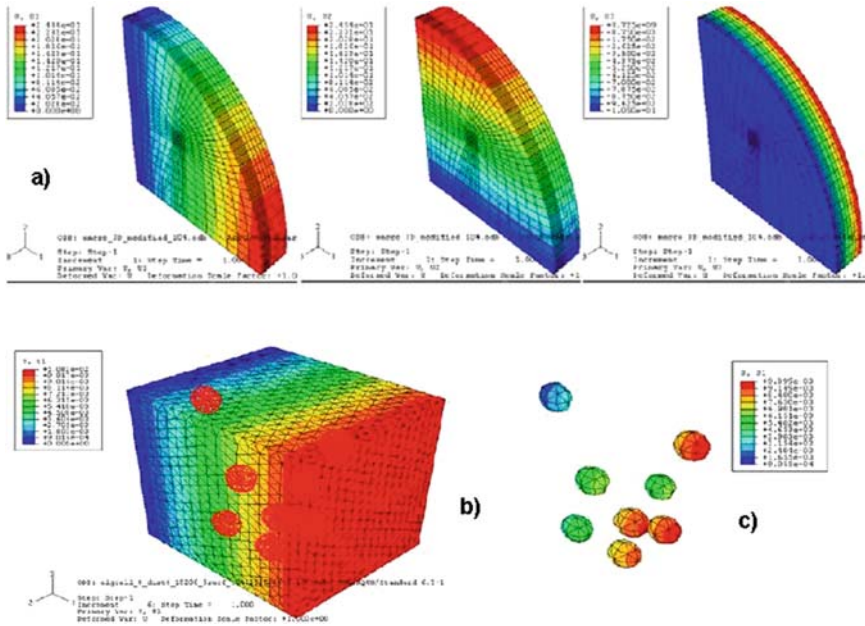


Fig. 13.27. (a) Macro-scale model displacement contour in the three principal directions, (b) displacement contour in multi-cellular model, scaffold and cells and (c) cells after the volume of alginate has been subtracted [234]

displacements from the macro-scale model were applied onto the multi-cellular volume. The 3D scenario allowed to simulate boundary and loading conditions that realistically mimic a natural environment within the body (Fig. 13.27b, c). To determine whether the cell apoptosis occurs, a strain energy density-based damage criterion has been formulated which, together with a stochastic approach ‘determining’ the cell viability within the 3D tissue construct under a prescribed mechanical load, allowed for pre-determining a cell viability index, predicting the cell viability (Fig. 13.28). Specifically, this method characterises the macro-scale structural behaviour of the tissue scaffold, analyses and quantifies the 3D stresses and deformations in the micro-scale level comprising multiple cells, and aids in determining cell damage by analysing a single cell model wherein individual cell components like the nucleus and cytoplasm are incorporated. This modelling methodology provides an effective means to quantify the stresses and strains at the cell’s micro-environment when tissue constructs are subjected to macro-field deformations providing a significant step towards a better understanding of cell mechanisms in TE applications. The computational modelling approach sets up a framework for the analysis of cell damage when tissue constructs are subjected to mechanical stimuli in different environments.

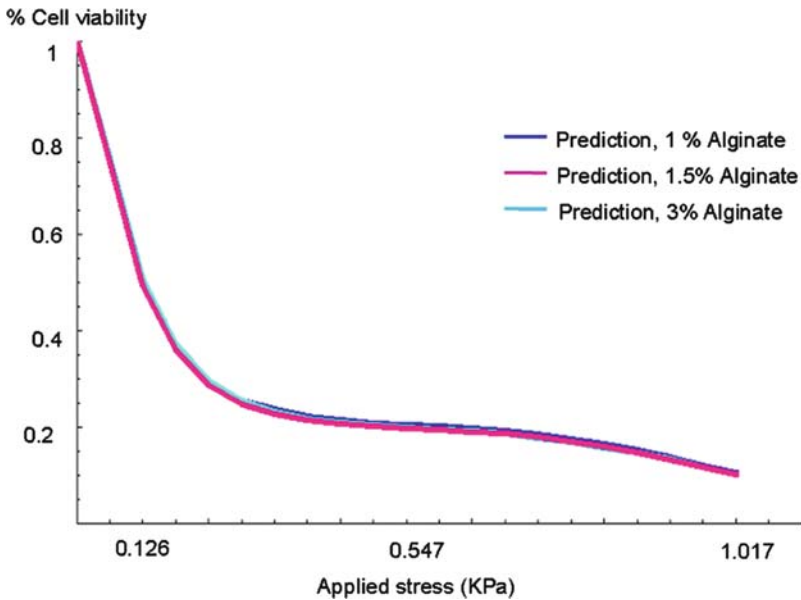


Fig. 13.28. Cell viability estimated as a function of the alginate concentration [1%, 3%, and 5% (w/v)] at different mechanical loading conditions [234]

13.4 Experimental Techniques for Scaffolds Characterisations

Several techniques have been developed in the attempt of evaluating all the most significant characteristics affecting the RP-based manufacture of a tissue-engineered scaffold [76]. Apart of micro-CT, they include (1) theoretical calculation, (2) scanning electron microscopy (SEM), (3) mercury and flow porosimetry, (4) gas pycnometry and (5) gas adsorption. In Table 13.4, a summary of the most significant architectural and structural properties of scaffolds that can be derived (denoted by +) by the individual techniques and those that cannot be (denoted by -) is furnished.

Briefly, theoretical methods can be applied in evaluating scaffold properties but most of these approaches are only capable of estimating scaffold porosity. The two main remarkable theoretical approaches are the ‘unit cube analysis’ and the ‘mass technique’ [76]. The first approach is commonly adopted for regular honeycombed scaffolds. In the unit cube method, the volume sum of the scaffold material and the associated pore spaces is calculated by taking linear measurements of the scaffold cube. The volume of the scaffold material is calculated from a known deposition pattern, so this method fails if applied for calculation of scaffolds with irregular geometry/architecture, as for the case of ones manufactured by means of extrusion techniques, since it is based on the assumption that the scaffold struts are maintained uniform during the

Table 13.4. A summary of the properties that can be derived (denoted by +) by the individual techniques and those that cannot be (denoted by -) [76]

Architectural and structural properties	Theoretical method	SEM	Mercury porosimetry	Gas pycnometry	Gas adsorption	Flow porosimetry	Micro-CT
Porosity	+	Qualitative	+	+	+	-	+
Surface area/volume	-	-	+	-	-	-	+
Pore size	-	+	+	-	+	+	+
Strut/wall thickness	-	+	-	-	-	-	+
Anisotropy	-	Qualitative	-	-	-	-	+
Cross-section area	-	Qualitative	-	-	-	-	+
Permeability	-	-	+	-	-	-	+
Other parameters that can be derived			<ul style="list-style-type: none"> • Pore volume vs. pore size • Pore cavity to pore throat ratio • Pore cavity sizes vs. pore throat sizes • Cumulative surface area vs. pore size • Pore tortuosity 		<ul style="list-style-type: none"> • Pore shapes 	<ul style="list-style-type: none"> • Effects of compressive stresses can be studied 	<ul style="list-style-type: none"> • Pore shape and morphology
					<ul style="list-style-type: none"> • Cumulative pore size distributions • Pore volume and areas of mesopores and micropores • Pore size distributions 		
			<ul style="list-style-type: none"> • Sample compressibility 				<ul style="list-style-type: none"> • Pore size distributions

(continued)

Table 13.4. (Continued)

Architectural and structural properties	Theoretical method	SEM	Mercury porosimetry	Gas pycnometry	Gas adsorption	Flow porosimetry	Micro-CT
Other drawbacks	<ul style="list-style-type: none"> Not precise due to the inaccurate measurements of the linear dimensions of the scaffold cube 	<ul style="list-style-type: none"> Physical sectioning required 	<ul style="list-style-type: none"> Does not account for closed pores 	<ul style="list-style-type: none"> Not precise due to inaccurate measurements of the linear dimensions of the scaffold cube Does not account for closed pores 	<ul style="list-style-type: none"> Erroneous measurements samples are encountered for low surface area the calculation Can be time consuming Does not account for closed pores 	<ul style="list-style-type: none"> Porosity and cross-sectional area values are prerequisites for of the surface area through pores Does not account for closed and blind pores 	<ul style="list-style-type: none"> Thresholding difficulty Beam hardening
Regular geometry scaffolds	Suitable	Qualitative	Suitable	Suitable	Suitable	Suitable	Suitable
Foams	Suitable	Qualitative	Suitable	Suitable	Suitable	Suitable	Suitable
Textile scaffolds	Not suitable	Qualitative	Not suitable	Not suitable	Suitable	Suitable	Suitable
Nanofibres	Not suitable	Qualitative	Not suitable	Not suitable	Suitable	Suitable	Suitable

manufacture. In the mass technique, the volume of the scaffold material is derived by dividing the mass of the scaffold cube by its material density. The ‘apparent’ scaffold cube volume is calculated from its linear measurements and scaffold porosity is derived. This method is applicable for scaffolds with controlled and uncontrolled geometries, but it is dependent on the accuracy of the linear measurements of the cube. As for SEM analysis, it complements these theoretical calculations of porosity and it allows direct measurements of pore size and strut/wall thickness [238, 239]. A visual estimation of interconnectivity, cross-section area and anisotropy can also be achieved. However, to examine the scaffold interior, physical sectioning is needed and this would introduce unnecessary compromising of the struts integrity. A qualitative analysis of foams, textiles and nanofibre meshes scaffolds can be provided by SEM imaging.

Mercury porosimetry is a well-known and established method which is often used to study porous materials. This technique employs mercury which is a non-wetting liquid that does not intrude into pore spaces except under sufficient pressure. In 1921, Washburn related this pressure to the size of the intruded pore space which is shown in the Washburn equation [240]. In this technique, the sample is introduced into a penetrometer and is subjected to gas evacuation while allowing mercury in flow. Ambient pressure is applied and the sample is enveloped by mercury. From the mercury volume measurements, the bulk volume of the sample is derived. Bulk volume is equivalent to apparent volume and it consists of scaffold material and pore spaces. Maximum pressure is attained through incremental steps so as to promote mercury intrusion. At maximum pressure, the total volume of the intruded mercury is measured, and this allows the derivation of total open pore volume and porosity. Using the Washburn equation, pore sizes and pore volume distribution by pore size are calculated [241]. Pore sizes smaller than $0.0018\ \mu\text{m}$ are not intruded with mercury and this is a source of error for porosity calculations. Furthermore, mercury porosimetry does not account for closed pores as mercury does not intrude into closed pores [242]. Because of the high pressure applied by the porosimeter (until 400 MPa), mercury porosimetry is not suitable for fragile compressible scaffolds such as flexible foams (with porosities higher than 90%), textiles and nanofibre sheets. Other concerns would include the fact that the Washburn equation assumes perfect cylindrical pores, which is not the case in reality.

Gas pycnometry is capable of measuring scaffold porosity without the need to measure the scaffold material volume [51, 59], employing a gas displacement method. This technique operates by detecting the change in pressure when the scaffold specimen displaces gas: a specimen of unknown volume is placed in the sample chamber of known volume; upon sealing, the chamber pressure is measured. A reference chamber, with known volume and pressure, is separated from the sample chamber by a closed valve. Both systems are allowed to equilibrate to a pressure when the valve is opened. Using gas law, the unknown volume of the specimen is obtained [243]. Once the scaffold material

volume is determined, using the cube approach allows for the determination of the scaffold porosity. There are typically two sources of error in this volumetric measurement; first, it does not account for the closed pores which the gas molecules cannot enter. Second, again, the reliability of this method is hampered by the difficulty in taking accurate linear dimensions of the unit scaffold cube. Furthermore, gas pycnometry does not inform the user of the other scaffold parameters such as surface area to volume ratio, interconnectivity, permeability and pore size; thus, it has to be coupled with other techniques (such as micro-CT) in order to achieve a more conclusive assessment.

In solids, electrical forces of attraction hold down the atoms in their fixed positions. As the outermost atoms have lesser neighbours than those beneath them, an imbalance of attractive forces occurs. To counter this, surface atoms attract surrounding gas molecules via van der Waals and electrical forces. This physical phenomenon is known as gas adsorption and it can be applied in the study of porous materials [244–246]. The gas adsorption procedure begins with the placement of the specimen in an evacuated chamber where a small amount of adsorbate gas (nitrogen, benzene vapor, argon and krypton) is introduced. Adsorption isotherms are derived from the pressure and volume measurements of the chamber. First, adsorbate molecules form an initial thin layer on the available surfaces: at this stage, the surface area can be calculated using Brunauer, Emmett and Teller (BET) theory [247] and the Langmuir model [248–250]. As gas adsorption continues, multi adsorption layers form and capillary condensation occurs: at this stage, pore sizes can be derived using the Barnett, Joyner and Halenda (BJH) method. When the pores are completely saturated with adsorbate molecules, the total pore volume of the scaffold is calculated. Gas adsorption setups are capable of assessing scaffolds with pore sizes ranging from 0.35 to 400 nm or 3.5 to 2,000 nm. Setups capable of small pore size measurements can be used to evaluate the architectural parameters of nanofibres, dense foams and textiles. However, gas adsorption is not a suitable approach in studying scaffolds with low specific surface area, as erroneous surface area measurements would be encountered for values lesser than $0.01 \text{ m}^2\text{g}^{-1}$. The gas adsorption analysis does not measure scaffold material volume; hence, it has to be complemented with the mass technique so that the surface area to volume ratio can be derived. Moreover, this technique does not evaluate other important architectural parameters such as interconnectivity, strut thickness, anisotropy, permeability and cross-sectional area. Closed pores do not allow the entry of adsorbate molecules; thus, gas adsorption fails to account for the presence of these pores [76].

Flow porosimetry is a non-destructive approach that has been used to measure the pore sizes of porous materials. Recently, these techniques have been also applied in biomedical applications [76]. In this technique, a fully wetted scaffold sample is sealed in a chamber and gas is allowed to pass through it. At the bubble point, there is sufficient gas pressure to overcome the fluid capillary action in the largest pore. The pressure is increased while the flow rate is measured till all the scaffold pores are empty and dry. Pore size distributions

and mean pore size can be derived once the flow rate and the applied pressure are known. Flow porosimetry is capable of measuring pore sizes within the range of 0.013–500 μm . Furthermore, compressive stresses can be applied onto the sample so as to study the effects of compression on pore size and distribution. Even though flow porosimetry is a relatively user-friendly technique, it is unable to determine several important scaffold properties such as porosity, interconnectivity, strut thickness, anisotropy, permeability and cross-sectional area. Despite this, flow porosimetry can be coupled with other scaffold evaluation techniques so as to determine the porosity and cross-sectional area of the scaffold. Sometimes, a combination of the previous techniques is required to achieve an in-depth study of the scaffold properties.

Moreover, great attention should be spent in investigating micro-structural and mechanical properties of scaffold by means of mechanical experimental tests since, as mentioned above, the scaffolds act as a physical support structure and a regulator of the cellular response. Mechanical characterisation of scaffolds is a useful tool to validate the mechanical feature of TE scaffolds. Typically, parameters that influence the global mechanical behaviour of a scaffold are (1) hydration level, (2) pore size, (3) cross-link density and (4) relative density, so both the scaffold micro-structure and mechanical properties significantly influence both the cell response and the scaffold in situ bioactivity [251–262]. In order to mechanically characterise a scaffold, tensile and compression tests should be performed [263]. These tests characterise the macroscopic behaviour of the materials constituting the scaffold under crushing loads: the specimen is stretched/compressed and the deformation at various loads is recorded. The tensile/compressive stress and strain are calculated and plotted as stress–strain diagram. As an example of a compression mechanical tests, in Fig. 13.29 the characteristics stress–strain curve of a collagen-glycosaminoglycal (CG) scaffold is shown [263]. Parameters indicative of the mechanical properties of the scaffold have been extracted as the linear elastic modulus (E^*), the elastic collapse stress and strain (σ^*_{el} , ε^*_{el}) and the collapse plateau modulus ($\Delta\sigma/\Delta\varepsilon$). Distinct linear elastic, collapse plateau and densification regimes were observed, as characteristics of low-density, open cell foams [263]. The E^* was determined via linear regression of the initial linear regime of the stress–strain curve, the σ^*_{el} , ε^*_{el} were determined from the intersection of the E^* and $\Delta\sigma/\Delta\varepsilon$ regression curves, and $\Delta\sigma/\Delta\varepsilon$ was determined via linear regression of the linear region following the ‘knee’ corresponding to the strut bending–buckling transition (Fig. 13.29b). A detailed mechanical characterisation, eventually coupled with modelling and simulation techniques, represents a pathway for standardising and defining the extra-cellular environment of the cells within the engineered constructs. As for the micro- and nano-characterisation of the mechanical properties of the engineered struts, scanning probe microscope (SPM) methods can be applied: SPM creates images using a physical probe that scans the specimen: an image of the surface is obtained by mechanically moving the probe in a raster scan

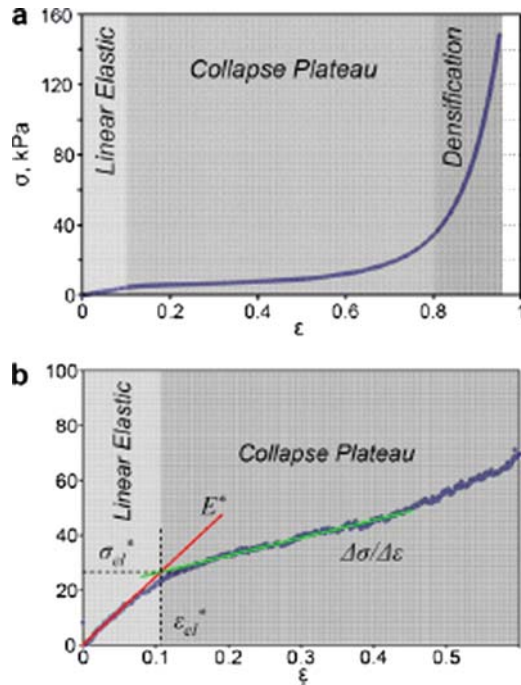


Fig. 13.29. Characteristic stress–strain curves observed for under compression tests for (a) the dry CG scaffold and (b) for a subregion of a hydrated CG scaffold. Distinct linear elastic, collapse plateau and densification regimes can be observed for the CG scaffolds regardless of mean pore size, relative density, cross-linking density, loading direction or level of hydration. In (b) linear regressions of the linear elastic and collapse plateau regimes are used to calculate the linear elastic modulus (E^*), the elastic collapse stress and strain (σ^*_{el} , ϵ^*_{el}), and the collapse plateau modulus ($\Delta\sigma/\Delta\epsilon$) [263]

of the specimen, line by line, and recording the probe–surface interaction as a function of position.

By using specific probes, atomic resolution is reachable applying SPM techniques, which owe this largely to the ability of piezoelectric actuators to execute motions with a precision and accuracy at the atomic level. In particular, in nanotechnology applications, the atomic force microscope (AFM) or scanning force microscope (SFM) is a very high-resolution type of SPM, with demonstrated resolution of fractions of a nanometre, more than 1,000 times better than the optical diffraction limit. The AFM consists of a micro-scale cantilever with a sharp tip (probe) at its end with a radius of curvature in the order of nanometres that is used to scan the specimen surface [264]. When the tip is brought into proximity of a sample surface, forces (mechanical contact, van der Waals interactions, chemical bonding, electrostatic or magnetic forces) between the tip and the sample lead to a deflection of the cantilever according

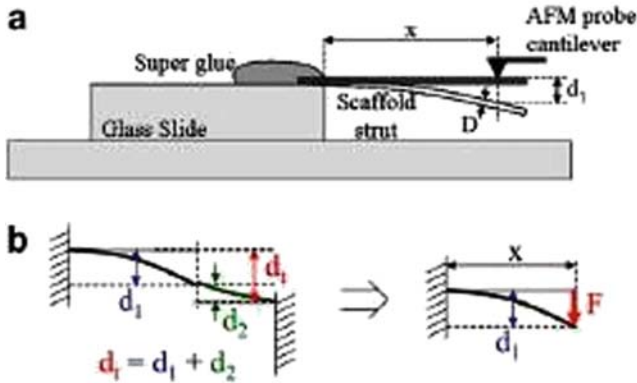


Fig. 13.30. (a) Experimental arrangement to perform bending tests on individual CG scaffold struts, (b) simplified beam bending system utilised to calculate CG scaffold strut modulus (E_s) [263]

to Hooke's law [264]. Typically, the deflection is measured using a laser spot reflected from the top surface of the cantilever into an array of photodiodes, or, alternatively, using optical interferometry, capacitive sensing or piezoresistive cantilevers, acting as a strain gauge. Using AFM techniques, the bending stiffness modulus (E_s) of individual struts of a scaffold can be measured, as schematically shown in Fig. 13.30 [263]. Experimental results are expressed in terms of the displacement of the AFM cantilever base (d_t) and the relative deflection of the free end (d_2) (Fig. 13.30b). Moreover, nanoindentation (NI) has proven to be a powerful technique for the measurement of mechanical properties in diverse biomaterials ranging from mineralised tissues to soft tissues. NI, also known as instrumented indentation testing (IIT), depth-sensing indentation, continuous-recording indentation, ultra-low-load indentation, is a relatively new form of mechanical testing that significantly expands on the capabilities of traditional hardness testing. Developed largely over the past two decades, NI employs high-resolution instrumentation to continuously control and monitor the loads and displacements of an indenter as it is driven into and withdrawn from a material. Depending on the details of the specific testing system, loads as small as 1 nN can be applied, and displacements of 0.1 nm (1 Å) can be measured. Properties such as hardness and reduced modulus are calculated from the unloading curves using well-established equations based on elastic contact theory [265]. This technique bridges the gap between (AFM) and macro-scale mechanical testing. NI improves upon the spatial, force and displacement resolutions of traditional techniques to provide a powerful tool for the characterisation of tissues and other biomaterials with submicron resolution. NI can be used to measure local material properties in small, thin and heterogeneous samples. NI is also useful for measuring mechanical properties of micro-structural features within bulk samples, characterising the properties of individual constituents within composite or heterogeneous

samples, or mapping mechanical properties across a sample surface. As a result, NI is emerging as a valuable mechanical testing technique for biomaterials. NI can serve as a complementary characterisation tool to other techniques that assess composition or structure with high spatial resolution, such as Raman spectroscopy, magnetic resonance imaging, micro-CT, histology or infrared spectroscopy. In this regard, NI has played a pivotal role in defining structure–property relationships for tissues and their constituents. In commercial nanoindenters, displacement is typically monitored by capacitance or inductance, while force actuation is provided through electrostatic force generation, magnetic coils or expansion of a piezoelectric element. A large portion of the current biomaterials nanoindentation literature focuses on the measurement of material properties such as Young’s modulus or elastic modulus (E) and indentation hardness (H) from indentation data. Doerner and Nix [266] published the first comprehensive experimental and analytical approach to a generalised form of NI analysis from load–displacement data for non-rigid indenters of all geometries. Oliver and Pharr [267] further generalised their approach, resulting in the widely used compliance method for indentation analysis. This experimental technique has been applied to a type component that can be used to realise a scaffold by Strasser et al. [268] who have determined the elastic properties of single collagen type I fibrils with the use of AFM used as a nanoindenter. Measurements were done on the outer shell and in the core of the fibril. These findings will open in future a new investigation area (molecular-scale) for the design of customised materials applied in bioactive scaffold manufacturing.

In conclusion, we have presented a wide, even if non–exhaustive, review of the state-of-art of experimental and modelling techniques which are currently applied for evaluating both mechanical and morphological properties of bioartificial scaffolds and matrices at different scales, aiming at furnishing to the reader a comprehensive description of biomimetic scaffolds modelling and design methodologies.

References

1. P. Akhyari, H. Kamiya, A. Haverich, M. Karck, A. Lichtenberg, Eur. J. Cardiothorac. Surg. **34**, 229–241 (2008)
2. D.W. Hutmacher, Biomaterials **21**, 2529–2543 (2000)
3. T. Nishida, K. Yasumoto, T. Otori, J. Desaki, Invest. Ophthalmol. Vis. Sci. **29**(12), 1887–1890 (1988)
4. T. Takezawa, Biomaterials **24**, 2267–2275 (2003)
5. D.W. Hutmacher, A. Kirsch, K.L. Ackermann, M.B. Huerzeler, in *Biological matrices and tissue reconstruction*, ed. by G.B. Stark, R. Horch, E. Tangos. (Springer, Heidelberg, Germany, 1998), pp. 197–206
6. D.W. Hutmacher, J. Biomater. Sci. Polymer Ed. **12**, 107–124 (2001)
7. C. Gomez, A Unit Cell Based Multi-scale Modeling and Design Approach for Tissue Engineered Scaffolds, Ph.D Thesis, Drexel University 2007

8. A.B. Saim, Y. Cao, Y. Weng, C.N. Chang, M.A. Vacanti, C.A. Vacanti, R.D. Eavey, *Laryngoscope* **110**, 1694–1697 (2000)
9. A. Atala, S.B. Bauer, S. Soker, J.J. Yoo, A.B. Retik, *Lancet* **367**, 1215–1216 (2006)
10. F.J. O'Brien, B.A. Harley, M.A. Waller, I.V. Yannas, L.J. Gibson, P.J. Prendergast, *Technol Health Care* **15**(1), 3–17 (2007)
11. M. Knothe-Tate, U. Knothe, *J.Biomech.* **33**, 247–254 (2000)
12. P.A. Netti, D.A. Berk, M.A. Swartz, A.J. Grodzinsky, R.K. Jain, *Can. Res.* **60**, 2497–2503 (2000)
13. C.A. Znati, M. Rosenstein, T.D. McKee, E. Brown, D. Turner, W.D. Bloomer, S. Watkins, R.K. Jain, Y. Boucher, *Clin. Can. Res.* **9**, 5508–5513 (2003)
14. J.M. Mansour, V.C. Mow, *J. Bone Joint Surg.* **58A**(4), 509–516 (1976)
15. I. Owan, D.B. Burr, C.H. Turner, J. Qiu, Y. Tu, J.E. Onyia, *Am. J. Physiol.* **273**(3), C810–C815 (1997)
16. M.V. Hillsley, J.A. Frangos, *Biotech. Bioeng.* **43**, 573–581 (1994)
17. P.J. Prendergast, R. Huiskes, K. Søballe, *J. Biomech.* **30**(6), 539–548 (1997)
18. C.M. Agrawal, J.S. McKinney, D. Lanctot, K.A. Athanasiou, *Biomater* **21**, 2443–2452 (2000)
19. L.I. Shihong, J.R.D. Wijn, L.I. Jiaping, P. Layrolle, K. De Groot, *Tiss. Eng.* **9**(3), 535–548 (2003)
20. J.R. Levick, *Quar. J. Exper. Physiol.* **72**, 409–438 (1987)
21. M.J. Dalby, *Med. Eng. Phys.* **279**(9), 730–742 (2007)
22. A. Katsumi, A.W. Orr, E. Tzima, M.A. Schwartz, *J Biol Chem* **279**(13), 12001–12004 (2004)
23. S. Huang, D.E. Ingber, *Nat Cell Biol* **1**(5), E131–8 (1999)
24. P.A. Anderson, A. Manring, J.R. Sommer, E.A. Johnson, *J Mol Cell Cardiol* **8**(2), 123–43 (1976)
25. G. Macchiarelli, O. Ohtani, *Heart* **86**(4), 416 (2001)
26. W. Sun, P. Lal, *Comput. Methods Programs Biomed.* **67**, 85–103 (2002)
27. N.J. Mankovich, D.R. Robertson, A.M. Cheeseman, *J. Digit Imaging* **3**(2), 69–80 (1990)
28. J.P. Hornak, *The Basics of MRI*, (1996) available at: <http://www.cis.rit.edu/htbooks/mri/>
29. H. Libouban et al., *J. Histotechnol.* **24**, 81–86 (2001)
30. N.E. Lane, W. Yao, J.H. Kinney, G. Modin, M. Balooch, T.J. Wronski, *J. Bone Miner Res* **18**(12), 2105–2115 (2003)
31. M.A. Marko, D. Leith, D. Parsons, *J. Electron Microsc. Tech.* **9**(4), 395–411 (1988)
32. M. Viceconti, C. Zannoni, L. Pierotti, *Comput. Methods Program Biomed.* **56**(3), 211–220 (1998)
33. M. Viceconti, M. Casali, B. Massari, L. Cristofolini, S. Bassini, A. Toni, *J. Biomech.* **29**(9), 1241 (1996)
34. M. Viceconti, C. Zannoni, D. Testi, A. Capello, *Comput. Methods Program Biomed.* **59**, 159–166 (1999)
35. F. Ko, C.T. Laurencin, M.A. Attawia, Fiber architecture for tissue engineering, MRS Meeting, Boston, December 1–5, 1997
36. F. Ko, C.T. Laurencin, M.D. Borden, M.A. Attawia, The role of fiber architecture in tissue engineering, *Proceeding of Composites: Design for Performance*, Lake Louise, Canada, October 12–17 (1997), p. 195

37. F. Ko, C.T. Laurencin, M.D. Borden, D. Reneker, *The Dynamics of Cell–Fiber Interaction*, Society of Biomaterials, (1998)
38. C.T. Laurencin, F. Ko, M.A. Attawia, M.D. Borden, *Cells Mater.* **8** 175–181 (1998)
39. C.T. Laurencin, F. Ko, M.D. Borden, J.A. Cooper, W.J. Li, M.A. Attawia, Fiber-based tissue engineered scaffolds for musculoskeletal applications: in vitro cellular response, MRS Symposium Proceedings, Pittsburgh, (1999)
40. A. Linder, H.P. Wolf, I.M. Watzke, Transfer of the habitual occlusion to 3-D reconstructed skull (stereolithography) with face bow and its application, Proceedings of the Fifth International Congress on Preprosthetic Surgery, Hofburg-Vienna, Austria, April 15–18, 1993
41. N. Mankovich, D. Curtis, T. Kugawa, *J. Prosthet. Dent.* **55**, 606–609 (1985)
42. N. Mankovich, A. Cheeseman, N. Stoker, *J. Digit. Imaging* **3**, 200–203 (1990)
43. N. Mankovich, A. Yue, *Solid Models for CT/MR image display: accuracy and utility in surgical planning*, in: *Medical Imaging V: Image Capture, Formatting and Display*, SPIE, San Jose, CA, (1991). pp. 2–8
44. L. Klimek, H.M. Klein, R. Mo, in *Computer Integrated Surgery*, ed. by R.H. Taylor, (1996), pp. 663–671
45. P. Quadrani, A. Pasini, M. Mattioli-Belmonte, C. Zannoni, A. Tampieri, E. Landi, F. Giantomassi, D. Natali, F. Casali, G. Biagini, A. Tomei-Minardi, *Med. Biol. Eng. Comput.* **43**, 196–199 (2005)
46. C.K. Chua, K.F. Leong, C.S. Lim, *Rapid Prototyping: Principles and Applications in Manufacturing*, (World Scientific, Singapore)
47. M.H. Too, K.F. Leong, C.K. Chua, C.M. Cheah, S.L. Ho, Feasibility of tissue engineering scaffolds fabrication using fused deposition modelling, Proceedings of the Seventh Australian and New Zealand Conference on Intelligent Information Systems (ANZIIS), (2001), pp. 433–8
48. M.H. Too, K.F. Leong, C.K. Chua, *Int. J. Adv. Manuf. Technol.* **19**, 217–23 (2002)
49. D.T. Pham, R.S. Gault, *Int. J. Mach. Tools Manuf.* **38**, 1257 (1998)
50. K.U. Koch, B. Biesinger, C. Arnholz, V. Jansson, *Creating of bio-compatible, high stress resistant and resorbable implants using multiphase jet solidification technology*, Rapid News Publication, 209–214 (1992)
51. C.X.F. Lam, X.M. Mo, S.H. Teoh, D.W. Hutmacher, *Mater. Sci. Eng. C* **20**, 49–56 (2002)
52. L. Lu, A.G. Mikos, *MRS Bull.* **11**, 28–32 (1996)
53. R.C. Thomson, M.J. Yaszemski, A.G. Mikos, in *Principles of tissue engineering*, ed. by R.P. Lanza, R. Langer, W.L. Chick, (Landes R.G. Co., Austin, TX, USA, 1997), pp. 263–272
54. M.S. Widmer, A.G. Mikos, in *Frontiers in tissue engineering* ed. by C.W. Patrick, A.G. Mikos, V. McIntirel (Elsevier Science, New York, USA, 1998), pp. 107–120
55. K.F. Leong, C.M. Cheah, C.K. Chua, *Tissue Eng.* **8**(6), 1113 (2002)
56. R. Petzold, F. Zeilhofer, W.A. Kalender, *Comput. Med. Imaging Graphics.* **23**, 277–284 (1999)
57. P. Potamianos, A.A. Amis, A.J. Forester, M. McGurk, M. Bircher, *Proc. Inst. Mech. Eng.* **212**, 383–393 (1998)
58. S. Yang, K.H.F. Leong, Z. Du, C.K. Chua, *Tissue Eng.* **7**, 679–689 (2002)
59. I. Zein, D.W. Hutmacher, K.C. Tan, S.H. Teoh, *Biomaterials* **23**, 1169–1185 (2002)

60. S.J. Hollister, R.A. Levy, T.M. Chu, J.W. Halloran, S.E. Feinberg, *Int. J. Oral Maxillofac. Surg.* **29**, 67–71 (2000)
61. M.W. Naing, C.K. Chua, K.F. Leong, Y. Wang, *Rapid Prototyping J.* **11**(4), 249–259 (2005)
62. C.M. Cheah, C.K. Chua, K.F. Leong, C.H. Cheong, M.W. Naing, *Tissue Eng.* **10**, 595–609 (2004)
63. J.M. Taboas, R.D. Maddox, P.H. Krebsbach, S.J. Hollister, *Biomaterials* **24**, 181–94 (2003)
64. K.F. Leong, C.M. Cheah, C.K. Chua, *Biomaterials* **24**, 2363–2378 (2003)
65. A.G. Mikos, J.S. Temenoff, *Electron. J. Biotechnol.* **3**(2), 114–119 (2000)
66. S.N. Bhatia, C.S. Chen, *Biomed. Microdevices* **2**(2), 131–44 (1999)
67. B.S. Chang, C.K. Lee, K.S. Hong, H.J. Youn, H.S. Ryu, S.S. Chung, K.W. Park, *Biomaterials* **21**, 1291–1298 (2000)
68. P. Habibovic, H. Yuan, C.M. van der Valk, *Biomaterials* **26**, 3565–3575 (2005)
69. M. Okamoto, Y. Dohi, H. Ohgushi, H. Shimaoka, M. Ikeuchi, A. Matsushima, K. Yonemasu, H. Hosoi, *J Materials Science Mater. Med.* **17**, 327–336 (2006)
70. H. Yuan, K. Kurashina, J.D. de Bruijn, Y. Li, K. de Groot, X. Zhang, *Biomaterials* **20**, 1799–1806 (1999)
71. R.C. Tsay, J. Vo, A. Burke, S.B. Eisig, H.H. Lu, R. Landesberg, *J. Oral Maxillofac. Surg.* **63**, 521–528 (2005)
72. M. Klein, H. Goetz, S. Pazen, B.I. AlNawas, W. Wagner, H. Duschner, *Clin. Oral Implants Res.* **20**(1), 67–74 (2009)
73. T. Hildebrand, P. Ruegsegger, *Comput. Methods Biomech. Biomed. Eng.* **1**, 15–23 (1997)
74. J. Rich, H. Korhonen, R. Hakala, J. Korventausta, L. Elomaa, J. Seppälä, *Macromol Biosci.* (2009)
75. A.C. Jones, C.H. Arns, A.P. Sheppard, D.W. Hutmacher, B.K. Milthorpe, A.M. Knackstedt, *Biomaterials* **28**, 2491–2504 (2007)
76. O. Gauthier, R. Müller, D. von Stechow, B. Lamy, P. Weiss, J.M. Bouler, E. Aguado, G. Daculsi, *Biomaterials* **26**(27), 5444–53 (2005)
77. S.T. Ho, D.W. Hutmacher, *Biomaterials* **27**(8), 1362–76 (2006)
78. A. Goyal, Y. Wang, H. Su, L.W. Dobrucki, M. Brennan, P. Fong, A. Dardik, G. Tellides, A. Sinusas, J.S. Pober, W.M. Saltzman, C.K. Breuer, *J. Pediatr. Surg.* **41**, 787–791 (2006)
79. L.E. Niklason, J. Gao, W.M. Abbott, K.K. Hirschi, S. Houser, R. Marini, R. Langer, *Science* **284**, 489–493 (1999)
80. G. Matsumura, N. Hibino, Y. Ikada, H. Kurosawa, T. Shin’oka, *Biomaterials* **24**, 2303–2308 (2003)
81. B.J.R.F. Bolland, J.M. Kanczler, D.G. Dunlop, R.O. Oreffo, *Bone* **43**, 195–202 (2008)
82. E.M. Brey, T.W. King, C. Johnston, L.V. McIntire, G.P. Reece, C.W. Patrick, *Microvasc Res* **63**, 279–94 (2002)
83. B.J.R.F. Bolland, S. Tilley, A.M. New, D.G. Dunlop, R.O. Oreffo, *Expert Rev Med Devices* **4**, 393–404 (2007)
84. H. Peng, A. Usas, A. Olshanski, A.M. Ho, B. Gearhart, G.M. Cooper, J. Huard, *J. Bone Miner Res* **20**, 2017–27 (2005)
85. J.M. Kanczler, J. Barry, P. Ginty, S.M. Howdle, K.M. Shakesheff, R.O. Oreffo, *Biochem Biophys Res Commun* **352**, 135–41 (2007)
86. A.C. Langheinrich, E.L. Ritman, *Invest Radiol* **41**, 645–50 (2006)

87. V. Djonov, A.C. Andres, A. Ziemiecki, *Microsc Res Tech* **52**, 182–9 (2001)
88. E. Toyota, K. Fujimoto, Y. Ogasawara, T. Kajita, F. Shigeto, T. Matsumoto, M. Goto, F. Kajiya, *Circulation* **105**(5), 621–6 (2002)
89. D.C. Moore, C.W. Leblanc, R. Muller, J.J. Crisco, M.G. Ehrlich, *J. Orthop Res* **21**(3), 489–96 (2003)
90. G.L. Kindlmann, D.M. Weinstein, G.M. Jones, C.R. Johnson, M.R. Capecchi, C. Keller, *Mol Imaging* **4**, 417–24 (2005)
91. S. Heinzer, T. Krucker, M. Stampanoni, R. Abela, E.P. Meyer, A. Schuler, P. Schneider, R. Muller, *Neuroimage* **32**(2), 626–36 (2006)
92. C.L. Duvall, W.R. Taylor, D. Weiss, R.E. Guldberg, *Am. J. Physiol Heart Circ Physiol* **287**, H302–10 (2004)
93. N.M. Malyar, L.O. Lerman, M. Gossl, P.E. Beighley, E. Ritman, *Circulation* **110**, 1946–52 (2004)
94. M. Cioffi, F. Boschetti, M.T. Raimondi, G. Dubini, *Biotechnol Bioeng.* **93**(3), 500–10 (2006)
95. I. Martin, B. Obradovic, S. Treppo, A.J. Grodzinsky, R. Langer, L.E. Freed, G. Vunjak-Novakovic, *Biorheology* **37**, 141–147 (2000)
96. I. Martin, D. Wendt, M. Heberer, *Rev Trends Biotechnol* **22**(2), 80–86 (2004)
97. B.D. Porter, A.S. Lin, A. Peister, D. Huttmacher, R.E. Guldberg, *Biomaterials* **28**(15), 2525–33 (2007)
98. G.N. Bancroft, V.I. Sikavitsas, J. van den Dolder, T.L. Sheffield, C.G. Ambrose, J.A. Jansen, A.G. Mikos, *Proc. Natl. Acad. Sci. USA* **99**(20), 12600–12605 (2002)
99. S.H. Cartmell, B.D. Porter, A.J. Garcia, R.E. Guldberg, *Tissue Eng.* **9**(6), 1197–203 (2003)
100. V.I. Sikavitsas, G.N. Bancroft, H.L. Holtorf, J.A. Jansen, *Proc. Natl. Acad. Sci.* **100**(25), 14683–14688 (2003)
101. J. Klein-Nulend, M.H. Helfrich, J.G. Sterck, H. MacPherson, M. Joldersma, S.H. Ralston, C.M. Semeins, E.H. Burger, *Biochem Biophys Res Commun* **250**(1), 108–14 (1998)
102. T.N. McAllister, T. Du, J.A. Frangos, *Biochem Biophys Res Commun* **270**(2), 643–8 (2000)
103. F. Boschetti, M.T. Raimondi, F. Migliavacca, G. Dubini, *J Biomech* **39**(3), 418–25 (2006)
104. B. Porter, R. Zauel, H. Stockman, R. Guldberg, *J. Biomech* **38**(3), 543–9 (2005)
105. C. Provin, K. Takano, Y. Sakai, T. Fujii, R. Shirakashi, *J. Biomech* **41**(7), 1436–49 (2008)
106. V. Mironov, T. Boland, T. Trusk, G. Forgacs, R.R. Markwald, *Trends Biotechnol.* **21**(4), 157–161 (2003)
107. C.J. Bettinger, E.J. Weinberg, K.M. Kulig, J.P. Vacanti, Y. Wang, J.T. Borenstein, R. Langer, *Adv. Mater.* **18**, 165–169 (2006)
108. T.H. Yang, H. Miyoshi, N. Ohshima, *J Biomed. Mater. Res.* **55**, 379–386 (2001)
109. J. Malda, T.B.F. Woodfield, F. van der Vloodt, F.K. Kooy, D.E. Martens, J. Tramper, C.A. van Blitterswijk, J. Riesle, *Biomaterials* **25**, 5773–5780 (2004)
110. V. Karageorgiou, D. Kaplan, *Biomaterials* **26**, 5474–5491 (2005)
111. T. Nakamura, K. Yoshimoto, Y. Nakayama, Y. Tomita, A. Ichihara, *Proc. Natl. Acad. Sci. USA* **80**, 7229–7233 (1983)
112. A. Ichihara, *Dig. Dis. Sci.* **36**(4), 489–493 (1991)

113. M. Dvir-Ginzberg, I. Gamlieli-Bonshtein, R. Agbaria, S. Cohen, *Tissue Eng.* **9**(4), 757–766 (2003)
114. E. Eschbach, S.S. Chatterjee, M. Nöldner, E. Gottwald, H. Dertinger, K.F. Weibezahn, G. Knedlitschek, *J Cell. Biochem.* **95**, 243–255 (2005)
115. T. Hongo, M. Kajikawa, S. Ishida, S. Ozawa, Y. Ohno, J. Sawada, A. Umezawa, Y. Ishikawa, T. Kobayashi, H. Honda, *J Biosci. Bioeng.* **99**(3), 237–244 (2005)
116. J.P. Chen, T.C. Lin, *Biochem. Eng. J.* **30**(2), 192–198 (2006)
117. M.E. Hoque, H.Q. Mao, S. Ramakrishna, *J Biomater. Sci. Polym. Ed.* **18**(1), 45–58 (2007)
118. F. Consolo, G.B. Fiore, S. Truscello, M. Caronna, U. Morbiducci, F.M. Montevecchi, A. Redaelli, *Tissue Eng. Part A.* (2008)
119. J. Park, F. Berthiaume, M. Toner, M.L. Yarmush, A.W. Tilles, *Biotechnol. Bioeng.* **90**(5), 632–644 (2005)
120. E. Leclerc, B. David, L. Griscom, B. LePiouffle, T. Fujii, P. Layrolle, C. Legallais, *Biomaterials* **27**(4), 586–595 (2006)
121. C.Y.J. Ma, R. Kumar, X.Y. Xu, A. Mantalaris, *Biochem. Eng. J.* **35**, 1 (2007)
122. F. Galbusera, M. Cioffi, M.T. Raimodi, R. Pietrabissa, *Comput. Methods Biomech. Biomed. Engin.* **10**, 279 (2007)
123. M.T. Raimondi, M. Moretti, M. Cioffi, C. Giordano, F. Boschetti, *Biorheology* **43**, 215 (2006)
124. H. Singh, S.H. Teoh, H.T. Low, D.W. Hutmacher, *J. Biotechnol.* **119**, 181 (2005)
125. K.A. Williams, S. Saini, T.M. Wick, *Biotechnol. Prog.* **18**, 951–963 (2002)
126. B.G. Sengers, C.C. van Donkelaar, C.W. Oomens, F.P. Baaijens, *Biotechnol. Prog.* **21**, 1252 (2005)
127. G. Mareels, P.P. Poyck, S. Eloot, R.A. Chamuleau, P.R. Verdonck, *Ann Biomed. Eng.* **34**, 1729 (2006)
128. Y. Zeng, T.S. Lee, P. Yu, P. Roy, H.T. Low, *J. Biomech. Eng.* **128**, 185 (2006)
129. Versteeg et al., *An Introduction to Computational Fluid Dynamics: The Finite Volume Method Approach*, (Prentice Hall, 1995)
130. C. Cuvelier, A. Segal, A.A. van Steenhoven, *Finite Element Methods and Navier-Stokes Equations*, (Kluwer, 1986)
131. J.H. Ferziger, M. Peric, *Computational Methods for Fluid Dynamics*, (Springer, 1996)
132. C. Hirsch, *Numerical Computation of Internal and External Flows. Vol. I and II*, (Wiley, Chichester, 1990)
133. P. Wesseling, *Principles of Computational Fluid Dynamics*, (Springer, 2001)
134. J. Donea, A. Huerta, *Finite Element Methods for Flow Problems*, (Wiley, 2003)
135. U. Morbiducci, R. Ponzini, G. Rizzo, M. Cadioli, A. Esposito, F. De Cobelli, A. Del Maschio, F.M. Montevecchi, A. Redaelli, *Ann. Biomed. Eng.* **37**(3), 516–531 (2009)
136. M. Cantini, G.B. Fiore, A. Redaelli, M. Soncini, *Tissue Eng. Part A* (2008)
137. Z. Li, L. Li, *Trends Biochem. Sci.* **31**, 589 (2006)
138. T. Yin, L. Li, *J. Clin. Invest.* **116**, 1195 (2006)
139. C. Smith, *Cancer Control* **10**, 9 (2003)
140. I. Bellantuono, *Int. J. Biochem. Cell. Biol.* **36**, 607 (2004)
141. J. Domen, A. Wagers, I.L. Weissman, in *Regenerative Medicine 2006* ed. by NIH, (National Institutes of Health, U.S. Department of Health and Human Services, Bethesda, MD, 2006) pp. 13–34

142. G.J. Cabrita, B.S. Ferreira, C.L. da Silva, R. Goncalves, G. Almeida-Porada, J.M. Cabral, *Trends Biotechnol.* **21**, 233 (2003)
143. N. Jelinek, S. Schmidt, U. Hilbert, S. Thoma, M. Biselli, C. Wandrey, *Eng. Life Sci.* **2**, 5 (2002)
144. M.G. Levee, G. Lee, S. Paek, B.O. Palsson, *Biotechnol. Bioeng.* **43**, 734 (1994)
145. C.L. McDowell, E.T. Papoutsakis, *Biotechnol. Bioeng.* **60**, 239 (1998)
146. P. Meissner, B. Schröder, C. Herfurth, M. Biselli, *Cytotechnology* **30**, 227 (1999)
147. N. Panoskaltzis, A. Mantalaris, J.H. Wu, *J. Biosci. Bioeng.* **100**, 28 (2005)
148. M. Horner, W.M. Miller, J.M. Ottino, E.T. Papoutsakis, *Biotechnol. Prog.* **14**, 689 (1998)
149. P. Sucusky, Flow characterization and modeling of cartilage development in a spinner-flask bioreactor, Ph.D. Thesis, Georgia Institute of Technology, 2005
150. M. Cioffi, J. Küffer, S. Ströbel, G. Dubini, I. Martin, D. Wendt, *J. Biomech.* **41**(14), 2918–2925 (2008)
151. A.M. Freyria, Y. Yang, H. Chajra, C.F. Rousseau, M.C. Ronziere, D. Herbage, A.J. El Haj, *Tissue Eng.* **11**, 674–684 (2005)
152. D. Wendt, S. Stroebel, M. Jakob, G.T. John, I. Martin, *Biorheology* **43**(3–4), 481–488 (2006)
153. T. Davisson, R.L. Sah, A. Ratcliffe, *Tissue Eng.* **8**, 807–816 (2002)
154. P. Sucusky, D.F. Osorio, J.B. Brown, G.P. Neitzel, *Biotechnol. Bioeng.* **85**(1), 34–46 (2004)
155. B. Obradovic, J.H. Meldon, L.E. Freed, G. Vunjak-Novakovic, *AIChE J* **46**, 1860–1871 (200)
156. M.T. Raimondi, F. Boschetti, L. Falcone, G.B. Fiore, A. Remuzzi, E. Marinoni, M. Marazzi, R. Pietrabissa, *Biomech. Modeling Mechanobiol.* **1**, 69–82 (2002)
157. M.T. Raimondi, F. Boschetti, L. Falcone, F. Migliavacca, A. Remuzzi, G. Dubini, *Biorheology* **41**, 401–410 (2004)
158. P.J. Prendergast, *Clin. Biomech.* **12**(6), 343–366 (2007)
159. E. Votta, E. Caiani, F. Veronesi, M. Soncini, F.M. Montevocchi, A. Redaelli, *Philos. Trans. R. Soc. A* **366**, 3411–3434 (2008)
160. J.K. Hastings, M.A. Juds, J.R. Brauer, in *Accuracy and Economy of Finite Element Magnetic Analysis*, 33rd Annual National Relay Conference, April 1985
161. R. Huiskes, E.Y.S. Chao, *J. Biomech.* **16**, 385–409 (1983)
162. W.A.M. Brekelmans, H.W. Poort, T.J. Slooff, *Acta Orthop. Stand.* **43**, 301–317 (1972)
163. J. Mackerle, *Erzg. Comput.* **9**, 403–435 (1992)
164. J. Mackerle, *Finite Elem. Anal. Design.* **16**, 163–174 (1994)
165. H. Grandin, *Fundamentals of the Finite Element Method*, (Macmillan, New York, 1986)
166. S.C. Cowin, *J. Biomech. Eng.* **115**, 528–533 (1993)
167. A.J.C. Lee, in *The Art of Total Hip Arthroplasty*, ed. by W.T. Stillwell, (Grune and Stratton, Orlando, 1987), pp. 33–39
168. J. Guedes, N. Kikuchi, *Comput. Methods Appl. Mech. Eng.* **83**, 143–198 (1990)
169. D.P. Byrne, D. Lacroix, J.A. Planell, D.J. Kelly, P.J. Prendergast, *Biomaterials* **28**, 5544–5554 (2007)
170. F.J. O'Brien, B.A. Harley, I.V. Yannas, L. Gibson, *Biomaterials* **25**(6), 1077–86 (2004)

171. M. Liebschner, M. Wettergreen, in *Topics in tissue engineering*, ed. by P. Ferretti, N. Ashammakhi, (2003)
172. S. Langstaff, M. Sayer, T.J.N. Smith, S.M. Pugh, *Biomaterials* **22**(2), 135–150 (2001)
173. C.J. Damien, J.R. Parsons, *J. Appl. Biomater.* **2**(3), 187–208 (1991)
174. S. Yamada, D. Heymann, J.M. Bouler, G. Daculsi, *Biomaterials* **18**(15), 1037–41 (1997)
175. J. Sanz-Herrera, J. Garcia-Aznar, M. Doblaré, *Biomech. Model Mechanobiol.* (2007)
176. L. Pothuau, J.C. Fricain, S. Pallu, R. Bareille, M. Renard, M.C. Durrieu, M. Dard, M. Vernizeau, J. Amédée, *Biomaterials* **26**(33), 6788–6797 (2005)
177. T. Adachi, Y. Osako, M. Tanaka, M. Hojo, S.J. Hollister, *Biomaterials* **27**(21), 3964–3972 (2006)
178. F. Pauwels, *Grundriss einer Biomechanik der Frakturheilung. 34e Kongress der Deutschen Orthopädischen*, (Ferdinand Engke, Gesellschaft Stuttgart, 1941), 464–508
179. D.R. Carter, P.R. Blenman, G.S. Beaupre', *J. Orthopaed. Res.* **6**(5), 736–748 (1988)
180. L.E. Claes, C.A. Heigele, C. Neidlinger-Wilke, D. Kaspar, W. Seidl, K.J. Margevicius, P. Augat, *Clin. Orthopaed. Related Res. (Suppl. 355)*, 132–147 (1998)
181. R. Huiskes, W.D.V. Driel, P.J. Prendergast, *J. Mater. Sci. Mater. Med.* **8**(12), 785–788 (1997)
182. L. Geris, A. Andreykiv, H. Van Oosterwyck, J.V. Sloten, F. van Keulen, J. Duyck, I. Naert, *J. Biomech.* **37**(5), 763–769 (2004)
183. M.A. Perez, P.J. Prendergast, *J. Biomech.* **40**(10), 2244–2253 (2007)
184. D. Lacroix, P.J. Prendergast, G. Li, D. Marsh, *Med. Biol. Eng. Comput.* **40**(1), 14–21 (2002)
185. D.J. Kelly, P.J. Prendergast, *Tissue Eng* **12**(9), 2509–2519 (2006)
186. C.E. Wilson, J.D. de Bruijn, C.A. van Blitterswijk, A.J. Verbout, W.J.A. Dhert, *J. Biomed. Mater. Res. Part A* **68A**(1), 123–32 (2004)
187. R.Y. Hori, J.L. Lewis, *J. Biomed. Mater. Res.* **16**, 911–927 (1982)
188. L.E. Claes, C.A. Heigele, *J. Biomech.* **32**, 255–266 (1999)
189. C.G. Armstrong, V.C. Mow, *J. Bone Joint Surg. Am.* **64**, 88–94 (1982)
190. J.A. Ochoa, B.M. Hillberry, Permeability of bovine cancellous bone, *Transactions of the 38th ORS Meeting*, 1992
191. S. Tepic, T. Macirowski, R.W. Mann, *Biophysics* **80**, 3331–3333 (1983)
192. H.C. Anderson, *J. Cell. Biol.* **35**(1), 81–101 (1967)
193. J.B. Richardson, J. Kenwright, J.L. Cunningham, *Clin. Biomech.* **7**(2), 75–79 (1992)
194. A. Boccaccio, P.J. Prendergast, C. Pappalettere, D.J. Kelly, *Med. Biol. Eng. Comput.* **46**(3), 283–98 (2008)
195. J.A. Sanz-Herrera, J.M. Garcia-Aznar, M. Doblaré, *Acta. Biomater.* **5**(1), 219–29 (2009)
196. S.J. Hollister, R.D. Maddox, J.M. Taboas, *Biomaterials* **23**, 4095–4103 (2002)
197. S.J. Hollister, *Nat. Mater.* **4**, 518–524 (2005)
198. J.A. Sanz-Herrera, J.M. Garcia-Aznar, M. Doblaré, *Comput. Methods Appl. Mech. Eng.* **197**, 3092–3107 (2008)
199. P.M. Suquet, In: Sanchez-Palencia E, Zaoui A, editors. *Lecture Notes in Physics*, vol. 272. Berlin: Springer-Verlag; 1983:193–278

200. K. Terada, T. Ito, N. Kikuchi, *Comput. Methods Appl. Mech. Engrg.* **153**, 223–257 (1998)
201. G.S. Beaupré, T.E. Orr, D.R. Carter, *J. Orthopaed. Res.* **8**, 651–661 (1990)
202. R.E. Guldberg, S.J. Hollister, G.T. Charras, *J. Biomech. Eng.* **27**, 433–444 (1998)
203. H. Ohgushi, A.I. Caplan, *J. Biomed. Mater. Res. B* **48B**, 913–27 (1999)
204. C.Y. Lin, N. Kikuchi, S.J. Hollister, *J. Biomech.* **37**, 623–36 (2004)
205. J. Yao, S.R. Turteltaub, P. Ducheyne, *Biomaterials* **27**(3), 377–387 (2006)
206. M.A. Hussein, S. Esterl, R. Portner, K. Wiegandt, T. Becker, *J. Biomech.* **41**(16), 3455–3461 (2008)
207. X. Zhang, C. Xie, A.S.P. Lin, H. Ito, H. Awad, J.R. Lieberman, P.T. Rubery, E.M. Schwarz, R.J. O’Keefe, R.E. Guldberg, *J. Bone. Miner. Res.* **20**(12), 2124–2137 (2005)
208. A. Gopferich, *Macromolecules* **30**, 2598–2604 (1997)
209. V.L. Tsang, S.N. Bhatia, *Adv. Drug. Deliv. Rev.* **56**(11), 1635–47 (2004)
210. J.T. Borenstein, E.J. Weinberg, B.K. Orrick, C. Sundback, M.R. Kaazempur-Mofrad, P. Vacanti, *Tissue Eng.* **13**(8), 1837–1844 (2007)
211. M. Radisic, H. Park, H. Shing, T. Consi, F.J. Schoen, R. Langer, L.E. Freed, G. Vunjak-Novakovic, *Proc. Natl. Acad. Sci. USA* **101**(52), 18129–18134 (2004)
212. G.M. Cooper, *Cell*, (Sinauer Associates, Sunderland, 2000)
213. D. Boal, *Mechanics of the cell*, (Cambridge University Press, 2002)
214. R.D. Vale, *Cell* **112**(4), 467–480 (2003)
215. A. Maniotis, C. Chen, D.E. Ingber, *Proc. Natl. Acad. Sci. USA* **94**, 849–854 (1997)
216. C.G. Galbraith, K.M. Yamada, M.P. Sheetz, *J. Cell. Biol.* **159**(4), 695–705 (2002)
217. H.F. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, J. Darnell, *Molecular Cell Biology*, (Scientific American Books, W. H. Freeman and Company, New York, 1995)
218. R.J. Pelham Jr, Y. Wang, *Proc. Natl. Acad. Sci. USA* **94**, 13661–13665 (1997)
219. N. Wang, D.E. Ingber, *Biophys. J.* **66**, 2181–2189 (1994)
220. N. Wang, D. Stamenovic, *Am. J. Physiol. Cell Physiol.* **279**, C188–C194 (2000)
221. B. Fabry, G.N. Maksym, J.P. Butler, M. Glogauer, D. Navajas, J.F. Fredberg, *Phys. Rev. Lett.* **87**, 148102 (2001)
222. D.E. Discher, D.H. Boal, S.K. Boey, *Biophys. J.* **75**, 1584–1597 (1998)
223. S.C. Liu, L.H. Derick, J. Palek, *J. Cell Biol.* **104**, 522–528 (1987)
224. B. Alberts, D. Bray, J. Lewis, M. Raff, J. Watson, *Molecular biology of the cell*, (Garland, New York, 1994)
225. D.E. Ingber, *J. Cell. Sci.* **116**, 1157–1173 (2003)
226. H. Baudriller, B. Maurin, P. Cañadas, P. Montcourrier, A. Parmeggiani, N. Bettache, *C. R. Mec.* **334**, 662–668 (2006)
227. D. Stamenovic, D.E. Ingber, *Biomech. Model. Mechanobiol.* **1**, 95–108 (2002)
228. B.G. Sengers, M. Taylorb, C.P. Plesec, R. Oreffo, *Biomaterials* **28**, 1926–1940 (2007)
229. G. Lemon, J.R. King, H.M. Byrne, O.E. Jensen, *J. Math. Biol.* **52**(5), 571–594 (2006)
230. J.M. Zaleskas, B. Kinner, T.M. Freyman, I.V. Yannas, L.J. Gibson, M. Spector, *Biomaterials* **25**(7–8), 1299–1308 (2004)
231. J.T. Schantz, A. Brandwood, D.W. Hutmacher, H.L. Khor, K. Bittner, *J. Mater. Sci. Mater. Med.* **16**(9), 807–819 (2005)

232. B. Leukers, H. Gulkan, S.H. Irsen, S. Milz, C. Tille, M. Schieker, H. Seitz, J. Mater. Sci. Mater. Med. **16**(12), 1121–1124 (2005)
233. K. Tuzlakoglu, N. Bolgen, A.J. Salgado, M.E. Gomes, E. Piskin, R.L. Reis, J. Mater. Sci. Mater. Med. **16**(12), 1099–1104 (2005)
234. K. Nair, Multi-Scale computational modeling and characterization of bio-printed tissue scaffolds, PhD Thesis, Drexel University, US, 2008
235. R.G. Breuls, C.V. Bouten, C.W. Oomens, D.L. Bader, F.P. Baaijens, J. Biomech. Eng. **125**(6), 902–909 (2003)
236. N. Caille, O. Thoumine, Y. Tardy, J.J. Meister, J. Biomech. **35**(2), 177–187 (2002)
237. M.A. Haider, R.C. Schugart, L.A. Setton, F. Guilak, Biomech. Model Mechanobiol. **5**(2–3), 160–171 (2006)
238. M. Kellomaki, H. Niiranen, K. Puumanen, N. Ashammakhi, T. Waris, P. Tormala, Biomaterials **21**, 2495–2505 (2000)
239. D. Walsh, T. Furuzono, J. Tanaka, Biomaterials **22**, 1205–12 (2001)
240. P.V. Jackson, J.A. Hunt, P.J. Doherty, A. Cannon, P. Gilson, J. Mater. Sci. Mater. Med. **15**(4), 507–511 (2004)
241. P.A. Webb, Micromeritics Instrum. Corp. (2001)
242. C.X.F. Lam, S.H. Teoh, D.W. Hutmacher, In vitro degradation studies of customized PCL scaffolds fabricated via FDM, International Conference on Biological and Medical Engineering, (2002)
243. A.W. Paul, Micromeritics Instrum. Corp. 1–16 (2001)
244. S. Brunauer, *Physical Adsorption*, (Princeton University Press, Princeton, NJ, 1945)
245. P. Atkins, *Physical Chemistry*, (Freeman, New York, 1978)
246. G.A. Somorjai, *Principles of Surface Chemistry*, (Prentice-Hall, Englewood Cliffs, NJ, 1972)
247. S. Brunauer, P.H. Emmett, E. Teller, J. Am. Chem. Soc. **60**(2), 309–319 (1938)
248. I. Langmuir, J. Am. Chem. Soc. **40**(9), 1361–1403 (1918)
249. I. Langmuir, J. Am. Chem. Soc. **54**(7), 2798–2832 (1932)
250. I. Langmuir, *Nobel lecture*, (1932)
251. S.R. Peyton, J. Cell. Physiol. **204**, 198–209 (2005)
252. F. Grinnell, C.H. Ho, Y.C. Lin, G. Skuta, J. Biol. Chem. **274**, 918–23 (1999)
253. F. Grinnell, C.H. Ho, E. Tamariz, D.J. Lee, G. Skuta, Mol. Biol. Cell. **14**, 384–395 (2003)
254. A. Engler, L. Bacakova, C. Newman, A. Hategan, M. Griffin, D. Discher, Biophys. J. **86**, 617–628 (2004)
255. T. Yeung, P.C. Georges, L.A. Flanagan, B. Marg, M. Ortiz, M. Funaki, N. Zahir, W. Ming, V. Weaver, P.A. Janmey, Cell. Motil. Cytoskeleton. **60**, 24–34 (2005)
256. H. Jiang, F. Grinnell, Mol. Biol. Cell. **16**, 5070–5076 (2005)
257. J. Pelham, J. Robert, Y.L. Wang, Proc. Natl. Acad. Sci. USA **9**, 13661–13665 (1997)
258. T.M. Freyman, I.V. Yannas, R. Yokoo, L.J. Gibson, Biomaterials **22**, 2883–2891 (2001)
259. D. Schulz-Torres, T.M. Freyman, I.V. Yannas, M. Spector, Biomaterials **21**, 1607–1619 (2000)
260. M.H. Zaman, L.M. Trapani, A. Siemeski, R.D. Kamm, A. Wells, D.A. Lauffenburger, Proc. Natl. Acad. Sci. USA **103**, 10889–108894 (2006)

261. I.V. Yannas, *Tissue and organ regeneration in adults*, (Springer, New York, 2001)
262. G.C. Babis, P.N. Soucacos, *Int. J. Care Injured* **36S**, S38–S44 (2005)
263. B.A. Harley, J.H. Leung, E.C.C.M. Silva, L. Gibson, *Acta. Biomaterialia* **3**, 463–474 (2007)
264. D. Sarid, *Scanning Force Microscopy, Oxford Series in Optical and Imaging Sciences*, (Oxford University Press, New York, 1991)
265. A.C. Fisher-Cripps, *Nanoindentation*, (Springer, Berlin, 2002)
266. M.F. Doener, W.D.J. Nix, *Mater. Res* **1**, 601 (1986)
267. W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **7**, 1564–83 (1992)
268. S. Strasser, A. Zink, M. Janko, W.M. Heckl, S. Thalhammer, *Biochem. Biophys. Res. Commun.* **354**, 27–32 (2007)

Quantized Mechanics of Nanotubes and Bundles

Nicola M. Pugno

Summary. In this chapter, the mechanics of carbon nanotubes and related bundles is reviewed, with an eye to their application as ultra-sharp tips for scanning probe “nanoscopy”. In particular, the role of thermodynamically unavoidable, atomistic defects with different sizes and shapes on the fracture strength, fatigue life, and elasticity is quantified, thanks to new quantized fracture mechanics approaches. The reader is introduced in a simple way to such innovative treatments at the beginning of the chapter.

Key words: QFM, quantized, fracture, mechanics, strength, toughness, defect, nanotube, bundle, space elevator, scanning probe, microscopy

14.1 Introduction

Fabrication of carbon nanotubes (CNT) as nanoprobes in Scanning Probe Microscopy (SPM) dates back just to a decade ago [1, 2]; and today, the tremendous importance of such ultra-sharp tips is emerging (for a review see [3]). In the era of nanotechnology, “nanoscopy” techniques are, in fact, becoming more important as a consequence of their capability of exploring extremely small-scale phenomena. SPM is one of the most widely utilized microscopy techniques because it is a versatile tool for not only measuring the topology of surfaces but also manipulating the nanostructures. In spite of this, conventional silicon tips can easily break during an impact on the scanned surface and are not sufficiently slender to measure the topography of high-aspect ratio surfaces. In contrast, a CNT protruding from a conventional cantilever tip can be used as a strong ultra-sharp slender probe. If larger diameters are required, a nanotube bundle could replace the single nanotube. As the mechanical strength, elasticity, and slenderness of CNT and bundles are very high [4, 5], the CNT-probed “nanoscopy” brought a breakthrough in the development of a microscopy technique. However, defects can dramatically affect the mechanics of nanotubes and bundles, thus strongly limiting their performances.

In this chapter, the role of defects on the strength of nanotubes and bundles, based on new theoretical, deterministic, and statistical approaches of quantized fracture mechanics (QFM) proposed by the Author [6–9], is reviewed. The role of thermodynamically unavoidable, atomistic defects with different sizes and shapes is thus quantified on brittle fracture, fatigue, and elasticity, for nanotubes and bundles.

The chapter is organized in 12 short sections, as follows: Introduction is reported as the first section and Conclusions is reported as the last section; in Sect. 14.2, we review the QFM approaches; in Sect. 14.3, we apply them to calculate the fracture strength of nanotubes and bundles, containing defects with given sizes and shapes; whereas in Sect. 14.4, the impact strength is reported; in Sect. 14.5, elastic-plastic or hyper-elastic materials, rough cracks, and finite domains are treated; in Sect. 14.6, the fatigue life time is estimated; in Sect. 14.7, the Young’s modulus degradation is quantified; in Sects. 14.8 and 14.9, we compare our results on strength and elasticity with atomistic simulations and nanotensile tests of CNT; in Sect. 14.10, we demonstrate that defects are thermodynamically unavoidable, evaluating the minimum defect size and the corresponding maximum achievable strength; in Sect. 14.11, we calculate the strength of nanotube bundles by using hierarchical simulations and also discuss the related size effect.

14.2 Quantized Fracture Mechanics Approaches

According to the classical, continuum-based, Linear Elastic Fracture Mechanics (LEFM, [10]), the strength of a structure can be computed by Griffith’s energy balance during a crack propagation or, equivalently, by setting the stress-intensity factor K equal to its critical value, the fracture toughness of the material K_C , i.e., $K = K_C$. The stress-intensity factor K , for crack propagations – mode I (opening), mode II (sliding), or mode III (tearing), is only a function of the geometry and applied loads. On the other hand, if the hypothesis of the continuous crack advancement is relaxed, and thus a quantized energy balance is assumed [6], a more general (QFM, [7]) is formulated. The crack propagation will take place when

$$K^* = \sqrt{\langle K^2 \rangle_l^{l+\Delta l}} = K_C; \text{ Modes I, II, III,} \quad (14.1)$$

where K^* is the square root of the mean value of the square of the stress-intensity factor along a fracture quantum Δl , for a crack of length l .

Analogously, for dynamic loads the “mean” value of the stress-intensity factor must be considered during the time quantum Δt , connected to the time $\Delta l / c$ – with c crack speed – to generate a fracture quantum. Accordingly, for Dynamic Quantized Fracture Mechanics (DQFM, [8]) the crack propagation will take place when

$$K_d^* = \sqrt{\left\langle \langle K^2 \rangle_l^{l+\Delta l} \right\rangle_{t-\Delta t}^t} = K_C; \text{ Modes I, II, III.} \quad (14.2)$$

Equation (14.2), in analogy to Quantum Mechanics (QM) that is erected on the Planck’s constant \hbar , is based on the existence of the action quantum $G_C \Delta l \Delta t$, where G_C is the fracture energy of the material ($G_C = K_C^2 / E$, where E is the Young’s modulus).

Note that the classical, continuum-based, Dynamic Fracture Mechanics (DFM, [11]) would imply $K = K_{dC}$, where K_{dC} is the a priori, unknown dynamic fracture initiation toughness, observed to be experimentally different from K_C , especially for severe loading rates, e.g., impacts. On the other hand, (14.2) reproduces very well the experimental observations on times to failure also for severe loading rates [8]. Thus, DQFM can treat severe loading rates, e.g., impacts, in contrast to DFM. DFM corresponds to the limit case of DQFM for $\Delta t \rightarrow 0$, and becomes not predictive for severe impacts, requiring an *ad hoc* dynamic, fracture initiation toughness. In the DQFM treatment, the dynamic, fracture initiation toughness is identical to its static value, as must physically be. Moreover, classical LEFM, which corresponds to the limit case of QFM with $\Delta l \rightarrow 0$, can be applied only to “large” and sharp cracks, i.e., to cracks having length larger than the fracture quantum and vanishing tip radius. In contrast, QFM has no restriction in treating defects with any size and shape.

Instead of a classical, continuum-based, maximum stress criterion, i.e., $\sigma_{\max} = \sigma_C$, where σ_{\max} is the maximum stress in the structure and σ_C is the strength of the material, the stress analog of QFM must be written as [12,13]

$$\sigma^* = \langle \sigma_{\text{tip}} \rangle_0^{\Delta l} = \sigma_C \text{ Mode I; for Modes II, III : } \sigma \rightarrow \tau, \quad (14.3)$$

where σ_{tip} is the opening – for mode I, stress at the tip of a defect, where is located the origin of the reference system; for modes II or III – the normal stress and strength must be evidently replaced by the corresponding shear stress τ and strength τ_C .

For dynamic loads this stress criterion has to be rewritten as [14]:

$$\sigma_d^* = \left\langle \langle \sigma_{\text{tip}} \rangle_0^{\Delta l} \right\rangle_{t-\Delta t}^t = \sigma_C \text{ Mode I; for Modes II, III : } \sigma \rightarrow \tau, \quad (14.4)$$

representing the stress analog of DQFM.

Equations (14.1) and (14.2) are based on stress-intensity factors, whereas (14.3) and (14.4) on stress. However, note that considering, for example QFM, the free-parameter Δl can be fixed to reproduce for $l \rightarrow 0$ the classical criterion $\sigma_{\max} = \sigma_C$: thus, QFM implies a smooth transition between the criteria of $\sigma_{\max} = \sigma_C$ for vanishing crack length to $K = K_C$ for large cracks (where $K^* \approx K$). Moreover, imposing that the criteria of (14.1) and (14.3) have to predict the same failure for each value of l corresponds to a mixed criterion [9, 15]. The corresponding fracture quantum capable of ensuring such an equality for each value of l can be derived consequently: in this case, the fracture quantum $\Delta l(l)$ becomes, more than a material constant, a well-defined

material/structural parameter. In formulae,

$$K^* = K_C \& \sigma^* = \sigma_C \Delta l : \text{same predictions, Mode I; or } \sigma \rightarrow \tau \text{ for Modes II, III.} \tag{14.5}$$

We finally note that (14.1) and (14.3) have been successfully applied also for fatigue limit predictions [16]. The corresponding, fatigue limit criteria can be formally written considering the variations of Δ in front of the symbols, interpreting ΔK^* and $\Delta \sigma^*$ as the amplitude ranges of K^*, σ^* in a cycle, ΔK_C as the threshold value of the stress-intensity factor, and $\Delta \sigma_C$ as the plain-specimen fatigue limit.

As suggested by (14.5), imposing the same strength and/or time to failure predictions from (14.2) and (14.4), a new dynamic mixed criterion, in which both fracture as well as time quanta are derived, to ensure the equality of such predictions, is formulated [9]:

$$K_d^* = K_C \& \sigma_d^* = \sigma_C \Delta l, \Delta t : \text{same predictions, Mode I;} \\ \text{or } \sigma \rightarrow \tau \text{ for Modes II, III} \tag{14.6}$$

In addition, substituting the stress with the corresponding strain in (14.3) and (14.4), normal, ε , for mode I; or tangential, γ , for mode II and III; a strain static

$$\varepsilon^* = \langle \varepsilon_{\text{tip}} \rangle_0^{\Delta l} = \varepsilon_C \text{ Mode I; for Modes II, III : } \varepsilon \rightarrow \gamma, \tag{14.7}$$

and dynamic criteria

$$\varepsilon_d^* = \left\langle \langle \varepsilon_{\text{tip}} \rangle_0^{\Delta l} \right\rangle_{t-\Delta t}^t = \varepsilon_C \text{ Mode I; for Modes II, III : } \varepsilon \rightarrow \gamma, \tag{14.8}$$

are derived [9].

The quantized criteria of (14.3–14.8) require in general the expression of the complete – and not only asymptotic – stress field around the tip of the defect, well-known only for the simplest cases. On the other hand, the criteria of (14.1) and (14.2) can be applied in a very simple way, by starting from the well-known solutions for the stress-intensity factors; for example, hundreds of static and dynamic solutions are reported in the classical Murakami’s Handbook [17]. Obviously, the predictions of different criteria are not coincident, but always similar, representing the asymptotic matching between the two classical regimes.

We note that not only (14.1) and (14.3) but all the criteria (1–8) can be rewritten for fatigue limit predictions, formally introducing the variations Δ in front of the symbols. However, such criteria estimate the beginning of the fatigue crack growth but not its evolution. On the other hand, regarding the evolution of the fatigue crack, substituting the stress-intensity factor K with its “quantized” version K^* in the Paris’ law – or in its classical extensions – a quantized Paris’ law, which has to be applied to short cracks and small systems too, is formulated [9] as

$$\frac{dl}{dN} \approx A (\Delta K^*)^\alpha, \tag{14.9}$$

where N is the number of cycles, A and α are the Paris' constants, and ΔK^* is the amplitude range of the "quantized" stress-intensity factor in a cycle.

For very short cracks $K^* \propto \sigma$ and (14.9) resembles the classical Whöler's law, i.e., $N_f (\Delta\sigma)^\beta = B$, with B, β constants and N_f life time. Note that, similarly to brittle fracture, the fracture quantum itself could be fixed in fatigue to derive from (14.9) and in the limit case of crack length tending to zero, the same Whöler's prediction for the life time N_f [18, 19]. On the other hand, for very large cracks ($K^* \approx K$) (14.9) becomes the classical Paris' law. Equation (14.9) is a stress-intensity factor-based, quantized criterion for a fatigue crack growth. The stress analog can be formulated substituting $\Delta\sigma$ with $\Delta\sigma^*$ in the classical Whöler's law, i.e. [9]

$$N_f (\Delta\sigma^*)^\beta = B. \tag{14.10}$$

Equations (14.9) and (14.10) correspond in fatigue crack growth to (14.1) and (14.3) or (14.2) and (14.4) in static or dynamic fracture. Accordingly, starting from these two analogs, it is clear that all the analogs of (14.1–14.8) can be easily formulated for fatigue crack growth also; for example, the mixed criterion of (14.5), ensuring the same fatigue life prediction N_f , has to be written as [9]:

$$\frac{dl}{dN} \approx A (\Delta K^*)^\alpha \ \& \ N_f (\Delta\sigma^*)^\beta = B \ \Delta l : \text{ same predictions.} \tag{14.11}$$

Finally, also the Weibull [20] statistical theory for the strength of solids can be quantized [9]. According to Weibull, the probability of failure P_f of a specimen of volume V under uniaxial stress σ is given by $P_f = 1 - \exp\left(-\frac{1}{V_0} \int (\sigma / \sigma_0)^m dV\right)$, where V_0 is a reference volume and σ_0 and m are two constants. If stress intensifications are present, as in cracked structures, the Weibull's integral does not converge: this represents a limit of the classical Weibull's statistics and can automatically be removed if instead of σ its "quantized" version σ^* (or σ_d^*) is considered, as

$$P_f = 1 - \exp\left(-\frac{1}{V_0} \int_V (\sigma^* / \sigma_0)^m dV\right). \tag{14.12}$$

Thus, the quantized crack advancement removes a paradox and a new statistics is generated.

In general, we conclude emphasizing that our definition of K^* and σ^* (also in dynamics) allow one to "quantize" classical, well-known criteria, based on stress and/or on stress-intensity factor. More powerful quantized approaches will result and the classical ones will be automatically recovered for the limit case of vanishing quanta, as required by the "Corresponding Principle." These

new criteria are ideal to study small-scale objects, where the intrinsic material strength cannot be considered further as infinitely high, as in nanotubes. The applications to nanotubes and bundles are presented in the following sections of the chapter.

14.3 Fracture Strength

By considering QFM [6–9], the failure stress σ_N for a nanotube having an atomic size q (the “fracture quantum”) and containing an elliptical hole of half-axes a , perpendicular to the applied load, or nanotube axis; and b can be determined including in the asymptotic solution [7] the contribution of the far-field stress. We accordingly derive [21]

$$\frac{\sigma_N(a, b)}{\sigma_N^{(\text{theo})}} = \sqrt{\frac{1 + 2a/q(1 + 2a/b)^{-2}}{1 + 2a/q}}, \sigma_N^{(\text{theo})} = \frac{K_{IC}}{\sqrt{q\pi/2}}, \quad (14.13)$$

where $\sigma_N^{(\text{theo})}$ is the theoretical (defect-free) nanotube strength (~ 100 GPa) and K_{IC} is the material fracture toughness. The self-interaction between the tips has been neglected here (i.e., $a \ll \pi R$, with R nanotube radius) and would further reduce the failure stress (finite domains are treated in Sect. 14.5).

Regarding the defect shape, for a sharp crack perpendicular to the applied load $a/q = \text{const}$ and $b/q \rightarrow 0$, thus $\sigma_N \approx \sigma_N^{(\text{theo})} / \sqrt{1 + 2a/q}$, and for $a/q \gg 1$, i.e., large cracks, $\sigma_N \approx K_{IC} / \sqrt{\pi a}$ in agreement with LEFM; note that LEFM can (1) only treat sharp cracks and (2) unreasonably predicts an infinite defect-free strength. On the other hand, for a crack parallel to the applied load $b/q = \text{const}$ and $a/q \rightarrow 0$ and thus, $\sigma_N = \sigma_N^{(\text{theo})}$, as it must be. In addition, regarding the defect size, for self-similar and small holes $a/b = \text{const}$ and $a/q \rightarrow 0$ and coherently $\sigma_N = \sigma_N^{(\text{theo})}$; furthermore, for self-similar and large holes $a/b = \text{const}$ and $a/q \rightarrow \infty$, and we deduce $\sigma_N \approx \sigma_N^{(\text{theo})} / (1 + 2a/b)$ in agreement with the stress concentration posed by elasticity; but elasticity (coupled with a maximum stress criterion) unreasonably predicts (3) a strength independent from the hole size and (4) tending to zero for cracks. Note the extreme consistency of (14.13), that removing all the discussed limitations (1–4) represents the first law capable of describing in a unified manner all the size- and shape effects for the elliptical holes, including cracks as limit case [21]. In other words, (14.13) shows that the two classical strength predictions based on stress intensifications (LEFM) or concentrations (elasticity) are only reasonable for “large” defects; (14.13) unifies their results and extends its validity to “small” defects (“large” and “small” are here with respect to the fracture quantum). It shows that even a small defect can dramatically reduce the mechanical strength of a nanotube, e.g., limiting the applications of ultra-sharp, nanotube-based tips, even if, in this context,

more significant are the impact strength and fatigue life of the nanotube tip, treated in Sects. 14.4 and 14.6, respectively.

For a nanotube bundle, an upperbound of the strength can be derived assuming the simultaneous failure of all the defective nanotubes present in the bundle. Accordingly, imposing the critical force equilibrium (mean-field approach) for a cable composed by nanotubes in numerical fractions f_{ab} , containing holes of half-axes a and b , we find the cable strength σ_C (ideal if $\sigma_C^{(\text{theo})}$) in the following form [21]:

$$\frac{\sigma_C}{\sigma_C^{(\text{theo})}} = \sum_{a,b} f_{ab} \frac{\sigma_N(a,b)}{\sigma_N^{(\text{theo})}}. \tag{14.14}$$

The summation is extended to all the different holes; the numerical fraction f_{00} of nanotubes is defect-free and $\sum_{a,b} f_{ab} = 1$. If all the defective nanotubes in the bundle contain identical holes $f_{ab} = f = 1 - f_{00}$ and $1 - \sigma_C / \sigma_C^{(\text{theo})} = f(1 - \sigma_N / \sigma_N^{(\text{theo})})$.

14.4 Impact Strength

Let us consider the simplest case of a semi-infinite crack in an otherwise unbounded body. The body is initially stress-free and at rest. At time $t = 0$ a pressure σ begins to act on the crack faces. In this case, as it is well-known, $K_I(t) = 2\sigma \frac{\sqrt{c_D t(1-2\nu)/\pi}}{(1-\nu)}$ [11], where c_D is the dilatational wave speed of the material and ν is its Poisson’s ratio. Applying (14.2) we find the failure for a given time $t_f > \Delta t$, satisfying [9]

$$K_I(t) = \frac{K_{IC}}{\sqrt{1 - \Delta t / (2t_f)}} \equiv K_{dIC}. \tag{14.15}$$

Thus, if classical DFM [11] is applied, i.e., $K_I(t) = K_{dIC}$, the “measured” fracture initiation toughness K_{dIC} will be observed, according to DQFM, time to failure dependent. In addition, note that, according to our time quantization, a minimum time to failure exists and it must be of the order of $t_{f \min} \approx \Delta t$. Considering very severe impacts ($t_f \rightarrow t_{f \min} \approx \Delta t$), the dynamic strength ($\propto K_{dIC}$) is expected for this scheme $\sqrt{2}$ times larger than its static value ($\propto K_{IC}$). For an applied pressure linearly increasing with time, the factor $\sqrt{2}$ is replaced by the factor 2 [8]. For different schemes a slightly different factor is expected. Roughly speaking, we could call this effect as the “doubling” of the impact strength (with respect to its static value). This is a well-known, experimental phenomenon [8] and is thus expected for nanotubes also.

14.5 Hyper-Elasticity, Elastic-Plasticity, Fractal Cracks, and Finite Domains

Equation (14.13) is based on linear elasticity, i.e., on a linear relationship $\sigma \propto \varepsilon$ between stress σ and strain ε . In contrast, let us assume $\sigma \propto \varepsilon^\kappa$, where $\kappa > 1$ denotes hyper-elasticity, as well as $\kappa < 1$ denotes elastic-plasticity. The power of the stress-singularity will accordingly be modified [22] from the classical value $1/2$ to $\alpha = \kappa / (\kappa + 1)$. Thus, the problem is mathematically equivalent to that of a re-entrant corner [23], and consequently we predict

$$\frac{\sigma_N(a, b, \alpha)}{\sigma_N^{(\text{theo})}} = \left(\frac{\sigma_N(a, b)}{\sigma_N^{(\text{theo})}} \right)^{2\alpha}, \alpha = \frac{\kappa}{\kappa + 1}. \quad (14.16)$$

A crack with a self-similar roughness, mathematically described by a fractal with non-integer dimension $1 < D < 2$, would similarly modify the stress-singularity, according to [24] $\alpha = (2 - D) / 2$; thus, with (14.16), we can also estimate the role of the crack roughness. Both plasticity and roughness reduce the severity of the defect, whereas hyper-elasticity enlarges its effect. For example, for a crack composed by n adjacent vacancies, we deduce $\sigma_N / \sigma_N^{(\text{theo})} \approx (1 + n)^{-\alpha}$. However, note that among these three effects only elastic-plasticity may have a significant role in CNT; in spite of this, fractal cracks could play an important role in nanotube bundles as a consequence of their larger size-scale, which would allow the development of a crack surface roughness. Hyper-elasticity is not expected to be relevant in this context.

Equation (14.13) does not consider the defect–boundary interaction. A graphene sheet having a finite width $2W$ can be treated by applying QFM starting from the related expression of the stress-intensity factor. However, to have an idea of the defect–boundary interaction, we can couple (14.13) with an approximated method [21], deriving the following correction $\sigma_N(a, b, W) \approx C(W) \sigma_N(a, b)$, $C(W) \approx (1 - a / W) / \left(\sigma_N(a, b) |_{q \rightarrow W-a} / \sigma_N^{(\text{theo})} \right)$ (note that such a correction is valid also for $W \approx a$, whereas for $W \gg a$ it becomes $C(W \gg a) \approx 1 - a / W$). Similarly, the role of the defect orientation β could be treated by QFM considering the related, stress-intensity factor; roughly, one could use in (14.13) the self-consistent approximation $\sigma_N(a, b, \beta) \approx \sigma_N(a, b) \cos^2 \beta + \sigma_N(b, a) \sin^2 \beta$.

14.6 Fatigue Life

The SPM nanotube tip is cyclically loaded during surface scanning, thus fatigue plays a major role. By integrating the quantized Paris' law of (14.9), we derive the following number of cycles to failure, or life time [21]:

$$\frac{N_N(a)}{N_N^{(\text{theo})}} = \frac{(1 + q / W)^{1-m/2} - (a / W + q / W)^{1-m/2}}{(1 + q / W)^{1-m/2} - (q / W)^{1-m/2}}, m \neq 2 \quad (14.17a)$$

and

$$\frac{N_N(a)}{N_N^{(\text{theo})}} = \frac{\ln \{(1 + q/W) / (a/W + q/W)\}}{\ln \{(1 + q/W) / (q/W)\}}, m = 2 \tag{14.17b}$$

where $m > 0$ is the material Paris' exponent. Note that according to Wöhler $N_N^{(\text{theo})} = K \Delta \sigma^{-k}$, where K and k are material constants and $\Delta \sigma$ is the amplitude of the stress range during the oscillations.

Only defects remaining self-similar during fatigue growth have to be considered, thus only a crack (of half-length a) is of interest in this context. By means of (14.17) the time to failure can be estimated, similarly to the brittle fracture treated by (14.13).

For a bundle, considering a mean-field approach (similarly to (14.14)) yields:

$$\frac{N_C}{N_C^{(\text{theo})}} = \sum_a f_a \frac{N_N(a)}{N_N^{(\text{theo})}}. \tag{14.18}$$

Better predictions could be derived integrating the quantized Paris' law for a finite width strip. However, we note that the role of the finite width is already included in (14.17), even if these are rigorously valid in the limit of W tending to infinity.

14.7 Elasticity

Consider a nanotube of lateral surface A under tension and containing a transversal crack of half-length a . Interpreting the incremental compliance, due to the presence of the crack, as a Young's modulus (here denoted by E) degradation we find $\frac{E(a)}{E^{(\text{theo})}} = 1 - 2\pi \frac{a^2}{A}$ [25]. Thus, recursively, considering Q cracks having sizes a_i or, equivalently, M different cracks with multiplicity Q_i ($Q = \sum_{i=1}^M Q_i$), noting that $n_i = \frac{2a_i}{q}$ represents the number of adjacent vacancies in a crack of half-length a_i , with q atomic size, and $v_i = \frac{Q_i n_i}{A/q^2}$ its related numerical (or volumetric) vacancy fraction, we find [25]:

$$\frac{E}{E^{(\text{theo})}} = \prod_{i=1}^Q \frac{E(a_i)}{E^{(\text{theo})}} \approx 1 - \xi \sum_{i=1}^M v_i n_i, \tag{14.19}$$

with $\xi \geq \pi / 2$, where the equality holds for isolated cracks. Equation (14.19) can be applied to nanotubes or nanotube bundles containing defects in volumetric percentages v_i .

Forcing the interpretation of our formalism, we note that $n_i = 1$ would describe a single vacancy, i.e., a small hole. Thus, as a first approximation, different defect geometries, from cracks to circular holes, e.g., elliptical holes,

could in principle be treated by (14.19); we have to interpret n_i as the ratio between the transversal and longitudinal (parallel to the load) defect sizes ($n_i = a_i / b_i$). Introducing the i -th defect eccentricity e_i as the ratio between the lengths of the longer and shorter axes, as a first approximation $n_i(\beta_i) \approx e_i \cos^2 \beta_i + 1 / e_i \sin^2 \beta_i$, where β_i is the defect orientation. For a single defect typology $\frac{E}{E^{(\text{theo})}} \approx 1 - \xi v n$, in contrast to the common assumption $\frac{E}{E^{(\text{theo})}} \approx 1 - v$, rigorously valid only for the material density, for which $\frac{\rho_C}{\rho_C^{(\text{theo})}} \equiv 1 - v$. Note that the failure strain for a defective nanotube or nanotube bundle can also be predicted, by $\varepsilon_{N,C} / \varepsilon_{N,C}^{(\text{theo})} = \left(\sigma_{N,C} / \sigma_{N,C}^{(\text{theo})} \right) / \left(E / E^{(\text{theo})} \right)$.

In contrast to what happens for fracture and fatigue, defects affect the elasticity in a significant but not dramatic way.

14.8 Atomistic Simulations

Let us study the influence on the strength of nano-cracks and circular nano-holes. n atomic adjacent vacancies perpendicular to the load, correspond to a blunt nano-crack of length $2a \approx nq$ and thickness $2b \approx q$ (or $2a \approx nq$ with a radius at tips of $b^2 / a \approx q / 2$). Similarly, nano-holes of size m can be considered: the index $m = 1$ corresponds to the removal of an entire hexagonal ring, $m = 2$ to the additional removal of the six hexagons around the former one (i.e. the adjacent perimeter of 18 atoms), $m = 3$ to the additional removal of the neighbouring 12 hexagonal rings (next adjacent perimeter), and so on (thus $a = b \approx q(2m - 1) / \sqrt{3}$) [7]. QM, semi-empirical (PM3 method), Molecular Mechanics (MM; with a modified Tersoff-Brenner potential of second generation (MTB-G2) or a modified Morse potential (M)) and coupled QM/MM calculations [26–30] are reported and extensively compared in Table 14.1 with the QFM predictions of (14.13) [7, 21]. The comparison shows a relevant agreement, confirming and demonstrating that just a few vacancies can dramatically reduce the strength of a single nanotube or of a nanotube bundle as described by (14.14) that predicts for $f \approx 1$, $\sigma_C / \sigma_C^{(\text{theo})} \approx \sigma_N / \sigma_N^{(\text{theo})}$. For example, assuming large holes ($m \rightarrow \infty$) and applying QFM to a defective bundle ($f \approx 1$), we predict $1 - \sigma_C / \sigma_C^{(\text{theo})} \approx 1 - \sigma_N / \sigma_N^{(\text{theo})} \approx 67\%$.

Note that an elastic ($\kappa \approx 1$) nearly perfectly plastic ($\kappa \approx 0$) behaviour, with a flow stress at ~ 30 – 35 GPa for strains larger than ~ 3 – 5% , has been recently observed in tensile tests of CNT [31], globally suggesting $\kappa \approx 0.6$ – 0.7 ; similarly, numerically computed stress-strain curves [32] reveal for an armchair (5,5) carbon nanotube $\kappa \approx 0.8$, whereas for a zig-zag (9,0) nanotube $\kappa \approx 0.7$, suggesting that the plastic correction reported in Sect. 14.5 could have a role.

Regarding elasticity, we note that (14.19) [25] can be viewed as a generalization of the approach proposed in [33] considering three different types of defects.

Table 14.1. Atomistic simulations [26–29] vs. QFM predictions, for nano-cracks of size n or nano-holes of size m

Nanotube Type	Nanocrack (n) and nanohole (m) sizes	Strength [GPa] by QM (MTB-G2) and MM (PM3; M) QM/MM atomistic or QFM calculations
[5,5]	Defect-free	105 (MTB-G2); 135 (PM3)
[5,5]	$n = 1$ (sym. + H)	85 (MTB-G2), 79 (QFM); 106 (PM3), 101 (QFM)
[5,5]	$n = 1$ (Asym. + H)	71 (MTB-G2), 79 (QFM); 99 (PM3), 101 (QFM)
[5,5]	$n = 1$ (Asym.)	70 (MTB-G2), 79 (QFM); 100 (PM3), 101 (QFM)
[5,5]	$n = 2$ (Sym.)	71 (MTB-G2), 63 (QFM); 105 (PM3), 81 (QFM)
[5,5]	$n = 2$ (Asym.)	73 (MTB-G2), 63 (QFM); 111 (PM3), 81 (QFM)
[5,5]	$m = 1$ (+H)	70 (MTB-G2), 68 for long tube, 79 (QFM); 101 (PM3), 101 (QFM)
[5,5]	$m = 2$ (+H)	53 (MTB-G2), 50 for long tube, 67 (QFM); 78 (PM3), 86 (QFM)
[10,10]	Defect-free	88 (MTB-G2); 124 (PM3)
[10,10]	$n = 1$ (sym. + H)	65 (MTB-G2), 66 (QFM)
[10,10]	$n = 1$ (Asym. + H)	68 (MTB-G2), 66 (QFM)
[10,10]	$n = 1$ (Sym.)	65 (MTB-G2), 66 (QFM); 101 (PM3), 93 (QFM)
[10,10]	$n = 2$ (Sym.)	64 (MTB-G2), 53 (QFM); 107 (PM3), 74 (QFM)
[10,10]	$n = 2$ (Asym.)	65 (MTB-G2), 53 (QFM); 92 (PM3), 74 (QFM)
[10,10]	$m = 1$ (+H)	56 (MTB-G2), 52 for long tube, 66 (QFM); 89 (PM3), 93 (QFM)
[10,10]	$m = 2$ (+H)	42 (MTB-G2), 36 for long tube, 56 (QFM); 67 (PM3), 79 (QFM)
[50,0]	Defect-free	89 (MTB-G2)
[50,0]	$m = 1$ (+H)	58 (MTB-G2); 67 (QFM)
[50,0]	$m = 2$ (+H)	46 (MTB-G2); 57 (QFM)
[50,0]	$m = 3$ (+H)	40 (MTB-G2); 44 (QFM)
[50,0]	$m = 4$ (+H)	36 (MTB-G2); 41 (QFM)
[50,0]	$m = 5$ (+H)	33 (MTB-G2); 39 (QFM)
[50,0]	$m = 6$ (+H)	31 (MTB-G2); 37 (QFM)
[100,0]	Defect-free	89 (MTB-G2)
[100,0]	$m = 1$ (+H)	58 (MTB-G2); 67 (QFM)
[100,0]	$m = 2$ (+H)	47 (MTB-G2); 57 (QFM)
[100,0]	$m = 3$ (+H)	42 (MTB-G2); 44 (QFM)
[100,0]	$m = 4$ (+H)	39 (MTB-G2); 41 (QFM)
[100,0]	$m = 5$ (+H)	37 (MTB-G2); 39 (QFM)
[100,0]	$m = 6$ (+H)	35 (MTB-G2); 37 (QFM)
[29,29]	Defect-free	101 (MTB-G2)
[29,29]	$m = 1$ (+H)	77 (MTB-G2); 76 (QFM)
[29,29]	$m = 2$ (+H)	62 (MTB-G2); 65 (QFM)
[29,29]	$m = 3$ (+H)	54 (MTB-G2); 50 (QFM)
[29,29]	$m = 4$ (+H)	48 (MTB-G2); 46 (QFM)
[29,29]	$m = 5$ (+H)	45 (MTB-G2); 44 (QFM)
[29,29]	$m = 6$ (+H)	42 (MTB-G2); 42 (QFM)

(Continued)

Table 14.1. (Continued)

Nanotube Type	Nanocrack (n) and nanohole (m) sizes	Strength [GPa] by QM (MTB-G2) and MM (PM3; M) QM/MM atomistic or QFM calculations
[47,5]	Defect-free	89 (MTB-G2)
[47,5]	$m = 1(+H)$	57 (MTB-G2); 67 (QFM)
[44,10]	Defect-free	89 (MTB-G2)
[44,10]	$m = 1(+H)$	58 (MTB-G2); 67 (QFM)
[40,16]	Defect-free	92 (MTB-G2)
[40,16]	$m = 1(+H)$	59 (MTB-G2); 69 (QFM)
[36,21]	Defect-free	96 (MTB-G2)
[36,21]	$m = 1(+H)$	63 (MTB-G2); 72 (QFM)
[33,24]	Defect-free	99 (MTB-G2)
[33,24]	$m = 1(+H)$	67 (MTB-G2); 74 (QFM)
[80, 0]	Defect-free	93 (M)
[80, 0]	$n = 2$	64 (M); 56 (QFM)
[80, 0]	$n = 4$	50 (M); 43 (QFM)
[80, 0]	$n = 6$	42 (M); 35 (QFM)
[80, 0]	$n = 8$	37 (M); 32 (QFM)
[40, 0]	Defect-free	99 (M)
(nested by a [32, 0])		
[40, 0]	$n = 2$	73 (M); 69 (QFM + vdW interaction ~ 10 GPa)
(nested by a [32, 0])		
[40, 0]	$n = 4$	57 (M); 56 (QFM + vdW interaction ~ 10 GPa)
(nested by a [32, 0])		
[40, 0]	$n = 6$	50 (M); 48 (QFM + vdW interaction ~ 10 GPa)
(nested by a [32, 0])		
[40, 0]	$n = 8$	44 (M); 44 (QFM + vdW interaction ~ 10 GPa)
(nested by a [32, 0])		
[100,0]	Defect-free	89 (MTB-G2)
[100,0]	$n = 4$	50 (M); 41 (QFM)
[10,0]	Defect free	124 (QM); 88 (MM);
[10,0]	$N = 1$	101 (QM) 95 (QM/MM) 93 (QFM); 65 (MM) 66 (QFM)

The QFM predictions are here obtained simply considering in (14.13) $2a/q = n$, $2b/q = 1$ for cracks of size n or $a/q = b/q = (2m-1)/\sqrt{3}$ for holes of size m . Quantum mechanics (QM) semi-empirical calculations (PM3 method), Molecular Mechanics (MM) calculations (modified Tersoff-Brenner potential of second generation (MTB-G2), modified Morse potential (M)) and coupled QM/MM calculations. The symbol (+H) means that the defect was saturated with hydrogen. Symmetric and asymmetric bond reconstructions were also considered; the tubes are “short”, if not otherwise specified. We have roughly ignored in the QFM predictions the difference between symmetric and asymmetric bond reconstruction, hydrogen saturation and length effect (for shorter tubes an increment in the strength is always observed, as an intrinsic size effect), noting that the main differences in the atomistic simulations are imputable to the used potential. For nested nanotubes a strength increment of ~ 10 GPa is here assumed to roughly take into account the van der Walls (vdW) interaction between the walls

14.9 Nanotensile Tests

The discussed tremendous defect sensitivity, described by (14.13), is confirmed by a statistical analysis based on Nanoscale Weibull Statistics (NWS, [34]) applied to the nanotensile tests. According to this treatment, the probability of failure P for a nearly defect-free nanotube under a tensile stress σ_N is independent from its volume (or surface), in contrast to classical Weibull Statistics [20], namely:

$$P = 1 - \exp -N_N \left(\frac{\sigma_N}{\sigma_0} \right)^w, \tag{14.20}$$

where w is the nanoscale Weibull modulus, σ_0 is the nominal failure stress (i.e. corresponding to a probability of failure of 63%) and $N_N \equiv 1$. In classical Weibull statistics $N_N \equiv V/V_0$ for volume dominating defects (or $N_N = A/A_0$ for surface dominating defects), i.e., N_N is the ratio between the volume (or surface) of the structure and a reference volume (or surface). The experimental data on CNT [4, 5] were treated [34] according to nanoscale and classical Weibull statistics, Fig. 14.1: the coefficients of correlation were found to be much higher for the nanoscale statistics than for the classical one (0.93 against 0.67, $w \approx 2.7$ and $\sigma_0 \approx 31 - 34$ GPa). Other data set on multi walled CNT tensile experiments [31, 35] are also treated in Fig. 14.1 [21].

Note that volume- or surface-based Weibull statistics are identical in treating the external wall of the tested nanotubes, just an atomic layer thick.

All these experimental data [4,5,31,35] are treated in Table 14.2, by applying QFM in the form of (14.13): non-linear multiple solutions for identifying

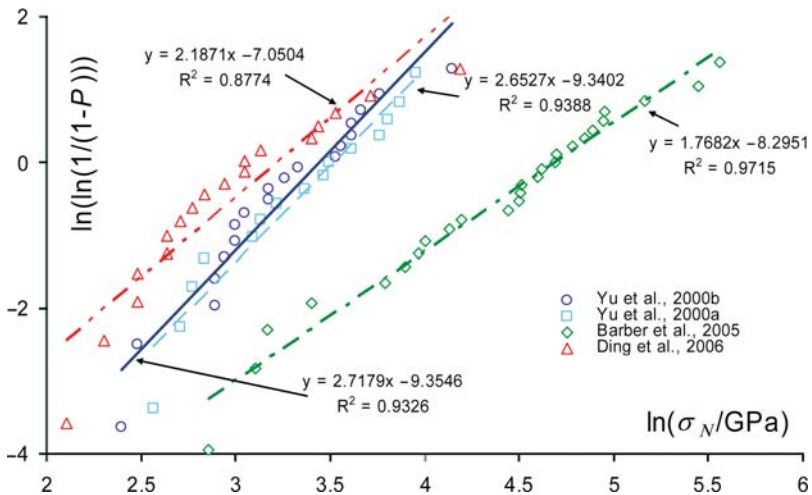


Fig. 14.1. Nanoscale Weibull Statistics, *straight lines*, applied to the new nanotensile experiments on carbon nanotubes

Table 14.2. Experiments vs. QFM predictions; strength reduction $\sigma_N(a, b) / \sigma_N^{(\text{theo})}$ derived according to (14.13)

$\sigma_N / \sigma_N^{(\text{theo})}$	$2a/q$		$2b/q$		0	1	2	3	4	5	6	7	8	9	10	∞
0	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00*	1.00
1	0.71*	0.75	0.79	0.82	0.85	0.87*	0.88*	0.90	0.91	0.91	0.91	0.91	0.91	0.91	0.92	1.00
2	0.58	0.60*	0.64*	0.68	0.71*	0.73	0.76	0.78*	0.79	0.81	0.82	1.00				
3	0.50	0.52	0.54*	0.58	0.61	0.64*	<u>0.66*</u>	0.68	0.70*	0.72	0.74	1.00				
4	0.45	0.46	0.48	0.51*	0.54*	0.56	0.59	0.61	<i>0.63</i>	0.65	0.67	1.00				
5	<u>0.41</u>	0.42	0.44*	0.46	0.48	0.51*	0.53*	0.55*	0.58	0.59	0.61	1.00				
6	0.38	0.38	0.40	0.42	0.44*	0.47	0.49*	0.51*	0.53*	0.55*	0.57	1.00				
7	<i>0.35</i>	0.36	0.37	<i>0.39</i>	<u>0.41</u>	0.43	0.45	0.47	0.49*	0.51*	0.53*	1.00				
8	0.33	<u>0.34</u>	<i>0.35</i>	0.37	0.38	0.40	0.42	0.44*	0.46	0.48	0.49*	1.00				
9	0.32	0.32	0.33	<u>0.34</u>	0.36	0.38	0.40	<u>0.41</u>	0.43	0.45	0.46	1.00				
10	<u>0.30*</u>	<u>0.30*</u>	<u>0.31</u>	0.33	<u>0.34</u>	0.36	0.37	<u>0.39</u>	<u>0.41</u>	0.42	0.44*	1.00				
11	0.29	0.29	<u>0.30*</u>	<u>0.31</u>	0.32	<u>0.34</u>	<i>0.35</i>	0.37	<u>0.39</u>	0.40	0.42	1.00				
12	<i>0.28</i>	<i>0.28</i>	0.29	<u>0.30*</u>	<u>0.31</u>	0.32	<u>0.34</u>	<i>0.35</i>	0.37	0.38	0.40	1.00				
13	0.27	0.27	<i>0.28</i>	0.29	<u>0.30*</u>	<u>0.31</u>	0.32	<u>0.34</u>	<i>0.35</i>	0.36	0.38	1.00				
14	<i>0.26</i>	<i>0.26</i>	0.27	0.27	0.29	<u>0.30*</u>	<u>0.31</u>	0.32	<u>0.34</u>	<i>0.35</i>	0.36	1.00				
15	0.25	0.25	<i>0.26</i>	0.27	0.27	0.29	<u>0.30*</u>	<u>0.31</u>	0.32	<u>0.34</u>	<i>0.35</i>	1.00				
16	0.24*	0.24*	0.25	<i>0.26</i>	0.27	<i>0.28</i>	0.29	<u>0.30*</u>	<u>0.31</u>	0.32	0.33	1.00				
17	0.24*	0.24*	0.24*	0.25	<i>0.26</i>	0.27	<i>0.28</i>	0.29	<u>0.30*</u>	<u>0.31</u>	0.32	1.00				
18	0.23	0.23	0.24*	0.24*	0.25	<i>0.26</i>	0.27	<i>0.28</i>	0.29	<u>0.30*</u>	<u>0.31</u>	1.00				
19	0.22*	0.22*	0.23	0.23	0.24*	0.25	<i>0.26</i>	0.27	<i>0.28</i>	0.29	<u>0.30*</u>	1.00				
20	0.22*	0.22*	0.22*	0.23	0.24*	0.24*	0.25	0.26	0.27	0.28	0.29	1.00				
21	<u>0.21</u>	<u>0.21</u>	0.22*	0.22*	0.23	0.24*	0.25	0.25	<i>0.26</i>	0.27	<i>0.28</i>	1.00				
22	<u>0.21</u>	<u>0.21</u>	<u>0.21</u>	0.22*	0.22*	0.23	0.24*	0.25	<i>0.26</i>	0.27	<i>0.28</i>	1.00				
23	<i>0.20</i>	<u>0.21</u>	<u>0.21</u>	<u>0.21</u>	0.22*	0.23	0.23	0.24*	0.25	<i>0.26</i>	0.27	1.00				
24	<i>0.20</i>	<i>0.20</i>	<i>0.20</i>	<u>0.21</u>	<u>0.21</u>	0.22*	0.23	0.24*	0.24*	0.25	<i>0.26</i>	1.00				
25	<i>0.20</i>	<i>0.20</i>	<i>0.20</i>	<i>0.20</i>	<u>0.21</u>	0.22*	0.22*	0.23	0.24*	0.25	<i>0.26</i>	1.00				
26	<u>0.19</u>	<u>0.19</u>	<i>0.20</i>	<i>0.20</i>	<i>0.20</i>	<u>0.21</u>	0.22*	0.22*	0.23	0.24*	0.25	1.00				
27	<u>0.19</u>	<u>0.19</u>	<u>0.19</u>	<i>0.20</i>	<i>0.20</i>	<u>0.21</u>	<u>0.21</u>	0.22*	0.23	0.24*	0.24*	1.00				
28	<u>0.19</u>	<u>0.19</u>	<u>0.19</u>	<u>0.19</u>	<i>0.20</i>	<i>0.20</i>	<u>0.21</u>	0.22*	0.22*	0.23	0.24*	1.00				
29	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	<u>0.19</u>	<i>0.20</i>	<i>0.20</i>	<u>0.21</u>	0.22*	0.23	0.23	1.00				
30	<i>0.18</i>	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	<u>0.19</u>	0.20	<u>0.21</u>	<u>0.21</u>	0.22*	0.23	1.00				
31	<i>0.18</i>	<i>0.18</i>	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	0.20	<u>0.21</u>	0.22*	0.22*	1.00				
32	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	<u>0.21</u>	<u>0.21</u>	0.22*	1.00				
33	0.17*	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	0.20	<u>0.21</u>	<u>0.21</u>	1.00				
34	0.17*	0.17*	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	0.20	<u>0.21</u>	1.00				
35	0.17*	0.17*	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	0.20	<u>0.21</u>	1.00				
36	0.16	0.16	0.17*	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	<u>0.20</u>	1.00				
37	0.16	0.16	0.16	0.17*	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	1.00				
38	0.16	0.16	0.16	0.16	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	0.20	1.00				
39	0.16	0.16	0.16	0.16	0.17*	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	1.00				
40	0.16	0.16	0.16	0.16	0.16	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	<u>0.19</u>	1.00				
41	0.15	0.15	0.16	0.16	0.16	0.16	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	1.00				
42	0.15	0.15	0.15	0.16	0.16	0.16	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	<u>0.19</u>	1.00				
43	0.15	0.15	0.15	0.15	0.16	0.16	0.16	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	1.00				
44	0.15	0.15	0.15	0.15	0.16	0.16	0.16	0.17*	0.17*	<i>0.18</i>	<i>0.18</i>	1.00				
45	0.15	0.15	0.15	0.15	0.15	0.16	0.16	0.16	0.17*	0.17*	0.17*	1.00				

(Continued)

Table 14.2. (Continued)

46	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.16</u>	<u>0.16</u>	<u>0.17*</u>	<u>0.17*</u>	<i>0.18</i>	1.00
47	<u>0.14</u>	<u>0.14</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.16</u>	<u>0.16</u>	<u>0.16</u>	<u>0.17*</u>	<u>0.17*</u>	1.00
48	<u>0.14</u>	<u>0.14</u>	<u>0.14</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.16</u>	<u>0.16</u>	<u>0.17*</u>	<u>0.17*</u>		1.00
49	<u>0.14</u>	<u>0.14</u>	<u>0.14</u>	<u>0.14</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.16</u>	<u>0.16</u>	<u>0.16</u>	<u>0.17*</u>		1.00
50	<u>0.14</u>	<u>0.14</u>	<u>0.14</u>	<u>0.14</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.15</u>	<u>0.16</u>	<u>0.16</u>	<u>0.17*</u>		1.00
∞	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	$(1 + 2a/b)^{-1}$

In **bold** type are represented the 15 different nanostrengths measured on single walled CNT in bundle [4]; whereas in *italic* we report the 19 nanostrengths measured on multi walled CNT[5], and in underlined type the most recent 18 observations [31]. All the data are reported with the exception of the five smallest values of 0.08, 0.10 [31], 0.11 [5], 0.12 [5,31] and 0.13 [4], for which we would need for example adjacent vacancies ($2b/q \sim 1$) in number $n = 2a/q = 138-176, 90-109, 75-89, 64-74$ and $55-63$ respectively. The 26 strengths measured in [35] are also treated (*asterisks*), simply assuming two interacting walls for $100 < \sigma_N^{(exp)} \leq 200$ gigapascals (thus $\sigma_N = \sigma_N^{(exp)} / 2$) or 3 interacting walls for $200 < \sigma_N^{(exp)} \leq 300$ gigapascals ($\sigma_N = \sigma_N^{(exp)} / 3$). All the experiments are referred to $\sigma_N^{(theo)} = 100\text{GPa}$ ($q \sim 0.25 \text{ nm}$). If all the nanotubes in the cable contain identical holes, $\sigma_C / \sigma_C^{(theo)} = \sigma_N / \sigma_N^{(theo)}$

the defects corresponding to the measured strength clearly emerge; however these are quantifiable, showing that a small defect is sufficient to rationalize the majority of the observed strong strength reductions.

Finally, the anomalous (due to plasticity) experimental results [31] are differently treated in Table 14.3, with respect to both strength and elasticity, assuming the presence of transversal nanocracks. The ideal strength is assumed to be of 100GPa and the theoretical Young’s modulus of 1.3 TPa; by (14.13) the crack length n is calculated and introduced in (14.19) to derive the related vacancy fraction v ($\xi = \pi / 2$).

Fracture in two cases was observed at the clamp; in one case the clamp itself failed, thus the deduced strength represents a lower bound of the nanotube strength. Three nanotubes were multiple loaded (in two a,b and A,B or in three I,II,III steps), i.e., after the breaking in two pieces of a nanotube, one of the two pieces was again tested and fractured at a higher stress. Two nanotubes displayed a plastic flow.

A vacancy fraction of the order of few $\%_{00}$ is estimated, suggesting that such nanotubes are much more defective than as imposed by the thermodynamic equilibrium, even if the defects are small and isolated. However, note that other interpretations are still possible, e.g., assuming the nanotube is coated by an oxide layer and rationalizing the ratio between the observed Young’s modulus and its theoretical value as the volumetric fraction (for softer coating layers) of carbon in the composite structure.

Table 14.3. The experiments [31] are here treated with respect to both strength and elasticity, assuming the presence of transversal nanocracks composed by n adjacent vacancies

MWCNT number and fracture typology	Strength [GPa]	Young's modulus [GPa]	κ	n	$v\%$
1 (multiple load A)	8.2	1,100	1.01	148	0.07
2 (clamp failed)	10	840	0.98	100	0.23
3	12	680	1.00	69	0.44
4 (failure at the clamp)	12	730	0.98	69	0.40
5 (multiple load B)	14	1,150	1.02	51	0.14
6 (multiple load a)	14	650	0.97	51	0.62
7	15	1,200	1.05	44	0.11
8	16	1,200	1.02	39	0.13
9	17	960	1.00	34	0.49
10	19	890	0.97	27	0.74
11 (multiple load b)	21	620	0.99	22	1.51
12 (multiple load I)	21	1,200	0.99	22	0.22
13 (multiple load II)	23	1,250	0.99	18	0.17
14	30	870	1.00	11	1.92
15 (plasticity observed)	31	1,200	0.59 (0.99)	10	0.49
16 (plasticity observed)	34	680	0.69 (1.02)	8	3.80
17 (multiple load III)	41	1,230	1.03	5	0.69
18 (failure at the clamp)	66	1,100	0.98	2	4.90

The constitutive parameter κ has been estimated as $\kappa \approx \ln(\varepsilon_N)/\ln(\sigma_N/E)$ for all the tests: note the low values for the two nanotubes that revealed plasticity (in brackets the values calculated up to the incipient plastic flow are also reported). The ideal strength is assumed to be of 100GPa and the theoretical Young's modulus of 1,300 GPa; by (14.13) the crack length n is calculated and introduced in (14.19) to derive the related vacancy fraction v ($\xi = \pi/2$)

14.10 Thermodynamic Limit

Defects are thermodynamically unavoidable, especially at the macroscale. At the thermal equilibrium the vacancy fraction $f = n/N \ll 1$, n is the number of vacancies and N is here the total number of atoms, is estimated as [36]:

$$f \approx e^{-E_1/(k_B T_a)}, \quad (14.21)$$

where $E_1 \approx 7$ eV is the energy required to remove one carbon atom and T_a is the absolute temperature at which the carbon is assembled, typically in the range between 2,000 and 4,000 K. Thus, $f \approx 2.4 \times 10^{-18} - 1.6 \times 10^{-9}$.

The strength of a bundle will be dictated by the largest transversal crack on it, according to the weakest link concept. The probability of finding a nanocrack of size m in a bundle with vacancy fraction f is $P(m) = (1-f)f^m$, and thus the number M of such nanocracks in a bundle composed by N atoms is $M(m) = P(m)N$. The size of the largest nanocrack, which typically occurs

once, is found from the solution to the equation $M(m) \approx 1$, which implies [37]:

$$m \approx -\ln[(1-f)N] / \ln f \approx -\ln N / \ln f \tag{14.22}$$

Inserting (14.21) and (14.22) into (14.13) evaluated for a transversal crack ($b \approx 0$ and $2a/q \approx m$), we deduce the statistical counterpart of (14.13) and thus the following thermodynamical maximum achievable strength [38]:

$$\frac{\sigma_N(N)}{\sigma_N^{(\text{theo})}} \leq \frac{\sigma_N^{(\text{max})}(N)}{\sigma_N^{(\text{theo})}} = \frac{1}{\sqrt{1 + \frac{k_B T_a}{E_1} \ln N}} \tag{14.23}$$

14.11 Hierarchical Simulations and Size Effects: from a Nanotube to a Megacable

To evaluate the strength of carbon nanotube cables, the SE³ algorithm, formerly proposed [39], has been adopted [40]. Multiscale simulations are necessary in order to tackle the size scales from a nanotube to a bundle. Let us consider the limit case of a nanotube based megacable, such as that of the space elevator [21, 38–40], Fig. 14.2.

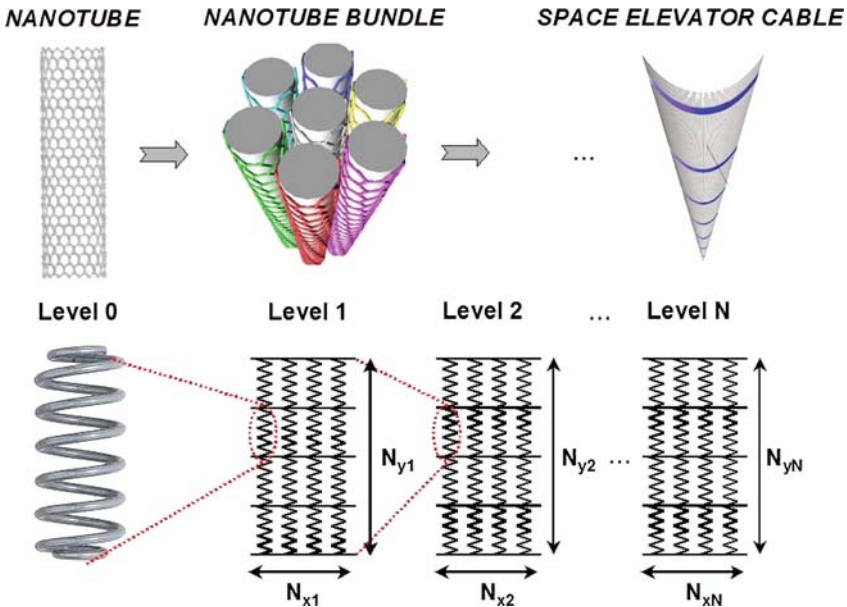


Fig. 14.2. Schematization of the adopted multiscale simulation procedure to determine the nanotube cable strength, from a nanotube to a megacable, here considered to be the limiting case of the space elevator cable. Here, $N = 5$, $N_{x1} = N_{x2} = \dots N_{x5} = 40$ and $N_{y1} = N_{y2} = \dots N_{y5} = 1,000$, so that the total number of nanotubes in the largest considered cable is $N_{\text{tot}} = (1,000 \times 40)^5 \approx 10^{23}$

The bundle is modelled as an ensemble of stochastic “springs”, arranged in parallel sections. Linearly increasing strains are applied to the fibre bundle, and at each algorithm iteration the number of fractured springs is computed (fracture occurs when local stress exceeds the nanotube failure strength) and the strain is uniformly redistributed among the remaining intact springs in each section.

In-silico stress-strain experiments have been carried out according to the following hierarchical architecture. Level 1: the nanotubes (single springs, Level 0) are considered with a given elastic modulus and failure strength distribution and composing a $40 \times 1,000$ lattice or fibre. Level 2: again a $40 \times 1,000$ lattice composed by second level “springs”, each of them identical to the entire fibre analysed at the first level, is analysed with in input the elastic modulus and stochastic strength distribution derived as the output of the numerous simulations to be carried out at the first level. And so on. Five hierarchical levels are sufficient to reach the size-scale of the megametre from that of the nanometre, Fig. 14.2.

The level 1 simulation is carried out with springs $L_0 = 10^{-7}$ m in length, $w_0 = 10^{-9}$ m in width, with Young’s modulus $E_0 = 10^{12}$ Pa and strength σ_f randomly distributed according to NWS [34] fitting to experiments [4,5], thus assuming $\sigma_0 = 34$ GPa and $m = 2.7$. Then the level 2 is computed, and so on. The results are summarized in Fig. 14.3, in which a strong size effect is observed, up to length of ~ 1 m.

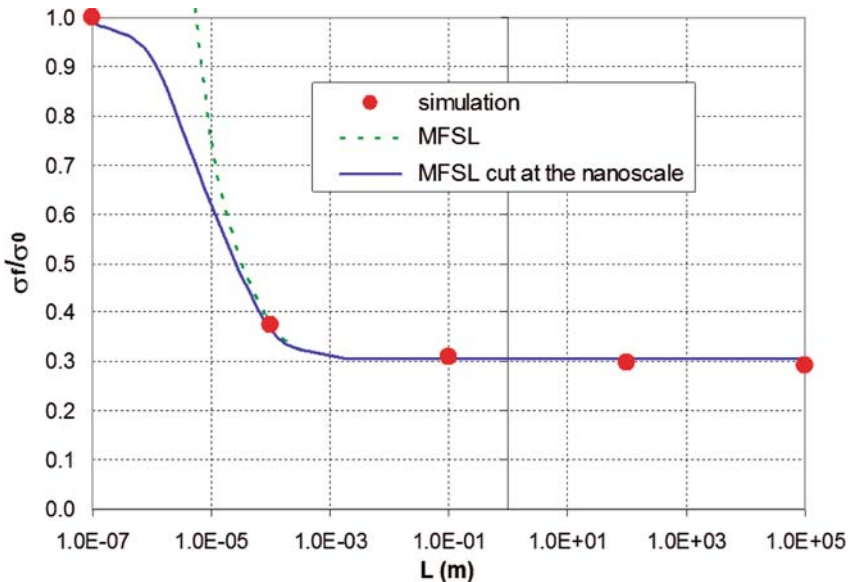


Fig. 14.3. Comparison between simulations and analytical scaling law, (14.24), for the failure strength of the nanotube bundle as a function of its length; the asymptote is at 10.20 GPa

Given the decaying σ_f vs. cable length L obtained from simulations, it is interesting to fit the behaviour with analytical scaling laws. Various exist in the literature, and one of the most used is the Multi Fractal Scaling Law (MFSL [41, 42]). This law has been recently extended towards the nanoscale [43], deriving (according to the theory of the geometrically necessary dislocations):

$$\frac{\sigma_f}{\sigma_{\text{macro}}} = \sqrt{1 + \frac{l_{\text{ch}}}{L + l_0}}, \quad (14.24)$$

where σ_f is the failure stress, σ_{macro} is the macrostrength, L is the structural characteristic size, l_{ch} is a characteristic internal length and l_0 is defined via $\sigma_f(l=0) = \sigma_{\text{macro}} \sqrt{1 + \frac{l_{\text{ch}}}{l_0}} \equiv \sigma_{\text{nano}}$, where σ_{nano} is the nanostrength. Note that for $l_0 = 0$ this law is identical to the MFSL [41, 42]. Here, we can choose σ_{nano} as the nanotube stochastic strength, i.e. $\sigma_{\text{nano}} = 34$ GPa. The computed macrostrength is $\sigma_{\text{macro}} = 10.20$ GPa. The fit with (14.1) is shown in Fig. 14.3 (“MFSL cut at the nanoscale”), for the various L considered at the different hierarchical levels (and compared with the classical “MFSL”). The best fit is obtained for $l_{\text{ch}} = 5 \times 10^{-5}$ m, where the analytical law is practically coincident with the simulated results. Thus, for a carbon nanotube bundle not shorter than a millimetre, the expected strength is $\sigma_C = \sigma_{\text{macro}} \approx 10 - 30$ GPa, where the upper limit would correspond to assume a theoretical value of $\sigma_{\text{nano}} = 100$ GPa.

14.12 Conclusions

The strength and fatigue life of a real, thus defective, carbon nanotube and bundle are expected to be strongly reduced with respect to their theoretical values. Less dramatic, but still significant, could be the reduction imposed by defects on their elasticity. Accordingly, we conclude that a proper and novel design of ultra-sharp tips based on CNT has to take into account the role of defects. QFM approaches are ideal to design flaw-tolerant nanomaterials and nanostructures, such as the discussed CNT tips for scanning probe “nanoscopy”.

References

1. H. Dai, J.H. Hafner, A.G. Rinzler, D.T. Colbert, R.E. Smalley *Nature* **384** 147 (1996)
2. S.S. Wong, A.T. Woolley, T.W. Odom, J.L. Huang, P. Kim, D.V. Vezenov, C.M. Lieber *Appl. Phys. Lett.* **73** 3465 (1998)
3. C.V. Nguyen Q.i. Ye, M. Meyyappan *Meas. Sci. Technol.* **16** 2138 (2005)
4. M.F. Yu, B.S. Files, S. Arepalli, R. Ruoff *Phys. Rev. Lett.* **84** 5552 (2000)
5. M.F. Yu, O. Lourie, M.J. Dyer, K. Moloni, T.F. Kelly R. Ruoff *Science* **287** 637 (2000)

6. N. Pugno *A quantized Griffith's Criterion, Fracture Nanomechanics, Meeting of the Italian Group of Fracture* September 25–26, 2002, Vigevano, Italy
7. N. Pugno, R. Ruoff *Philos. Mag.* **84** 2829 (2004)
8. N. Pugno *Int. J. Fract.* **140** 158 (2006)
9. N. Pugno *Int. J. Fract.* **141** 311 (2006)
10. A.A. Griffith *Phil. Trans. R. Soc. Lond. Biol. Sci* **A221** 163 (1921)
11. L.B. Freund *Dynamic Fracture Mechanics* (Cambridge University Press, Cambridge, 1990)
12. H. Neuber *Theory of Notch Stresses* (Springer, Berlin, 1958)
13. V. Novozhilov *Prikl. Mat. Mek.* **33** 212 (1969)
14. N.F. Morozov, V. Yu Petrov, A.A. Utkin *Sov. Phys. Dokl.* **35** 646 (1990)
15. P. Cornetti, N. Pugno, A. Carpinteri, D. Taylor *Eng. Fract. Mech.* **73** 2021 (2006)
16. D. Taylor, P. Cornetti, N. Pugno *Eng. Fract. Mech.* **72** 1021 (2005)
17. H. Murakami *Stress Intensity Factors Handbook* (Pergamon Press, Oxford, 1986)
18. N. Pugno, P. Cornetti, A. Carpinteri *Eng. Fract. Mech.* **74** 595 (2007)
19. N. Pugno, M. Ciavarella, P. Cornetti, A. Carpinteri *J. Mech. Phys. Solids* **54** 1333 (2006)
20. W. Weibull *Handlingar*, Nr 151 (1939)
21. N. Pugno *Acta Mater.* **55** 5269 (2007)
22. J.R. Rice, G.F. Rosengren *J. Mech. Phys. Solids* **16** 1 (1968)
23. A. Carpinteri, N. Pugno *Eng. Fract. Mech.* **72** 1254 (2005)
24. A. Carpinteri, B. Chiaia *Int. J. Fract.* **76** 327 (1996)
25. N. Pugno *Appl. Phys. Lett.* **90** 043106 (2007)
26. S.L. Mielke, D. Troya, S. Zhang, J.L. Li, S. Xiao, R. Car, R. Ruoff, S., G.C. Schatz, T. Belytschko *Chem. Phys. Lett.* **390** 413 (2004)
27. T. Belytschko, S.P. Xiao, R. Ruoff *Effects of defects on the strength of nanotubes: experimental-computational comparisons*, Los Alamos National Laboratory, Preprint Archive, Physics, arXiv:physics/0205090 (2002)
28. S. Zhang, S.L. Mielke, R. Khare, D. Troya, R.S. Ruoff, G.C. Schatz, T. Belytschko *Phys. Rev. B* **71** 115403 1 (2005)
29. R. Khare, S.L. Mielke, J.T. Paci, S. Zhang, R. Ballarini, G.C. Schatz, T. Belytschko *Phys. Rev. B* **75** 075412 (2007)
30. M. Ippolito, A. Mattoni, L. Colombo, N. Pugno *Phys. Rev. B* **73** 104111–1/6 (2006)
31. W. Ding, L. Calabri, K.M. Kohlhaas, X. Chen, D.A. Dikin, R.S. Ruoff *Modulus Exp. Mech.* **47** 25 (2006)
32. M. Meo, M. Rossi *Eng. Fract. Mech.* **73** 2589 (2006)
33. M. Sammalkorpi, A. Krasheninnikov, A. Kuronen, K. Nordlund, K. Kaski *Phys. Rev. B* **70** 245416–1/8 (2004)
34. N. Pugno, R. Ruoff *J. Appl. Phys.* **99** 1 (2006)
35. A.H. Barber, I. Kaplan-Ashiri, S.R. Cohen, R. Tenne, H.D. Wagner *Compos. Sci. Technol.* **65** 2380 (2005)
36. C. Kittel *Introduction to Solid State Physics* (Wiley, New York, 1966)
37. P.D. Beale, D.J. Srolovitz *Phys. Rev. B* **37** 5500 (1988)
38. N. Pugno *Nano Today* **2** 44 (2007)
39. N. Pugno *J. Phys. Condens. Matter* **18** S1971 (2006)
40. N. Pugno, F. Bosia, A. Carpinteri *Small* **4** 1044 (2008)
41. A. Carpinteri *Int. J. Solid Struct.* **31** 291 (1994)
42. A. Carpinteri *Mech. Mater.* **18** 89 (1994)
43. N. Pugno *Acta Mater.* **55** 1947 (2007)

Spin and Charge Pairing Instabilities in Nanoclusters and Nanomaterials

Armen N. Kocharian, Gayanath W. Fernando, and Chi Yang

Summary. Our exact studies are focused on the understanding of pairing and magnetism arising from local electron correlations in tunneling probe responses in nanoclusters, assembled clusters of various geometries, nano and heterostructured materials, atomic scale experiments in inhomogeneous high T_c cuprates, manganites, and other concentrated transition metal oxides. Interpretation of many body physics and local density states anomalies in scanning tunneling spectroscopy is given for electron charge and spin pairing instabilities in real space. Near crossing degeneracies or quantum critical points the electron configurations of the lowest energy levels control the low temperature physics of spontaneous transitions, phase separation, spatial inhomogeneities, and pairing that can be the important steps in deciphering the mystery of high T_c superconductivity. The thermodynamic phase diagrams, obtained from exact canonical and grand canonical ensemble calculations in small clusters resemble a number of inhomogeneities, coherent and incoherent, paired and unpaired nanoscale phases found recently in Nb and Co nanoparticles, ultracold fermionic atoms, doped high T_c cuprates, spin-charge separation in manganites, and multiferroic nanomaterials probed by scanning tunneling microscopy. We review experimental and theoretical descriptions of electron instabilities and intrinsic inhomogeneities in an ensemble of clusters, ultrasmall nanoparticles, and eventually bulk nanomaterials, which require modifications of the standard theories to implement the local electron correlations in the local building block structure to exhibit energy level discreteness and spatial inhomogeneities in novel nonstoichiometric nanomaterials.

Key words: High T_c superconductivity, Coherent and incoherent pairing, Charge and spin pseudogaps, Spin-charge separation, Quantum critical points, Phase separation, Magnetism

15.1 From Atoms to Solids

Interest in the physics of condensed matter at various scales has grown rapidly, due to the increasing realization of new physical properties and inherent collective phenomena found in small and large systems. The ensembles

of nanoparticles and nanoclusters still reveal important features known for inhomogeneous bulk nanostructures, different from those named as conventional solid materials. The application of this concept provides a link between cluster, nuclear, and condensed-matter physics. A detailed study of a quantum gas of clusters can be used to bridge the gap of knowledge by clustering atoms together and studying their condensed cooperative characteristics. For example, the nucleation and the onset of a phase transition in small clusters of metal-rich titanium selenides turned out to be a playground for condensed cluster phases [1]. The study of atomic and molecular clusters also benefits the developing field of nanotechnology. If new materials are to be made out of nanoscale particles (for possible use in quantum computation and entanglement), the properties of the nanoparticles must first be understood on a single atom, molecule and cluster levels [2]. The properties of bulk elements formed from atoms, molecules and clusters can be divided into two categories: (a) a stable regime wherein properties vary smoothly until they reach the bulk limit and (b) unstable regime wherein the variation of properties is highly non-monotonic. In this latter case, unusual things can and often do arise because of quantum confinement, electron doping and interaction, intra-site couplings, and boundary conditions. The stability of clusters can be altered by changing the number of electrons and/or atoms. Furthermore, it was predicted [3] and later verified [4] that the dominant channel in the fragmentation of a cluster, whether it is neutral or charged, will usually involve a magic cluster. Evidence was found for Coulomb explosion of doubly charged microclusters in [5]. This process is analogous to nuclear fission where magic nuclei appeared in the fission process. For cluster sizes above n_c , however, evaporation of a neutral monomer is the dominant channel for delayed dissociation.

Under certain circumstances, the charged (doped) clusters of nonmagnetic elements become magnetic, semiconducting materials exhibit metallic properties, metallic systems become semiconducting or superconducting, the nonmagnetic systems become magnetic with size or electron/hole doping, etc. These properties arise because of the unusual structure of the clusters and because the electrons belong to molecular orbitals and exhibit an energy excitation gap, referred to as the gap between highest occupied and lowest unoccupied molecular orbital (HOMO-LUMO). Often the bonding orbitals are filled and separated from the vacant antibonding ones by a significant HOMO-LUMO gap. On the other hand, the HOMO-LUMO degeneracy can lead to electron instabilities which can be driven by local Coulomb interaction, Jahn-Teller effect, Peierls distortion and inter-site coupling. For many systems this gap serves as a useful guide to favored geometry. The magnitude of the HOMO-LUMO gap varies with size and composition of the cluster, electron interaction, boundaries and the way these orbitals are filled determines not only the energy stability of the clusters but also their magnetic and electronic properties. A systematic study of the structure, which includes electron instabilities, and properties of clusters composed of a variety of elements has bridged many fields of physics, particularly atomic, molecular,

nuclear, and condensed-matter physics. Electron correlations and changes of electron count with composition in transition-metal clusters can have clear implications in nanochemistry and fabrication of novel nanomaterials. The metal/insulator, superconducting/normal, magnetic/nonmagnetic transitions are a few typical examples of cooperative phenomena, which often develop in an aggregate of a large number of atoms in bulk materials. Generally speaking, physical properties and excitation gaps of clusters are strongly correlated to their geometrical and electronic structures. It is commonly believed that there is a critical length parameter (size) for small clusters beyond which a ferromagnetic (ferroelectric) material is considered to be super-paramagnetic rather than ferromagnetic (ferroelectric). Earlier, it was thought that many body phenomena break down for very small cluster sizes. The question then was asked, for example: How many atoms do we need in order to obtain the collective metallic, superconducting, or magnetic properties of a solid? Soon after the first cluster sources had been developed in the early eighties, it led to the discovery of ferromagnetism and superconductivity in small, finite-size fullerenes and carbon nanotubes [6]. An electronic dipole moment was also found in smaller Nb_N ($N \leq 28$) exhibiting a pronounced even-odd oscillation of the dipole moment values. Size evolution and the increase of binding energy at magic small numbers such as 4, 7, 10, 15, and 22 [7] lead to enhancement of the dipole moment. More importantly, it is found that the ferroelectricity and ferromagnetism can coexist in the transition metal inhomogeneous materials composed of the clusters. These novel materials in which magnetic and electric orders coexist, termed “multiferroics” or “magneto-electrics,” [8,9], offer another possibility to prepare new bulk materials with both ferromagnetic and ferroelectric orders. Finally, a *quantum* macroscopic phenomenon in solids, such as superconductivity, is in fact an electron instability that can also exist locally in complex organic molecules, carbon nanotubes (CNs) and nanoclusters. The mechanism of high T_c superconductivity also can be controlled by local electron structure and correlations in building blocks on nanoscale (see Sect. 15.7.7). Thus, the studies of superconductivity can bridge the existing gap between the micro and macro world. Nanomagnetism, nanoelectricity or the understanding and design of complex multicomponent magnetic and superconducting nanostructures, is one of the current frontiers in condensed matter physics. Spin-polarized tunneling electrons give rise to a powerful technique which uses a real space imaging on the atomic scale [10] for the understanding of complicated magnetic configurations of nanomagnets and corresponding magnetic superstructures and heterostructures.

15.1.1 Discreteness of Spectrum

One of the most fundamental features of quantum mechanics is the discrete or quantized energy spectrum of a system of particles confined to a small region with the typical spacing between energy levels increasing with decreasing system size. In atomic and nuclear physics, spectroscopic techniques for

measuring and analyzing discrete spectra have for decades been a major source of detailed information on the correlations between the particles. In condensed matter physics, however, it has been much more difficult to study spectroscopically the discrete spectrum of an individual sample, since system sizes were typically so large that discrete eigenenergies could not be resolved on the energy scale set by the temperature and boundaries. This changed in the course of the last two decades due to advances in microfabrication techniques, which made it possible to study individual systems of mesoscopic or nanoscopic dimensions, especially with discovery of new nano and heterostructured materials with strong inhomogeneities, where the characteristic length scales range from a few micrometers down to few nanometers. If metallic clusters contain a small number of electrons, one might think that they do not display superconducting properties, because the average level spacing greatly exceeds the pairing energy gap. However, the situation here is more complicated. As a matter of fact, there exist clusters in which the pattern of electronic states is very different from that of a simple equally spaced level distribution. They contain highly degenerate electronic levels (Sect. 15.7.3), or groups of very close energy levels that can be beneficial for electron pairing. There is a common belief that electron repulsion generally increases level separation in atomic scale, which prevents electron pairing and superconductivity in finite clusters [11]. This is certainly true at half filling or for even number of electrons in bipartite geometries at strong Coulomb interaction, wherein electron localization and level separation suppresses electron pairing. However, in some frustrated cluster geometries, in spite of strong electron correlations, we found that in contrast the density of states near certain odd or even integer number of electrons increases due to the high degeneracy of energy levels. This factor can lead, at suitable model parameters and cluster geometry, to electron pairing, pairs (mobility) metallization and possible superconductivity (Sect. 15.7.7).

15.1.2 Electron Spectroscopy

In the mid-nineties, a similar advance was achieved with metals, when Ralph, Black, and Tinkham [12,13] succeeded in performing single-electron-tunneling spectroscopy on individual ultrasmall metallic grains. These experiments have shown for the first time that mean-field BCS theory is actually invalid to apply to nanoscale and exact (finite-size) study is necessary to conduct a reliable analysis. The limitations of the BCS pairing are also well understood for description of odd-even effects and binding energies, when applied to nuclear matter. This opened up a new frontier in the study of electron correlations in metals, since accurately resolved discrete energy levels allows the nature of electron correlations to be studied in unprecedented detail. During the last several years, single-electron tunneling spectroscopy of ultrasmall metallic grains has been used to probe superconducting pairing correlations in Al grains and ferromagnetic correlations in Co grains [14,15]. Under certain

conditions, such clusters should display a drastic increase in the values of superconducting parameters such as the critical temperature and the energy gap [16, 17]. The advent of beams of atom clusters with selected sizes allowed the physics and chemistry of these confined ensembles to be critically explored, leading to an increased understanding of their potential, particularly as the constituents of new materials, including metals, ceramics, and composites of these materials. In addition, the effects of confinement on interfaces also play an important and, sometimes, dominant role in cluster-assembled nanophase materials, as well as in nanostructured multilayers. The smooth evolution from strong ferromagnetism for localized moments in small cobalt clusters to itinerant magnetism in the bulk appears to be related to the increase in size. The ferroelectric state in niobium clusters can be responsible for electron pairing, which can be beneficial to superconductivity observed in the bulk materials [18, 19].

15.1.3 Electron Correlations in Clusters

There were also many earlier attempts [20] to get the complete spectrum, and perform accurate calculations of thermodynamics of interacting electrons in exactly solvable models and exact diagonalization using the compelling simplicity of finite-size Hubbard clusters [21–26]. In condensed matter, this model was introduced to describe the behavior of narrow-band phenomena, with the main emphasis given to the explanation of the metal-insulator transition, anti-ferromagnetism, and ferromagnetism. The Hubbard-like lattice models with local interactions are intrinsically different from models in continuum space because the Galilean invariance is explicitly broken. Although the Hubbard model is oversimplified, exact homogeneous solution for coupled nonlinear equations was found only in the one-dimensional Hubbard model with and without a magnetic field [27–37]. For higher dimensions, the study is far from being completed, and only a few rigorous results are known [38]. Therefore, one can try to get deep insight into the system by solving the problem in special cases exactly. Then these rigorous results in two- and three-dimensional cluster geometries can serve as tests to the other results derived by approximate methods.

The dynamic mean-field theory (DMFT) has been proven successfully to exhibit the Mott-Hubbard transition and dual (local and band) nature of the electronic states at half filling, in spite its inability to describe away from half filling the atomic limit correctly. The DMFT is exact only in the limit of large lattice dimensionality or large connectivity. In addition, spatial fluctuations included in the DMFT self-energy correction seems have no sufficient influences on the one-particle spectrum, which is a serious problem for studies of inhomogeneous systems especially in the vicinity of the quantum critical points. Among other analytical methods, such as slave bosons, and random phase approximation (RPA), surprisingly insufficient attention was given to

a detailed analysis of local electron correlations in mean-field theory. On the other hand, the generalized mean-field approximations allows to formulate a problem in a physically transparent way and get an analytical solution in any spatial dimension [39, 40]. For the case of one dimension, an analytical generalized self-consistent field (GSCF) theory can be compared for repulsive ($U > 0$) and attractive ($U < 0$) interactions with the exact solution [41–44]. This analytical approach provides an inhomogeneous solution and also allows to construct a more accurate converging perturbation theory about the self-consistent field solution [45]. The techniques, by far, most intensively used for numerical treatment of finite systems are the Lanczos algorithm [46] and Quantum Monte Carlo (QMC) methods [47]. However, the Lanczos method has poor convergence, while QMC contains severe limitations in the presence of magnetic field or away from half filling due to the notorious sign problem. On the other hand, the exact numerical calculations require a considerable numerical effort, which grow exponentially with the system size. The studies of thermodynamic properties are considerably simplified, if all the eigenvalues of a Hamiltonian are calculated numerically, using exact diagonalization techniques. Thus, the exact nonperturbative study of thermodynamics in small clusters with few interacting electrons is a formidable theoretical challenge, given that we do not know any exact solution for interacting particles at finite temperatures. The Hubbard model is a paradigm for theoretical studies of electron correlations in small and large thermodynamic systems. In particular, interesting results were obtained by exact diagonalization on small clusters. Studies of electron instabilities using exact expressions for grand canonical and canonical partition functions provide important insights into numerous many body problems in condensed matter physics [48–57]. We shall see in Sect 15.7.1 that the character of these instabilities also strongly depends on a phase of the order parameter or sign of the energy gaps. Moreover, these exact cluster studies of charge and spin excitation gaps (Sect. 15.8.1–15.8.3) go far beyond the reach of approximate schemes such as QMC, DMFT, etc that haven't been able to capture or reproduce some of our results at low-temperatures and away from half filling. Particularly exciting, in this context, are ultracold fermionic atoms in optical lattices, which allow versatile tuning of the model parameters in studying various novel phases and phase transitions [58]. The study of clusters and nanoparticles allows us to unravel important details that depend on the local geometry and local electron interaction, to investigate the specific features of inhomogeneous, systems when studying corresponding frustrated clusters (triangles, tetrahedrons) as compared with bipartite systems (consisting of building blocks such as dimers, squares, or cubes). This approach gives important insights not only for isolated ensemble of clusters with given average number of electrons (for which it is exact), but also for respective spatially inhomogeneous solid materials. The tendency of pairing and ferromagnetism, which can be observed on the example of isolated clusters in thermodynamic equilibrium, is preserved for respective values of parameters also for concentrated systems. Exact calculations of electron correlations in

large and finite-size systems can unveil many body problems in physics such as magnetism, electron pairing, ferroelectricity, superconductivity, and display remarkable analogies between ultrasmall and the bulk materials. The local electronic properties and finite-size effects in nanoscale can be used to interpret the spatial peak structure obtained by scanning-tunneling-microscopy for open chains [59].

15.2 Transition Metal Oxides

A vital feature of the cuprate superconductors is their ability to accommodate chemical substitutions; i.e., atoms that (a) replace one of the atoms of the parent compound without disrupting the short-range lattice order and (b) have a different number of electrons in their outer shells. The excess electrons may enter the copper-oxide planes (electron doping), or electrons can be taken away from the copper-oxide planes (hole doping), as a result of such chemical substitution. One can eliminate next nearest neighbor couplings by replacing the planar square lattice with independent Mott-Hubbard plaquettes and treat them as an ensemble of clusters immersed in a particle heat reservoir, wherein electrons can transfer from cluster to cluster because of the thermal fluctuations. Concentration (counter) of excess electrons or holes (in short, doping) in “impurities” (plaquette) is one of the most important parameters that determine the low-energy properties of the cuprate compounds. The stable Mott-Hubbard behavior in parental high T_c cuprates at half filling is spatially homogeneous. In the absence of electron-hole symmetry for frustrated systems the electronic and magnetic properties are different in electron/hole doped cuprates. Although the origin for breaking the symmetry in electron/hole doping is not yet fully understood, the reason can be the sign of the hopping term or an additional variation of the inter-site coupling parameter c under electron doping.

15.2.1 Spin-Charge Separation

A wide variety of metallic ferromagnets and antiferromagnets have been observed to develop quantum critical behavior when their magnetic transition temperature is driven to zero (close to the quantum critical point) through the application of pressure, chemical doping, or magnetic fields. In these cases, the properties of the metal are radically transformed by the critical fluctuations, departing qualitatively from the standard (normal) Fermi liquid behavior, to form a metallic state sometimes called a “non Fermi liquid” (Luttinger-Liquid) or a “bad” or “strange metal” [60–66]. The “non Fermi liquid” behavior is generic to one-dimensional interacting electron systems [67,68]. One of its key features is a spin-charge separation; instead of quasiparticles, collective excitations of charge (without spin) and spin (without charge) are formed, which are quite independent and moves with different velocities. Therefore, differences

in spin and charge degrees can be captured by studying the spin and charge excitation characteristics (gaps and susceptibilities). The spin-charge separation is absent in highly doped, metallic, systems. Experimental confirmation of spin-charge separation by probing charge and current in quasi-one dimensional organic conductors, the organic Bechgaard salts, etc. is reported in [69]. There is particular interest in these unusual bad metallic states in ceramic high T_c cuprates with local electron interactions. Since the initial suggestion, considerable efforts have been devoted to interpret the existing experimental data using spin-charge separation based scenarios. On the other hand, the initial data have also been analyzed without invoking spin-charge separation. Anderson has proposed a non-Fermi liquid (Luttinger liquid) theory for describing the electronic properties or anomalous features in HTSCs [70,71]. For example, the unconventional Hall and longitudinal resistivity effects in textured HTSCs [72] can be explained by the assumption of two different scattering rates, a transport scattering time and a Hall (transverse) scattering time [73]. A particularly significant dilemma is whether or not spin-charge separation exists in higher dimension. In the absence of inhomogeneity, the differences in the behavior of spin and charge degrees (provided by calculations of spin and charge gaps and corresponding susceptibilities) is very indicative for the understanding of partial or full spin-charge separation. Recently, the spin-charge separation phenomenon has been proven to exist in clusters in various two and three dimensional cluster topologies [49]. In grand canonical ensemble the spin-charge separation seen in susceptibilities at relatively low temperatures is often accompanied by Mott-Hubbard-like metal-insulator or from antiferromagnetic into spin liquid transitions. The separation of charge and spin degrees of electrons confined in building blocks depends on cluster geometry and local interactions [48]. There is a full reconciliation of spin and charge degrees at low temperatures in frustrated clusters with particular sign of hopping term ($t > 0$) and bipartite clusters at weak coupling, $U < U_c$. On the other hand, experimental observation of spin-charge separation still remains a great challenge for large thermodynamic systems in higher dimensions and HTSCs. Experiments show no any sign of spin-charge separation in the close vicinity to T_c in HTSCs, although there is an indication of a partial spin-charge separation at higher temperatures $T > T_c$ (see Sect. 15.9.1). Experimentally, charge ordering of stripes is always observed at a higher temperature than the magnetic ordering, which implies the differences in behavior of charge and spin degrees, i.e., partial spin-charge separation above T_c [74]. The incipient spin-charge separation phenomenon found in [48] can also lead to the formation of unidirectional nanostripes and other imperfections and inhomogeneities in HTSCs. The feasibility of experimentally studying the spin-charge separation effect is discussed in [75,76].

Doped Cuprates and Manganites

Transition metal perovskite oxides exhibit numerous extraordinary phenomena under doping. The best known examples are: high T_c superconductivity in copper oxides, colossal magnetoresistance, and ferroelectricity in manganese oxides (“manganites”). The observation of charge ordering and phase separation is possible in real space with atomic-scale resolution using the scanning tunneling microscope and spectroscopic imaging. By combining images and current-voltage spectroscopy data, it is possible to analyze the correlation between charge and structural order and the local conductive state (either metallic or insulating). These experiments provide an atomic-scale basis for descriptions of manganites as mixtures of electronically and structurally distinct phases. The STM images display many of the phenomena that have been proposed for the manganites, most notably charge ordering. The STM is the most direct probe due to its sensitivity to the outer valence electrons. In addition, diffraction and transmission electron microscopy, which have been the tools for the initial exploration of the charge ordering phenomena, are also primarily sensitive to atomic core positions and electron charge [77]. Ferroelectrics, magnetoelectrics, and multiferroics with magnetic and electric ordering can also exhibit spontaneous transitions and local micromagnetic inhomogeneities. Perovskite ($\text{La}_2\text{NiMnO}_4$, $\text{Bi}_2\text{NiMnO}_6$, and $\text{Bi}_2\text{FeCrO}_6$) materials in thin film and nanostructured form also display spatial inhomogeneities, and phase separation which can be easily detected [78].

Electronic Characteristics

Strongly correlated electrons in nonstoichiometric cuprates, manganites, and other transition metal oxides exhibit high T_c superconductivity, magnetism and ferroelectricity accompanied by spatial inhomogeneities at the nanoscale level [79–87]. The ongoing search for new superconductors has recently yielded a new family of Fe-based compounds, which upon doping exhibit superconducting transition temperatures T_c above 50 K [88]. Thus, the traditional paradigm of superconductivity in conventional materials has been broken again. It is textbook knowledge that superconductivity and ferromagnetism are mutually exclusive. Iron is a crucial component of these compounds (iron pnictides). The Fe-based family in many respects is similar to the cuprates and share a common behavior [89]. For instance, a spin density wave was observed in the parent compound LaOFeAs , which vanishes in the high temperature phase (above 150 K) and large doping. These materials display a phase diagram close to the cuprate family (for example, $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$). However, unlike the cuprates, whose parent compound is an insulator, LaOFeAs is a semimetal. Theoretically, the transition temperature estimated based on the electron-phonon coupling seems unlikely to explain the observed superconductivity, thus suggesting these materials might be unconventional and non-electron-phonon mediated. Despite the tremendous effort made to

understand the interesting physics of this new family of superconductors, the microscopic mechanisms that lead to superconductivity and ferroelectricity are still not known. There is a strong belief that electron instabilities observed in cuprates and manganites are crucial for providing mechanisms of superconductivity and ferromagnetism. The developed bottom up cluster approach in Sect. 15.6 applied in real space provides an understanding of electron instabilities and the importance of local real space inhomogeneities for the coexistence of superconductivity and low spin ferromagnetism for attractive electrons [57]. Although the experimental determination of various inhomogeneous phases in the cuprates is still somewhat controversial [90, 91], the underdoped high T_c superconductors (HTSCs) have many common features and are often characterized by crossover temperatures below which excitation (pseudo)gaps in the normal state are seen to develop [48]. The detailed manner in which T_c and other critical temperatures for various competing phases are changed under the variation of electron concentration, magnetic field, or pressure (Coulomb interaction) is also of fundamental interest for the formulation of the microscopic models responsible for superconductivity [92]. The additional nonstoichiometry and local disorder are introduced to the system because of the electron or hole doping resulting in many profound implications. In the optimally doped cuprates, the correlation length of dynamical spin fluctuations is very small [93], and hence short-range fluctuations are dominant over long-range electron correlations. Therefore, a microscopic theory, with local short-range dynamical correlations, even on a small (cluster) scale can give useful insight into Cooper type pairing correlations and possible superconductivity in real r-space and the rich physics observed in the high- T_c cuprates, nanoparticles, etc.

Phase Separation

The electron instabilities under the variation of doping, pressure, and local Coulomb interaction in cuprates and manganites are manifested by strong charge and spin fluctuations at some critical parameters, where charge or magnetic order disappears. The critical values for these parameters in the ground state can also be found from the level crossings at quantum critical points (QCPs), where the energy gaps or difference between the lowest canonical energies introduced in Sect. 15.7.2 disappear. Near the QCPs these materials exhibit density phase separation into hole-rich and hole-poor regions, similar to the phases in the spinodal region in abrupt phase transitions. The local charge fluctuations and coexistence of segregated phases leads to various imposed inhomogeneities at thermodynamic equilibrium. There is growing experimental and theoretical evidence suggesting that the superconducting state in the cuprates is inhomogeneous such that the locally defined charge density at low temperatures varies across the sample. The superconducting electrons in these materials also exhibit a short nanometer coherence length in contrast to conventional superconductors. One of

the defining characteristics of the HTSCs is the atomic-scale disorder. Spatially inhomogeneous features in the spin and charge distribution have been indicated in a number of experiments for both cuprates and manganites. The implication of doping induces additional disorder at the atomic scale in real space. The simplest realization of this phase is the so-called nanostripe phase in which charged clusters in nanoscale show distinct patterns, and the remainder of the sample is essentially an antiferromagnetically correlated Mott-Hubbard-like insulator (see Sect 15.7.1). Similar stabilization of inhomogeneity in the ferromagnetic phase of the manganites can also be related to the various instabilities. This picture corresponds to a nanoscale distribution of charge and spin, rather than a global phase separation. These experiments lead us to a central question: Are the found intrinsic inhomogeneities in superconducting and ferromagnetic states in cuprates and manganites of the same origin? We believe that the mechanism of phase separation and electron inhomogeneities in cuprates and manganites are of different nature in contrast to what is usually discussed in the literature, such as in [94,95]. The mechanism of local inhomogeneities in manganites is a spin driven phase separation in contrast to the cuprates, where the phase separation and inhomogeneities occur from the variation in the charge density [48]. The phase separation here is often first order phase transition due to spinodal phase segregation that occurs at a small length scale and can be accomplished in a perovskite by lattice instabilities and cooperative local oxide-ion displacements (see Sect. 15.8.3) similar to Jahn-Teller deformation. In $\text{Ln}_{1-x}\text{A}_x\text{MnO}_3$ perovskites the Coulomb repulsion also induces charge ordering (CO) below a certain temperature, leading to a rich variety of charge-ordered/orbital-ordered (CO/OO) structures and spatial inhomogeneities.

In the cuprates, the doped charge can be distributed inhomogeneously. Depending on the doping level, the inhomogeneous charge distribution can lead to the formation at nanoscale and a subnanoscale of charged stripes and clusters. Clusters containing the hole-poor (charge neutral) phase in the CuO_2 planes in T_c show proximity to half filled ordered Mott-Hubbard-like phase. In our opinion, this spontaneous phase separation in the normal state at nanoscale in the cuprates is the key point for understanding the mechanism of unconventional superconductivity in materials with strong spatial inhomogeneities (see Sect. 15.5.5). In many respects, the phase separation here resembles instabilities observed in spinodal region in the first order phase transitions. The doped holes in the CuO_2 planes have a pancake-like shape and are distributed inhomogeneously in these clusters. This is usually called topological doping. The charge distribution in cuprates is inhomogeneous both on a microscopic and a macroscopic scale. In high T_c cuprates neutral (hole-poor) clusters coexist with hole rich charge-stripe domains. The separation of charge creates an inhomogeneous redistribution of hole pairs that can become delocalized to render the compound conducting and, at low temperatures, superconducting properties.

Therefore, the insulating charge-stripes (proximity to Mott-Hubbard-phase) surrounded by conducting hole-rich domains can provide favorable conditions for (electron) pairs to propagate through inhomogeneous media [48]. The topological order and stripes are observed also in MnO_2 planes in manganites. With pirovskite structures at relatively strong electron interaction. In manganites under doping spin phase separation inhomogeneities in Sect. 15.5.5 can be accompanied with the Mott-Hubbard-like insulating behavior and ferromagnetism [48, 49, 56].

15.2.2 BCS Versus High T_c Superconductivity

The electron instability with regard to pairing and formation of electron Cooper pairs, which is in fact an instability of electrons in solids, is called superconductivity. Many of the experimental work on the high temperature superconductors can still be interpreted according to the BCS quasiparticle picture. The superconductivity in conventional BCS theory for itinerant electrons has been established in the k-space. The BCS superconductivity implies electron pairing and also the long-range phase coherence. Cooper quasiparticles with coupled charge and spin thus can behave as bosons, and so can collectively form a coherent state at low temperatures. The attractive electron-electron interactions between states of opposite spin and k result in the formation of coherent electron pairing with spin-zero (singlet) state with a unique bound energy gap Δ for quasiparticles or so-called Cooper pairs. The BCS theory fully relies on a momentum picture, where Cooper pairs consist of two electrons with opposite spins and wavevectors k and $-k$. For classical superconductors, the “coherence” length, or the approximate spatial extent of Cooper pairs, is usually hundreds of nanometers, so that these paired states make the r-space local picture of pairing largely irrelevant. The mean-field BCS theory has come under close scrutiny quite a while ago due to its application in the theory of metallic nanograins [13–14]. In recent years high T_c superconductivity in YBCO cuprates also was found dramatically different from the BCS picture. The superconductivity is confined to the Cu-O layers sandwiched between Y layers. In contrast, the correlation length of dynamical spin fluctuations in the optimally doped cuprates is very small [93], and, hence, short-range fluctuations are more dominant than long-range ones. In addition, inhomogeneities in high T_c superconductivity make k-space picture for uniform pairing inadequate. A typical $T - x$ (temperature-carrier concentration) phase diagram near half filling displays antiferromagnetic behavior. With increased hole doping, the long-range order fades away, while short-range antiferromagnetic order persists in the region where superconductivity prevails at appropriate doping. The STM measurements have shown that electronic states form standing wave patterns at temperatures far below T_c , which clearly indicate that four unit-cell periodicity patterns observed in the superconducting state persist well above T_c , in the pseudogap state. In this regime, many different experimental studies

of the cuprates have long claimed the absence of well-defined quasiparticles [86]. Quasiparticle dispersions along the nodal directions as measured by ARPES in the families cuprates of cuprates such as $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$, $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+d}$, $\text{Bi}_2\text{Sr}_2\text{CuO}_{6+d}$, show a clear “kink” or break in the dispersion around 70 meV. If the pseudogap regime is indeed devoid of any coherent quasiparticles, then these results suggest that other phenomena, such as stripes or other inhomogeneities, must most likely play a role in the formation of these patterns. The nature of the orbital structure of these electron Cooper pairs also remains one of the central questions in theory of superconductivity (see Sect. 15.2.1). The spatially local character of Cooper pairs with atomic-scale variations in confined inhomogeneities makes the r-space picture in high T_c superconductors more relevant (Sect. 15.7.7). The current description of the superconductivity in novel superconducting materials (CN, HTSC) is mainly based on the traditional BCS theory, assuming that the Cooper pair scenario is the only possible mechanism for superconductivity. However, in HTSCs the BCS quasiparticle paradigm for electron pairing doesn’t comply with the experimental observation in STM measurements of the two energy gaps, two characteristic temperatures of condensation, pseudogap, incoherent pairing above T_c , magnetic dormant states, etc. In fact, we analyzed earlier in [41] some of these drawbacks of the BCS theory by comparison of the GSCF approach in attractive Hubbard model ($U < 0$) with the exact Bethe-ansatz solution. The issue of symmetry of the order parameter or the sign (phase) and magnitude of energy gap related to the electron interaction are still being debated. The so-called s-wave gap requires electron attraction, while d-wave can take advantage of electron repulsion. The ratio of the gap over T_c exceeds the BCS value of $3.5 k_B T_c$. In our opinion, thermodynamic properties of small Hubbard clusters under variation of composition, size, structure, temperature and magnetic field have not been fully explored, although there have been numerous exact calculations [20–26]. From this perspective, the exact diagonalization in real space of small Hubbard *nanoclusters* can give useful insights into the origin of superconductivity, ferromagnetism and ferroelectricity in an ensemble of clusters, nanoparticles, and eventually, nanomaterials [48].

The intrinsic spatial variations in the electronic properties may play a key role in determining the nature of collective versus quasiparticle excitations in the high-temperature superconductors [96–101]. To understand the pairing mechanism and to search for new, even higher T_c superconductors, the excited-state electronic structure must be understood in both r-space and k-space. These excited charge and spin states can be studied by two closely related single-particle spectroscopies: angle-resolved photoemission spectroscopy (ARPES) and spectroscopic imaging scanning tunneling microscopy (SI-STM). ARPES directly probes the single particle spectral function and, therefore, can offer a complete picture of the many body interactions in materials with the perfect crystalline structures [69]. Therefore, due to the incoherent nature of the excitations, which do not live long near the

Fermi surface, the applicability of these traditional techniques for “nonideal” crystalline structures with spatially averaged quantities over atomic scale here is limited. In contrast, with the data in the superconducting state, where the standing wave pattern along $\langle 0, \pi \rangle$ persists also in the pseudogap regime is difficult to understand based on the quasi-particle scattering scenario. There is also evidence for the lack of well-defined quasiparticles at temperatures above T_c in the optimally doped cuprates. The STM measurements clearly indicate that some component of the modulated electronic structure does not rely on the formation of well-defined quasiparticles, but is due to the spatial ordering of charges [86]. The resonance valence bond (RVB) theory also provides spin-charge separation for charge and spin excitations, i.e., holon and spinon. In ARPES, a real electron - containing both spin and charge of course - is ejected. One doesn’t know a priori whether any ARPES peak results from (a) a convolution involving a spinon, a holon, or both; (b) a convolution involving other objects of exotic quantum numbers; or (c) simply a quasiparticle-like excitation. From this perspective, ARPES does not provide the quantum numbers of the elementary excitations and, hence, does not directly probe spin-charge separation. Further discussions along this line can be found in [75, 76]. A more appealing approach for studies of spin-charge separation phenomenon can be electron tunneling experiments in HTSCs. This scanning probe microscopy (SPM) technique was successfully used for imaging the flow of electrons (conductance) through a quantum point contact in quantum dots in nanotubes [102]. To understand the collective and localized behavior and spontaneous transitions between them, it is convenient to use various approaches and practices developed for the understanding of homogeneities and inhomogeneities in mixed valence compounds.

15.2.3 Localized Versus Itinerant Behavior

Generally speaking, in solid-state physics, there are already sub-fields wherein these two different descriptions in real and momentum space are typically applied. The first is a local picture, in which one visualizes the quantum states of electrons in atomic orbitals (Hitler-London) or at impurity atoms in real space. The second is the Bloch momentum or reciprocal space picture, in which itinerant electrons are delocalized and collectivized throughout the whole crystal. Understanding of these two extreme features is essential for condensed matter physics, but it has not been a priority to combine both pictures to understand the electronic structure and excitations in cuprates and manganites. There is direct evidence for the existence of dynamic charge stripes in the superconducting cuprates. However, the inhomogeneities in manganites fluctuate slowly in time and space, while in the cuprates charge stripes resonate much faster. It makes the observation of charge inhomogeneity in manganites much easier than that in the cuprates. Therefore, it is useful to distinguish between fast and slow measurements in the studies of inhomogeneous phases. For example, in mixed valence compounds fast experimental studies of X-ray

absorption and photoelectron emission (snapshots) have revealed the presence of double peaks separated by an energy gap between different electron configurations [103,104], while slow measurements give rise to a single resonance peak in the intermediate metallic state. It is also important to distinguish between homogeneous and spatially ordered inhomogeneous mixed valence compounds. In the former case, the average valence, due to fluctuations between $4f^n$ and $4f^{n-1}s$ configurations, is spatially uniform and not an integer. In the case of inhomogeneous mixed valence, the alternating in space ionic (integer) configurations are well-defined [105,106]. The manganites are most likely in an intermediate regime between homogeneous and inhomogeneous mixed valence states [107]. Earlier, we have considered a local “impurity” resonant model for mixed valence state and, in a generalized Hartree-Fock approximation, found solutions that, when generalized to the many site problem, displayed continuous and discontinuous (spontaneous) transitions from ground states of integer to intermediate valence states [108,109]. Although it is beyond the scope of this paper to discuss the various electron instabilities and inhomogeneities in rare earth materials, it should be noted that the first order sharp (spontaneous) and second order smooth transitions are analogous to electron instabilities observed in finite size clusters [108] (see Sects. 15.7.4 and 15.7.6).

15.3 Scanning Tunneling Experiments

The spectroscopic imaging scanning tunneling microscopy (SI-STM) has undergone rapid development to allow the space electronic structure of cuprates to be studied. For example, energy-resolved local density of states (LDOS) imaging [110–112], using the STM measurements of differential tunneling conductance, yield a spatial image proportional to the local density of electronic states, LDOS. The resulting gap map technique [99] images superconducting energy-gap variations with atomic resolution by measuring the peak-to-peak energy separation in a typical spectrum. A Fourier transform scanning tunneling spectroscopy (FT-STs) technique [100,101] uses \mathbf{q} vectors, where spatial modulations are determined from the locations of peaks in conductance; the magnitude of Fourier transform of conductance is used to determine elements of the k-space electronic structure [113].

15.3.1 Pseudogap and Gap

In high- T_c superconductors an energy gap is already exist at the Fermi surface in the normal state. This is known as a pseudogap, and its origin and relationship to superconductivity is still open question in the physics of HTSCs and represents the hot issue in current theoretical debate. According to one view, this pseudogap is intrinsic to pairing (superconducting) gap, reflecting the existence of electron pairing without global phase coherence. The actual transition to high- T_c superconductivity occurs at lower

temperature, when phase coherence is established. In an alternative view, the pseudogap is entirely irrelevant to the mechanism of superconductivity. In contrast to conventional superconductors, the low-energy electronic excitations in HTSCs are insensitive to disorder in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+x}$, manganites [114], ruthenocuprates, and other cuprates leading to characteristic signatures in atomic scale quasi-particle scattering [112], scanning tunneling spectroscopy [115] (STM), and angle resolved photoemission (ARPES) [116–118]. The spectroscopy in $\text{Bi}_2\text{Sr}_2\text{CuO}_{6+\delta}$ yields new evidence for a common origin of the pseudogap and the superconducting gap scaling with each other, providing strong support for notion that they have a common origin [119]. The square pattern (checkerboard) modulations have been visualized by STM in [100,101]. STM measurements in underdoped Bi2212 provide information about two energy scales or gaps (Sect. 3.2): coherent pairing gap in nanoscale patches and preformed pair gap between the patches, where incoherent Cooper pairs are present. There is also some evidence for the presence in the normal state of two pseudogaps for spin and charge excitations (susceptibilities). These results also reveal pseudogap phase and static electronic orderings. STM measurements can determine the temperature dependence of the pairing energy gaps and electronic excitations in the absence of coherent pairing [84, 110, 120, 121].

The average gap appears to evolve continuously from the superconducting state into the pseudogap region, and the energy gap in the local density of states vary widely in real space, and many local regions remain gapped well above T_c . The larger the gap in a given nanoscale region, the higher the temperature at which the gap closes. These large inhomogeneous variations of electron charge distribution on the nanoscale level are among the most prominent features in this class of superconducting materials. Despite the inhomogeneity, the pairing gap develops locally at a temperature T_c , following the BCS-like relation $2\Delta/T_c \approx 7.9 \pm 0.5$ [84], which is approximately twice the BCS value. The phonon sidebands seen in STM experiments show that electron-lattice interaction in HTSCs is not sufficient evidence for electron pairing and possible superconductivity. Instead, the proximity to the Mott-Hubbard insulator under doping plays a crucial role in the mechanism of superconductivity based on electron repulsion. Another interesting observation is a spin pseudogap in the normal state well above T_c on the underdoped side of the HTSC phase diagram. STM studies of cuprate superconductors revealed that the most prominent phenomena of hole-doped (p-type) cuprates, including the predominant $d_{x^2-y^2}$ pairing symmetry, spin fluctuations, and pseudogap, are absent in the infinite-layer electron-doped (n-type) cuprates $\text{Sr}_{1-x}\text{La}_x\text{CuO}_2$ [122]. The optimally doped $\text{Sr}_{0.9}\text{La}_{0.1}\text{CuO}_2$ exhibits s-wave pairing symmetry, with a superconducting transition temperature $T_c = 43.0\text{ K}$, a large energy gap of order, $13.0 \pm 1.0\text{ meV}$, and no pseudogap in the normal state.

15.3.2 Two Energy (Temperature) Scales

Recently, the attention to thermal excitations in HTSCs has shifted to the number of energy gaps in the system. As established by a number of spectroscopic probes, primarily angle-resolved photoemission spectroscopy, the pseudogap manifests itself as a suppression of the normal state electronic density of states at Fermi energy. The self-energy for quasi-particles has strong momentum dependence, as clearly shown experimentally and theoretically in the high-temperature superconductors. The cluster approach allows for the best nanoscale resolution. This is particularly important for the ARPES pseudogap in electron-doped cuprates that has quite a detailed momentum and real space structures. The antinodal states are quasilocalized in r -space and cannot, therefore, be well-defined in k -space. Simultaneous electronic structure images from SI-STM reveal the spatial structure of the antinodal or pseudogap excitations breaking translational symmetry, and reduce the C_4 rotational symmetry of the electronic structure in each Cu plaquette to C_2 in Cu-O-Cu bond-centered patterns without long-range order [115]. On the other hand all the experiments support electron pairing in real space and, therefore, cluster calculations of local charge and spin densities in real space might be quite adequate for the interpretation of ARPES data. Both electron-(n-type) and hole-doped (p-type) high T_c cuprates exhibit close proximity to a quantum critical point and in general there is no electron-hole symmetry breaking in bipartite lattices. However, a symmetry breaking is still possible under doping in some frustrated geometries at appropriate sign of coupling, t (Sect. 15.2.8). Using the temperature dependent STM measurements [110] a previously unobserved narrow and homogeneous (spin) gap that vanishes near T_c was found in $(\text{Bi}_{1-y}\text{Pb}_y)_2\text{Sr}_2\text{CuO}_{6+x}$. In addition, a broad (charge) energy gap, which exists below and above T_c , is inhomogeneous and only weakly temperature dependent. These results not only support the two-gap picture, but also explain previously troubling differences between STM and other experimental measurements. These features give important information for the emergence of two energy scales and two characteristic critical temperatures leading (correspondingly) to incoherent pairing far above T_c and coherent condensation of superconducting electron pairs below T_c in HTSCs. Recently, the latter aspect of the HTSCs has been discussed based on STM data in [123]. Thus, spectroscopic probe measurements provide some key answers to the nature of the elusive pairing mechanism, existence of local inhomogeneities and two energy gaps in the superconducting cuprates. These data show that the mechanism of electron pairing, i.e., scenario with two temperatures of condensations in high T_c superconductors, is different from conventional picture.

Here, we also identify a new class of nanoscale experiments on cuprate superconductors in the current-carrying state [124]. These experiments are aimed at studying how the high T_c order parameter responds, in both real and reciprocal space, when driven by a sizable phase gradient. Superconductors

are well known for their ability to carry current without dissipation due to the macroscopic phase coherence between paired electrons. In reduced dimensions, wherein fluctuation effects can become pronounced, the loss of phase coherence and the appearance of dissipation under an applied current are important. It is believed that high T_c superconductors are inherently fragile against phase fluctuations by virtue of their low superfluid densities. Considering all these factors, it would be very revealing to probe the short-range phase dynamics of cuprate superconductors in the current-carrying state. Such phase fluctuations, when thermally and magnetically driven, could weaken the superconducting coherent pairing.

The scenario for electron pairing without coherent condensation has in fact been invoked to explain the well-known pseudogap phenomenon, in relation to the notion of incoherent bosonized pairs. However, despite a wealth of experimental data indicating the presence of phase fluctuations in the cuprates, there has also been much inconclusive and even contrary evidence. Novel scanning tunneling spectroscopy (STS) technique of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ (YBCO) in thin films strips carrying an applied current at 4.2 K shows that phase-sensitive spectral (spin) features are systematically suppressed, while the amplitude-dependent (charge) features remain largely unchanged. These results can be understood in the general context of opposite spin pairing rigidity of the high T_c order parameter [48]. These observations indicate that a supercurrent can weaken the local phase coherence while preserving the charge pairing in YBCO, suggesting that phase fluctuations in spin pairing do play a significant role in high-temperature superconductivity.

15.3.3 Coherent Versus Incoherent Condensation

Few recent problems in science have generated so many controversial discussions as the problem of high-temperature superconductivity since its experimental discovery. The fundamental question here is following: Are the superconducting state and the superfluid density in these materials homogeneous? In the BCS theory the coherent Cooper pairs in the momentum k -space are “bosonic-like” composite quasiparticles with double charge and integer spin (0 or 1). The superconductivity implies the (homogeneous) isotropic real space or ideal crystalline structure. Therefore, the BCS theory is invalid for studies of “dirty” superconductivity in inhomogeneous systems. Using the cluster calculations in Sect. 15.7.6, we assert that the coherent state responsible for superconducting behavior in fact can be “inhomogeneous.” Moreover, the simple cluster approach shows that existing inhomogeneities in electron distribution can become a driving force (mechanism) for electron pairing. Electrons (holes) are transferred more effectively through these inhomogeneities (from one cluster to another) in pairs with opposite spins. This factor also points to a spin origin of possible pairing mechanism in HTSCs. This implies that the depression of T_c by overdoping is associated with a decrease of the superconducting condensate density n_s in spite of the increasing normal-state

carrier density [125]. In the undoped state, cuprates behave as antiferromagnetic (AF) Mott-Hubbard insulators, and it is precisely upon doping with holes that these strongly correlated materials demonstrate various electron instabilities and in some range of parameters become superconductors.

Recent experiments seem to indicate that inhomogeneous textured (intrinsically nanoscale) phases characterized with the stabilization of coherent quantum BCS-like state exist in high-temperature superconductors. This is not surprising in retrospect since these are complex materials with competing time and length scale inhomogeneities and spatial boundaries arising from different electron concentrations. A relevant and non trivial question is whether those intrinsic and local inhomogeneities are essential to drive the system into a degenerate coherent state, i.e., a state with the BCS-like full phase coherency between electron charge and spin degrees with the formation of coherent condensate with unique energy gap at rather low temperatures [48]. With improved sample quality and resolution, there is reliable evidence for the existence of spin rigidity in the (zero) spin susceptibility. On the other hand, recent angle-resolved photoemission spectroscopy (ARPES) data suggest an inhomogeneous electronic distribution consistent with clustering of charge carriers into quasi one-dimensional structures [126]. Therefore, although the orientation, width, length, and dynamics of these structures remain to be elucidated, both the above classes of experiments appear to confirm a new paradigm for origination of the “nanostripe” phase because of the partial separation in real space of spin and charge degrees in high T_c superconductors. Motivated by this newly introduced paradigm, a class of novel inhomogeneous structures in an ensemble of clusters was found, which captures the magnetic and superconducting properties of these strongly correlated materials. Although the origin(s) of the mesoscopic skeleton of complex stripe segments in the CuO_2 planes is presently unclear, simple cluster approach reveals some important regularities in their distribution, as discussed in Sect.15.7.6. In Sect.15.7.4, we will show that the local Coulomb interaction in an ensemble of clusters can be a source responsible for the formation of spatially local charge inhomogeneities and spatially homogeneous opposite spin distribution.

We shall see in particular, that appropriate inter-configuration fluctuations can introduce an inhomogeneous electron charge distribution and opposite spin pairing that can preserve magnetic symmetries at low temperatures, which are distinctive in inducing substantial pair binding of holes, as well as explaining the spin rigidity in response to magnetic neutron scattering. Moreover, based upon the exact calculations of a microscopic model, we have developed a microscopic theory that provides an approximate scenario for the occurrence of a coherent superconducting state. We also see the connection of the resulting inhomogeneity induced superconductivity to recent experimental evidence. The local repulsive interaction in some cluster geometries can produce bound states of two holes. In the present manuscript, we discuss an alternate approach in which inter-configuration fluctuations in an ensemble of clusters lead to the formation of inhomogeneities generated dynamically

by the redistribution of holes. The pairing mechanism for repulsive Hubbard clusters is provided by the coexistence of clusters with homogeneous spin and inhomogeneous charge distributions. Hole pairs can transfer due to these inhomogeneities in the system as a consequence of a singlet background (avoiding a possible global phase separation).

15.3.4 Modulated Pairs in Cuprates

The hole concentration in the CuO_2 planes of BSCCO is proportional to the number of out-of-plane dopant atoms, which is also responsible for disorder. This has led to the proposition that poorly screened electrostatic potentials of the dopant atoms can generate a variation in the local doping concentration and thus give rise to the gap modulations observed in STM [127,128]. Using STM in [129], the authors recently found a direct correlation between the size of the energy gap characterizing the superconducting state and a modulation of the atomic positions in the material. In doing so, they have also made the first definitive observation at zero magnetic field of a modulated superconducting state known as a “pair density wave”. This shows that the modulated state has no connection with vortices or magnetic field. The copper atoms can be planar coordinated, pyramidal-coordinated, or octahedral-coordinated; the nonplanar oxygens are known as apical oxygens. It has been appreciated for some time that a direct correlation exists between T_c and the apical oxygens [130]. However, only very recently has it been possible to measure correlations between the inhomogeneities observed in STM and positions of dopant atoms [131], thus providing a clue to the relation between disorder and doping in this compound, as well as a means to examine the above proposals. As it is shown in [48], the modulated electron pairing due to the doping could arise at a microscopic level from local distortions in pyramidal and octahedral (perovskite) structures surrounding the dopant atoms (see also Sect. 15.8.3). In particular, now it is well established that the dominant factor for pairing is the couplings between the apical and base oxygen atoms.

15.3.5 Inhomogeneities

In the cuprates and manganites the distribution of charge and spin carriers is not homogeneous in comparison with ordinary materials. These charge and spin structural inhomogeneities take place at the nanoscale and subnanoscale, resulting in charge clusters (stripes) and magnetic domains, respectively. There is more evidence suggesting that “inhomogeneities” at the nanoscale level, in the so-called stripes surrounded by essentially neutral correlated Mott-Hubbard-like antiferromagnetic insulators [132, 133], play a defining role for electron pairing and origin of superconductivity on atomic scale in HTSCs [134,135]. Besides charge pairing, the inhomogeneities of purely electronic nature in cuprates and manganites under doping can exist in the form

of spatial phase separation [94,95]. The non monotonous behavior of the chemical potential found in self-consistent field theory points to a possible electron instability for the charge degrees in repulsive Hubbard model [41]. On the other hand, magnetic inhomogeneities in the spin channel were also found in the attractive Hubbard model in low spin region [42,43,45,136] (see Sect. 15.5). The classification of phase transitions, mechanisms of inhomogeneities and possible superconductivity in conventional and high T_c superconductors are debated hotly. Experiments indicate that the superconducting transition is second order, while it becomes first order in the presence of an external magnetic field. The superconducting state with an inhomogeneous periodic spatial variation of the order parameter becomes stable when a singlet superconductor is subject to a large Zeeman splitting [137, 138]. The Zeeman splitting could be due to either a strong magnetic field or an internal exchange field. Under such a strong magnetic or exchange field, there is a splitting of the Fermi surfaces for spin-up and spin-down electrons, and condensed pairs of electrons with opposite spins across the Fermi surface may be formed to lower the free energy from that of a normal spin-polarized state. Thus, according to Fulde-Ferrell-Larkin-Ovchinnikov [137,138] the phase separation with unequal numbers of spin up and down electrons can still be found within the BCS-like mean-field theory. For example, beyond a critical polarization, the attractive electrons can be separated into a phase that is consistent with a superfluid paired phase surrounded by normal unpaired (magnetic) electrons [57].

The magnetic inhomogeneities seen also in other transition metal oxides at the nanoscale level, widely discussed in the literature [139–142], can be crucial for the spin pairing instabilities and the origin of ferromagnetism in the spin subsystems [38, 143]. A phase separation of ferromagnetic clusters embedded in an insulating matrix is believed to be essential to the colossal magnetoresistance (CMR) related properties in manganese oxides. At sufficiently low temperatures, the spin redistribution in an ensemble can produce inhomogeneities in the ground state and finite temperatures [49, 56, 57]. A strong connection between electron correlation, geometric and electronic structure was found. It is generally believed that a strong on-site Coulomb interaction favors ferromagnetism and is detrimental for electron pairing and superconductivity in finite clusters [11]. Our exact studies of pseudogaps in finite-size systems have uncovered some important answers related to the problem of spin-charge separation, pairing, and thermal condensation of electron charge and spin. Despite this, there still remains a vast amount of knowledge that needs to be unravelled. The following questions are central to our current study: (a) What are the conditions for the phase separation and spin/charge inhomogeneities? (b) What is the role of inhomogeneities, and are these spatial spin/charge *inhomogeneities* crucial for the pairing mechanism in these compounds? (c) When treated exactly, what essential features can the Hubbard clusters capture that are in common with the “large” concentrated transition metal oxides?

From this perspective, exact studies of electron charge and spin instabilities at various inter-site couplings c and cluster topologies can give important clues for understanding the importance of charge/spin inhomogeneities and local deformations for the mechanism of pairings and magnetism in “large” concentrated systems whenever correlations are local. The existence of intrinsic inhomogeneities has been established in nanoclusters for ordered Mott-Hubbard-like phases in manganites [139–142] in the high hole-doping region above the Curie temperatures. Heterogeneity in charge distributions, such as charge segregation or charge ordering, is also one of the key aspects for understanding the new properties of several families of oxides exhibiting colossal magnetoresistance (CMR), “stripe” structures, etc. The spin and charge inhomogeneities in nanoclusters are believed to be crucial for understanding CMR in oxides and disproportionate charge in multiferroics [144, 145]. In this article, we review recent developments in the understanding of the relationship between the coordinate r -space and momentum k -space electronic spectroscopies that are used to explore various degrees of freedom in high-temperature superconductivity.

15.4 Bethe-Ansatz and GSCF Theories

There are many open questions and unresolved problems with regard to the understanding of inhomogeneities and excitation gaps in the minimal Hubbard model with a few adjustable parameters. Much progress has been achieved in the understanding of many body correlations in infinite and finite 2d lattices, using mean-field, slave bosons, RPA, exact diagonalization, and QMC calculations. For evaluation of mean-field theory and development of more accurate RPA-type perturbation techniques by including the fluctuations about mean-field solution, systematic numerical studies of mean-field ground state properties are necessary. It is the lack of an exact solution in the thermodynamic limit in 2d case that puts forward the test of the most general linear mean-field theory, using the Bethe-ansatz theory. From strong quantum disorder, the universal mean-field scheme with continuous broken symmetry is traditionally considered inadequate for strongly correlated systems in 1d case. It is certainly true for long-range characteristics, while short-range correlations can address some ground state properties quantitatively correct in wide range of $U > 0$ and $U < 0$ Hubbard model even in 1d case [41–43, 45].

Here, we present the results of numerical calculations of the Bethe-ansatz and GSCF in parameter space of U/t , magnetic field h and electron concentration $n = (N_{\text{latt}} - N)/N_{\text{latt}}$ (where N_{latt} and N are the total number of lattice sites and electrons). The local correlations in the GSCF approach are a straightforward extension of existing mean-field solutions for general U/t , n and h . Many authors have discussed the feasibility of a phase separation instability within the Hubbard, t - J , and s - d models [94, 95, 141, 142]. The recent interest in phase separation in the HTSC materials made us undertake

a detailed examination of this feature within the GSCF approximation. We show that the GSCF within a certain range of parameters, U/t and n , displays an electron instability, while the Bethe-ansatz gives, by virtue, only a stable monotonous solution everywhere.

15.5 Hubbard Model

Although there is no direct evidence to decide whether the Hubbard model is a valid description of realistic systems, many physicists intuitively believe that it is of high relevance to understanding the many body concept of local electron correlations, electron pairing, and magnetic instabilities responsible for high-temperature superconductivity and ferromagnetism. Thus, we consider a single band Hubbard model of purely electronic nature described by

$$H = - \sum_{\mathbf{r}, \delta, \sigma} t_{\mathbf{r}}(\delta) c_{\mathbf{r}\sigma}^{\dagger} c_{\mathbf{r}+\delta\sigma} + U \sum_{\mathbf{r}} c_{\mathbf{r}\uparrow}^{\dagger} c_{\mathbf{r}\downarrow}^{\dagger} c_{\mathbf{r}\downarrow} c_{\mathbf{r}\uparrow} \quad (15.1)$$

The first term describes the tunneling of electrons. Here, $c_{\mathbf{r}\sigma}$ is the electron destruction (creation) operator at lattice site \mathbf{r} with spin (or magnetic sub-level) σ , $t_{\mathbf{r}}(\delta)$ is the hopping amplitude for a particle at site \mathbf{r} hopping to site $\mathbf{r} + \delta$, where parameter δ (coordination number) defines a nearest neighbor sites, and U is the usual on-site (local) electron interaction energy.

Below we show that the GSCF variational energy for arbitrary $h \geq 0$ and $0 \leq n \leq 1$ is exact in both limiting cases $U/t \rightarrow 0$ and $U/t \rightarrow \infty$. First, this substantial input allows the formulation of a more accurate and rigorous perturbation procedure about the mean-field solution, which can establish the relationship for general U/t , h , and n between the converging perturbation theory using the variational principle [45]. Second, by analyzing GSCF solution, we can understand the drawbacks of the homogeneous Bethe-ansatz solution near half filling to formulate the bottom up approach (see Sect. 15.6) as an appealing alternative for studies of collective phenomena, and electron instabilities and inhomogeneities. It is interesting that the phase separation in the optimally doped region becomes apparent already within the GSCF approximate treatment of the local Coulomb interaction near half filling.

15.5.1 GSCF Decoupling Scheme

For completeness and to facilitate the comparison between the Bethe-ansatz and GSCF theory, we first briefly summarize the main results obtained earlier for the one-dimensional repulsive and attractive Hubbard Hamiltonian with a magnetic field ($h \geq 0$) [41–43, 45]. According to Wick's theorem, the following single-particle decoupling is assumed for the interaction term

$$c_{j\uparrow}^{\dagger} c_{j\downarrow}^{\dagger} c_{j\downarrow} c_{j\uparrow} \approx \langle c_{j\uparrow}^{\dagger} c_{j\uparrow} \rangle c_{j\downarrow}^{\dagger} c_{j\downarrow} + \langle c_{j\downarrow}^{\dagger} c_{j\downarrow} \rangle c_{j\uparrow}^{\dagger} c_{j\uparrow} - \langle c_{j\uparrow}^{\dagger} c_{j\uparrow} \rangle \langle c_{j\downarrow}^{\dagger} c_{j\downarrow} \rangle$$

$$-\langle c_{j\uparrow}^+ c_{j\downarrow} \rangle c_{j\downarrow}^+ c_{j\uparrow} - \langle c_{j\downarrow}^+ c_{j\uparrow} \rangle c_{j\uparrow}^+ c_{j\downarrow} + \langle c_{j\uparrow}^+ c_{j\downarrow} \rangle \langle c_{j\downarrow}^+ c_{j\uparrow} \rangle. \quad (15.2)$$

The approximation (15.2) takes into account the effect of electron-electron and electron-hole interaction in a linear approximation as average single-particle (mean-field) terms. The average local electron numbers $n_{j\sigma}^c \equiv \langle c_{j\sigma}^+ c_{j\sigma} \rangle$ or corresponding concentration of electrons and average spin are

$$\begin{aligned} n &\equiv \langle \hat{n} \rangle \equiv \frac{1}{N_{\text{latt}}} \sum_j \langle c_{j\uparrow}^+ c_{j\uparrow} + c_{j\downarrow}^+ c_{j\downarrow} \rangle, \\ s &\equiv \langle \hat{S}_z \rangle \equiv \frac{1}{2N_{\text{latt}}} \sum_j \langle c_{j\uparrow}^+ c_{j\uparrow} - c_{j\downarrow}^+ c_{j\downarrow} \rangle. \end{aligned} \quad (15.3)$$

The introduced local electron-hole order parameter $\Delta_j^{(+)} \equiv \langle c_{j\uparrow}^+ c_{j\downarrow} \rangle$ is connected with $\Delta_{\mathbf{q}}^{(+)} \equiv -(2U/N_{\text{latt}}) \sum_{\mathbf{k}} \langle c_{\mathbf{k}\uparrow}^+ c_{\mathbf{k}+\mathbf{q}\downarrow} \rangle$

$$\Delta_j^{(+)} = -\frac{\Delta_{\mathbf{q}}^{(+)}}{2U} \exp i(\mathbf{q} \cdot \mathbf{r}_j). \quad (15.4)$$

In GSCF, it is assumed that the average $\Delta_j^{(-)} \equiv \langle c_{j\uparrow} c_{j\downarrow} \rangle$ for $U < 0$ is zero (can be neglected). Here and below, the superscripts “ \pm ” signify solutions for $U > 0$ and $U < 0$, respectively [41–45]. There are direct relationships (see Sect. 15.7.3) between the expectation values for the local spin components $\langle \hat{s}_{jx} \rangle$, $\langle \hat{s}_{jy} \rangle$, $\langle \hat{s}_{jz} \rangle$ and the parameters n_{\uparrow}^c , n_{\downarrow}^c , $\Delta_j^{(+)}$. Namely,

$$\langle \hat{s}_{jx} \rangle = \text{Re} \Delta_j^{(+)}, \quad \langle \hat{s}_{jy} \rangle = \text{Im} \Delta_j^{(+)}, \quad \langle \hat{s}_{jz} \rangle = s = \frac{1}{2}(n_{\uparrow}^c - n_{\downarrow}^c), \quad (15.5)$$

where the transverse local spin $s_{j\perp} = |\Delta_j^{(+)}|$ components are modulated by a wave number q (see (15.4))

$$\langle \hat{s}_{jx} \rangle = -\frac{\Delta_q^{(+)}}{2U} \cos(qj), \quad \langle \hat{s}_{jy} \rangle = -\frac{\Delta_q^{(+)}}{2U} \sin(qj). \quad (15.6)$$

Taking into account (15.6), for the average transverse lattice spin $S_{\perp} \equiv (\langle \hat{S}_x \rangle^2 + \langle \hat{S}_y \rangle^2)^{1/2}$, we get

$$S_{\perp} = 0 \quad \text{if } q \neq 0, \quad (15.7)$$

$$S_{\perp} = \frac{|\Delta_0^{(+)}|}{2U} \quad \text{if } q = 0. \quad (15.8)$$

The expectation value of the lattice spin component along the magnetic field is a longitudinal magnetization s . The solution with $q \neq 0$ describes a spatially inhomogeneous state, while at $h = 0$, the uniform state with $q = 0$ may be conditionally called a “transverse ferromagnetic” state ($s = 0$, $S_{\perp} = n/2$). This state has the same energy as that of the fully polarized ferromagnetic

state ($s = n/2$, $S_\perp = 0$, q is arbitrary) at the same U/t and n in the presence of infinitesimal magnetic field $h \rightarrow 0$ (Sects. 15.5.3 and 15.5.6). A “transverse spiral incommensurate” phase with $0 < q < \pi$ exists only at $n \neq 1$ (see below) as in higher dimensions [39, 146–148]. The spin of conduction electron feels the spatially varying effective spins of other electrons (15.6) similar to the spiral phase in the Kondo model [149].

We consider the homogeneous spatial distribution of electron density and put $n_\sigma^c = n_{j\sigma}^c$ independently of the lattice site. The solution with the total momentum of the center of mass $\mathbf{q} \neq 0$ for the electron-hole pair corresponds to the spatially inhomogeneous spiral spin density waves. We call this approach developed for arbitrary $U \geq 0$, $h \geq 0$, and all $0 \leq n \leq 1$ the generalized self-consistent field (GSCF) approximation.

In the GSCF approach, the concentration of double occupied sites $D^{(+)}$ can be reduced to

$$D^{(+)} \equiv \frac{1}{N_{\text{latt}}} \sum_j \langle c_{j\uparrow}^+ c_{j\downarrow}^+ c_{j\downarrow} c_{j\uparrow} \rangle = D_0 - \frac{(\Delta_{\mathbf{q}}^{(+)})^2}{4U^2}. \quad (15.9)$$

Here, $D_0 \equiv n^2/4 - s^2$ is the limiting value of $D^{(+)}$ for all U/t , n , and s while $\Delta_{\mathbf{q}}^{(+)} = 0$. The magnitude $D^{(+)}$ in (15.9) depends on \mathbf{q} and decreases in expense of the density of bound electron-hole pairs $(\Delta_{\mathbf{q}}^{(+)})^2/4U^2$.

15.5.2 Canonical Transformation

Using the standard Bogoliubov canonical transformation to the new Fermi quasi-particle operators

$$\begin{aligned} \alpha_{\mathbf{kq}+}^+ &= w_{\mathbf{kq}} c_{\mathbf{k}\uparrow}^+ + y_{\mathbf{kq}} c_{\mathbf{k}+\mathbf{q}\downarrow}^+, \\ \alpha_{\mathbf{kq}-}^+ &= y_{\mathbf{kq}} c_{\mathbf{k}\uparrow}^+ - w_{\mathbf{kq}} c_{\mathbf{k}+\mathbf{q}\downarrow}^+ \end{aligned} \quad (15.10)$$

($w_{\mathbf{kq}}^2 + y_{\mathbf{kq}}^2 = 1$) the single-particle (quadratic) Hamiltonian $H_{\text{GSCF}}^{(+)}$ is diagonalized and the average energy $E_h^{(+)} \equiv \langle H_{\text{GSCF}}^{(+)} \rangle$ is reduced to

$$E_h^{(+)} = \frac{1}{N_{\text{latt}}} \sum_{\mathbf{k}, \lambda} \mathbf{E}_{\mathbf{k}\lambda}^{(+)}(\mathbf{q}) n_{\mathbf{kq}\lambda}^{(+)} - UD^{(+)}, \quad (15.11)$$

where $n_{\mathbf{kq}\lambda}^{(+)}$ are the occupation numbers of the quasi-particle states

$$n_{\mathbf{kq}\lambda}^{(+)} \equiv \langle \alpha_{\mathbf{kq}\lambda}^+ \alpha_{\mathbf{kq}\lambda} \rangle = 1 \text{ (if } \mathbf{k} \in \mathbf{K}_F) \text{ or } 0 \text{ (if } \mathbf{k} \notin \mathbf{K}_F), \quad (15.12)$$

and \mathbf{K}_F is the Fermi region (do not mix $n_{\mathbf{kq}\lambda}^{(+)}$ with the electron occupation number $n_{\mathbf{k}\sigma}^c \equiv \langle c_{\mathbf{k}\sigma}^+ c_{\mathbf{k}\sigma} \rangle$). Indices $\lambda = \pm$ in (15.11) denote the two quasi-particle states and differ from electron spin states σ . The two quasi-particle sub-bands in (15.11) are

$$\mathbf{E}_{\mathbf{k}\lambda}^{(+)}(\mathbf{q}) = \frac{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}+\mathbf{q}} + nU}{2} - \lambda E_{\mathbf{k}}^{(+)}(\mathbf{q}), \quad (15.13)$$

where the dispersion relation for electrons is

$$\epsilon_{\mathbf{k}} = -2t \cos(\mathbf{k} \cdot \mathbf{R}) \quad (15.14)$$

with $\mathbf{R} \equiv \mathbf{r}_{j+1} - \mathbf{r}_j$. The energy difference between these sub-bands is

$$2E_{\mathbf{k}}^{(+)}(\mathbf{q}) = \sqrt{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}} - 2Us - h)^2 + (\Delta_{\mathbf{q}}^{(+)})^2}. \quad (15.15)$$

15.5.3 Order Parameter $\Delta_{\mathbf{q}}^{(+)}$

Both the order parameter $\Delta_{\mathbf{q}}^{(+)}$ and the average spin s are continuous functions of U/t , n , and h (Fig. 15.1). From (15.9), we obtain a simple relationship between $|\Delta_{\mathbf{q}}^{(+)}|$ and $D^{(+)}$:

$$|\Delta_{\mathbf{q}}^{(+)}| = 2U\sqrt{D_0 - D^{(+)}}, \quad (15.16)$$

valid in the entire space of U/t , n , and s values. The order parameter $|\Delta_{\mathbf{q}}^{(+)}|$ away from half filling does not have a corresponding analog in the exact theory. However, we can compare $|\Delta_{\mathbf{q}}^{(+)}|$ with analogous parameter $|\Delta_{\text{ex}}^{(+)}| \equiv 2U\sqrt{D_0 - D}$, using the Bethe-ansatz result for D .

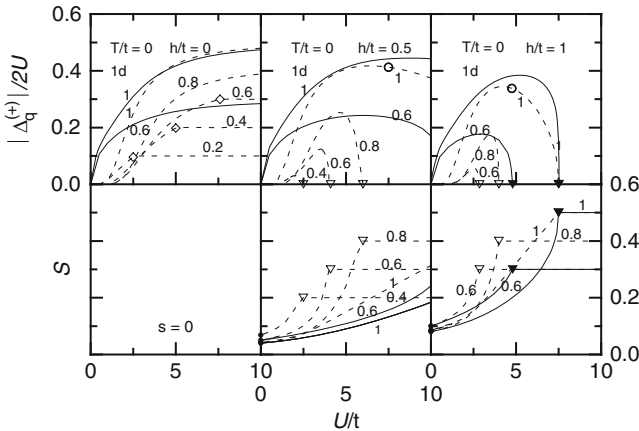


Fig. 15.1. The ground state longitudinal (s) and the local transverse ($|\Delta_{\mathbf{q}}^{(+)}|/2U$) spin components as a function of U/t for various n (figures label the curves) and h in the GSCF approach (*dashed curves*) along with the one-dimensional Bethe-ansatz result for $n = 1$ and $n = 0.6$ (*solid curves*). The triangles and rhombuses mark the longitudinal and transverse spin saturation correspondingly

In Fig. 15.1, at $h = 0$, we have $s = 0$, while $|\Delta_q^{(+)}|/2U$ increases monotonously with U/t and becomes saturated (if $n < 1$) as U exceeds some U_{homog} (marked by rhombuses). At $h = 0$ and $n = 1$, we have $|\Delta_q^{(+)}|/2U \rightarrow 1/2$ at $U/t \rightarrow \infty$. At $h > 0$, the spin s increases with U/t and becomes saturated ($s = n/2$) as U exceeds U_{sat} (marked by downward-pointing triangles), while $|\Delta_q^{(+)}|/2U$ vanishes at $U \geq U_{\text{sat}}$. The GSCF results (dashed curves) qualitatively agree with the exact ones (solid curves) at large U/t and strong field h . Our calculations show that the total spin

$$S_{\text{tot}} \equiv (s^2 + S_{\perp}^2)^{1/2} \quad (15.17)$$

monotonously increases with U/t and h as in the exact theory.

In general, the behavior of wave number q versus U/t and n in 1d case qualitatively resembles the corresponding one obtained for a square lattice [146–148]. In fact, the parameter q linearly increases with electron concentration in some range of U/t , and exactly at $n = 1$, it is independent of U/t (Fig. 15.3). However, at $h = 0$ and $n < 1$, the GSCF ground state is an incommensurate phase, while paramagnetic phase is a more stable mean field solution in higher dimensions [39, 150]. As in higher dimensions, there is a GSCF uniform solution $S_{\perp} = |\Delta_0^{(+)}|/2U$ with $q = 0$, and maximum transverse magnetization is stable at $h = 0$ and larger $U \geq U_{\text{homog}}$ (Fig. 15.1). The GSCF solution in the presence of magnetic field becomes fully saturated ($s = n/2$) and degenerate ($|\Delta_q^{(+)}|/2U = 0$ with q arbitrary) at $U \geq U_{\text{sat}}$ (on the right of triangles in Figs. 15.1 and 15.2) or $n < n_{\text{sat}}$ (on the left of the triangles in Fig. 15.3). Even though at $h \neq 0$, there is no true long-range order in 1d case;

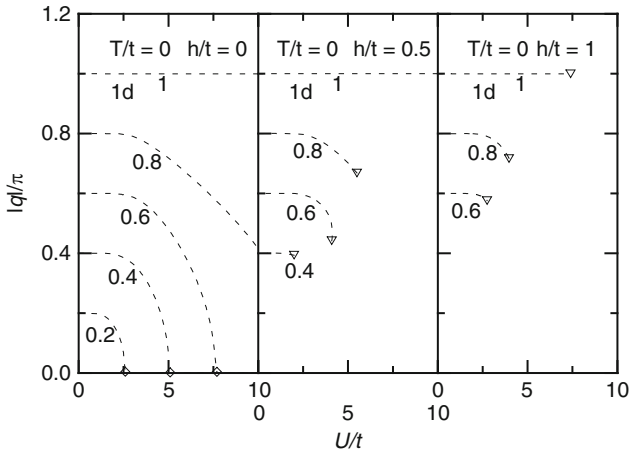


Fig. 15.2. The GSCF optimized wave number q of local transverse spin modulation as a function of U/t for various n (figures label the curves) and h/t . The triangles and rhombuses mean the same as in Fig. 15.1

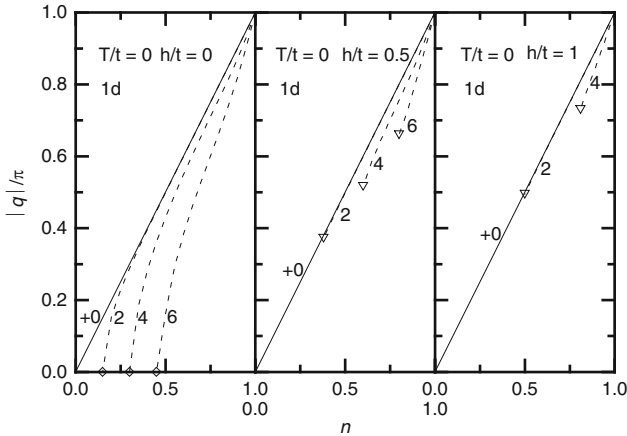


Fig. 15.3. The GSCF optimized wave number q as a function of n for various U/t (figures label the curves) and h/t . The triangles and rhombuses mean the same as in Fig. 15.1

by comparison of the GSCF quasiparticle spectrum, ground state properties, and the phase diagram to the exact results, we can estimate the short-range electron correlations (Sect. 15.5.6).

15.5.4 Quasi-Particle Spectrum

In the common case of $0 \leq n < 1$, the situation $\min \{ \mathbf{E}_{k^-}^{(+)}(q) \} < \max \{ \mathbf{E}_{k^+}^{(+)}(q) \}$ may occur. But we never found the case when the Fermi energy was higher than the minimum of the upper band, i.e., the corresponding upper band is always empty for all U/t , h , and $n \neq 1$. The corresponding gap at $n = 1$ for the quasi particle-hole excitations is $E_{\text{gap}}^{(+)} = 2 \min \{ E_{\mathbf{k}}^{(+)}(\pi) \}$. At all n , there is a correlation in the position of the minimum energy difference (15.15) between the two sub-bands and the wave number q shown by rhombuses. As Fig. 15.4 makes it apparent, arbitrary U removes the Kramers degeneracy, so that the quasiparticles near half filling fill the “anomausely” large region in the momentum space ($k \in \mathbf{K}_F$). In both figures, the occupied Fermi region is shown by thick curves. Although in many cases there are only two Fermi-impulses (k_{F1} , k_{F2}) corresponding to the same Fermi energy E_F , in some cases, we found at $h \neq 0$ also more than the two Fermi-impulses (k_{F1} , k_{F2} , k_{F3} , k_{F4} , ...). In general, the occupied part of the k -space within the first Brillouin zone happens to be closed or open, symmetric or non symmetric with respect to the origin of the k -space (Fig. 15.4). Thus, it is more convenient to simply introduce the width of the occupied part in the momentum space ($k \in \mathbf{K}_F$)

$$W_{\text{Fermi}} \equiv (k_{F2} - k_{F1}) + (k_{F4} - k_{F3}) + \dots \tag{15.18}$$

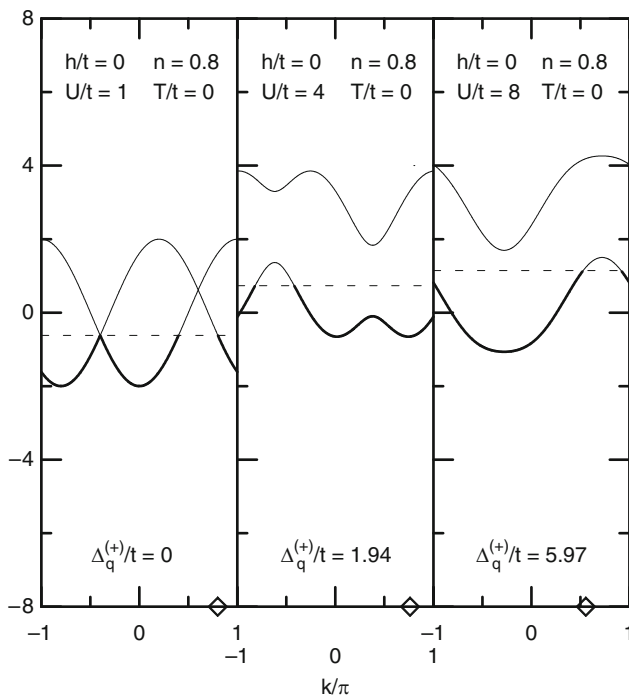


Fig. 15.4. The GSCF quasi-particle energy spectrum (15.13) for $n = 0.8$, $h/t = 0$, and $U/t = 1.0, 4.0, 8.0$. The notation is the same as for Fig. 15.4. In the represented cases, $s = 0$; $|\Delta_{\pi}^{(+)}|/t = 0.00, 1.94, 5.97$; $D^{(+)} = 0.16, 0.10, 0.02$, correspondingly. The Fermi regions are open and non symmetric with respect to $k = 0$

At $n \leq 1$, the width of the occupied region in the lower band $W_{\text{Fermi}} = n\pi$ is independent of U/t and h/t . This result is consistent with the Luttinger theorem [60–66]. In the limit $U/t \rightarrow 0$ and $n \leq 1$, the quasiparticles $n_{kq+}^{(+)}$ occupy Fermi region in the k -space between the Fermi-points $k_{F1} \leq k \leq k_{F2}$ symmetric with respect to k_0 , where $k_0 = 0$ at $n = 1$. In the following paragraphs, we compare the GSCF and exact ground state properties in the entire space of U/t , $h \geq 0$, and $0 \leq n \leq 1$.

15.5.5 Chemical Potential

In Figs. 15.5 and 15.6, the Bethe-ansatz results of μ versus U/t and n at given h/t (the solid curves) are compared with the GSCF ones (the dashed curves). The exact μ is a monotonous and smooth function of U/t , n , and h except the sharp kinks at the critical interaction strengths U_{sat} (triangles in Fig. 15.5) or at the critical electron concentrations n_{sat} (triangles in Fig. 15.6). In contrast, the GSCF $\mu^{(+)}$ is non monotonous in some range of U/t , n , and h . The

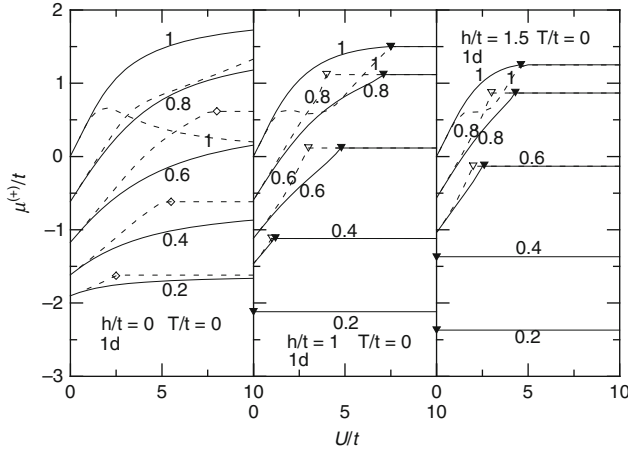


Fig. 15.5. The ground state chemical potential $\mu^{(+)}$ as a function of U/t for various n and h/t . The triangles and rhombuses mean the same as in Fig. 15.1

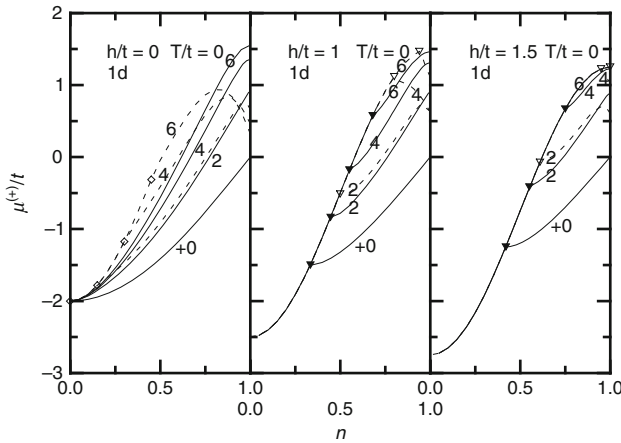


Fig. 15.6. $\mu^{(+)}$ as a function of n for various U/t and h/t . The triangles and rhombuses mean the same as in Fig. 15.1

variation of $\mu^{(+)}$ versus U/t in Fig. 15.5 shows monotonous behavior for all $n \neq 1$, and as in exact theory, $\mu^{(+)}$ at weak interaction increases with $U/t \ll 1$ linearly. Also, except for the $n = 1$ case, the numerical agreement between $\mu^{(+)}$ and μ becomes more accurate with increasing U/t . For $n \neq 1$, the GSCF result for $\mu^{(+)}$ everywhere slightly overestimates electron correlations (Fig. 15.5). But $\mu^{(+)}$ at $n = 1$ is non-monotonous in intermediate region of U/t , and as $U/t \rightarrow \infty$ and $h = 0$, we find $\mu^{(+)} \rightarrow 0$, while $\mu \rightarrow 2t$ is finite. Thus, at $n = 1$ and large U/t , $\mu^{(+)}$ underestimates electron correlations and

differs significantly from the exact result (Fig. 15.5). Analogous non monotonous behavior $\mu^{(+)}$ versus n also takes place at intermediate U/t near $n = 1$ (Fig. 15.6).

The transition near half filling in Fig. 15.6 is of the first order and it is accompanied by a (spinodal) region of phase separation between the two phases with different densities. The non monotonous behavior of $\mu^{(+)}$ versus n in one dimension (Fig. 15.5), observed within the universal GSCF theory, is a signature of electron instability that is expected also in higher dimensions [141, 142]. This instability from incommensurate state into the phase-separated phase in the Hubbard model is similar to the one observed in the doped manganites by treating with the Kondo lattice (s-d model). There are also considerable amount of experimental data pointing on the presence of a stable phase separation regime in doped cuprates, which also contain quasi-hole-rich ferromagnetic and quasi-hole free antiferromagnetic separated phases [95]. While both antiferromagnetic and ferromagnetic states at least correspond to a local minima in an energy, a spiral phase near $n = 1$ is absolutely unstable. The most energetically favorable state corresponds to ferromagnetic metallic states embedded into the matrix of antiferromagnetic dielectric [94, 95]. Within the GSCF theory, the magnetic field noticeably shrinks the region of electron instability phase with the GSCF negative compressibility $\chi_{\text{ch}}^{(+)} \equiv \partial\mu/\partial n < 0$ and shifts the maximum to the right, closer to $n = 1$ (Fig. 15.7). Notice that the electron instability at $h = 0$ in Fig. 15.6 takes place in the spiral phase far from transverse spin saturation (rhombuses). Such behavior is also in agreement with the experimental observation that the manganites are being in separated phase only if the magnetic moment is unsaturated [141, 142]. Thus, according to the GSCF theory, the phase separation in ferro- and antiferromagnetic states can be strongly suppressed by applied magnetic field. The full-range Coulomb interaction can also strongly suppress any local fluctuation toward the phase-separated state [151].

15.5.6 Ground State Phase Diagram

The Bethe-ansatz ground state at $h = 0$ is a singlet irrespective of n [27, 28, 29], while the GSCF theory at large U/t shows the tendency toward the Nagaoka-like transverse ferromagnetism (spatial homogeneous phase) for all $n \neq 1$ and $h = 0$ (Sect. 15.5.3). In the spiral phase, the longitudinal spin is non saturated ($s < n/2$). However, at sufficiently large $U \geq U_{\text{sat}}$, the system at $h \neq 0$ in Fig. 15.7 undergoes a transition into $s = n/2$ and $s_{j\perp} = |\Delta_q^{(+)}|/2U = 0$ with q arbitrary (degeneracy). In the presence of an infinitesimal magnetic field, this degeneracy results in a sharp transition (spin reorientation) from saturated transverse $s_{j\perp} = n/2$ ($s = 0$) into fully polarized longitudinal $s = n/2$ ($s_{j\perp} = 0$) phase. The U - n phase diagram at certain range of parameters U and n (or μ) displays electron spin instability. The thick dotted curve ($h = 0$) divides the entire U - n plane on two parts with $q > 0$ (the left upper part) and $q = 0$ (the right lower part). From self-consistent equations

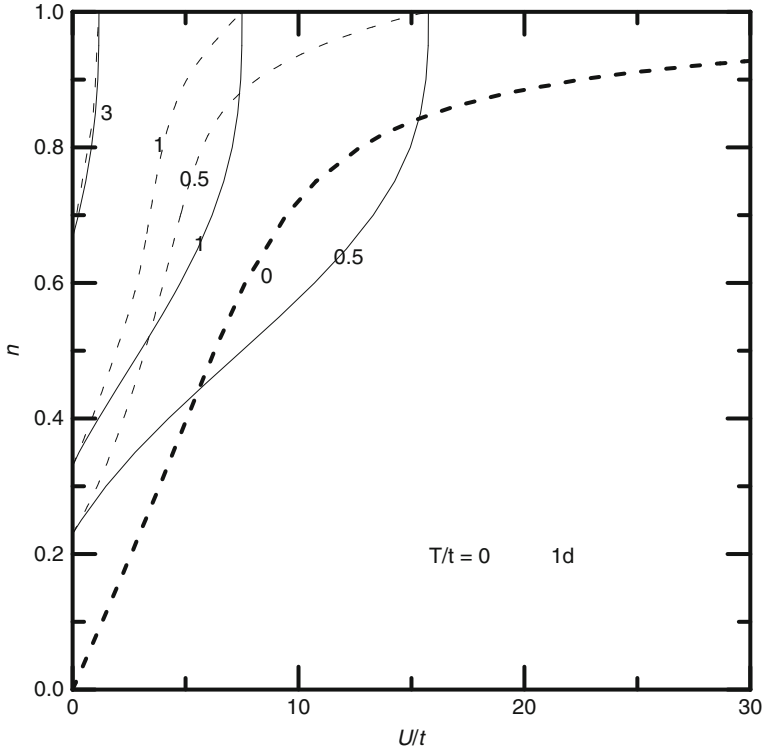


Fig. 15.7. The phase diagram in the U - n plan for various h/t (figures label the curves) in the Bethe-ansatz (the *solid curves*) and GSCF (the *thin dashed curves*) approaches. On the right of these curves, we have the saturated “longitudinal ferromagnetic” phase with $s = n/2$, $\Delta_q^{(+)} = 0$ (q is arbitrary) at the corresponding $h/t > 0$; on the left, we have the “transverse spiral incommensurate” phase with $q \neq 0$, $s < n/2$, $\Delta_q^{(+)} \neq 0$. The thick dashed curve corresponds to the boundary between the “transverse spiral incommensurate” phase with $s = 0$ and the spatially homogeneous “transverse ferromagnetic” phase (on the right of the curve) with the GSCF $\Delta_q^{(+)} = Un$, $q = 0$ and $s = 0$. At infinitesimal h/t , the thin dashed curve coincides with the thick dashed one

in [34] at $h = 0$ and $q \rightarrow 0$, we find the equation of the phase boundary in U - n plane, U_{homog} versus n between the spiral and homogeneous (ferromagnetic) phases

$$U_{\text{homog}} = 2\pi t \left(\frac{1}{\sin(n\pi)} - \frac{\cos(n\pi)}{n\pi} \right). \tag{15.19}$$

Below the curve U_{homog} versus n , we have the homogeneous (transverse ferromagnetic) state (at $h = 0$) or saturated longitudinal spin state (at $h \neq 0$). The spatial homogeneity at $h = 0$ prevails at large U/t and small n , while

the spiral phase dominates at weak interaction and in close vicinity of $n = 1$. In contrast to the exact theory, the GSCF phase diagram at $h = 0$ consists exclusively of spatially ordered magnetic phases and is similar to 2d case at $n \neq 1$ [146–148], where incommensurate phase in 2d case eventually loses its stability to the paramagnetic phase at $U/t \leq 1$ [39]. It is interesting that the GSCF phase diagram for the Hubbard model is similar to that for 1d Kondo model at $h = 0$. In fact, (15.19) is close to the corresponding boundary between ferromagnetic and spiral phases in 1d Kondo model with renormalized parameter $J_c = 2U_{\text{homog}}\pi n$ [149].

In the Bethe-ansatz approach, the critical magnetic field h_{sat} of phase transition is given by the formula ($U \geq 0$) [33, 34]

$$h_{\text{sat}} = U \left(-n + \frac{\gamma}{2} \right) - \frac{U\gamma}{\pi} \arctan \left(\frac{1}{\gamma \tan n\pi} \right) - \frac{4\theta t \cos n\pi}{\pi}, \quad (15.20)$$

where $\gamma = \sqrt{16t^2 + U^2}/U$, $\theta = \arctan(4t \sin n\pi/U)$. Particularly, at $U/t \rightarrow \infty$, we have $h_{\text{sat}} \rightarrow 0$ for all n . From (15.20), we can see that the exact h_{sat} decreases with U/t for all n and increases with n for all U/t . At $n < 1$ in the GSCF approach, $h_{\text{sat}} = 0$ at sufficiently large but finite U/t value, which implies that the longitudinal spin saturation may occur at infinitesimal $h \rightarrow 0$, because of the degeneracy mentioned above in this section. The phenomenon of spin-charge separation occurs at low energies away from half filling in the exact Bethe-ansatz solution and the GSCF approach. For repulsive interactions, the half filled band is gapped in the charge sector, while the spin excitations remain gapless. The mirror-inverted situation occurs also in attractive interactions wherein the gap opens in the spin excitation spectrum, while charge excitations are gapless [34].

15.5.7 GSCF Phase Diagram at $T \neq 0$

Much progress has been achieved in the understanding of the ground state properties and electron correlations in the low-dimensional Hubbard model by comparison of its predictions in generalized self-consistent field approach with exact diagonalization, Bethe-ansatz and QMC calculations. It is common knowledge that mean-field theory produces continuous broken symmetry ground states, while temperature (Mermin-Wagner theorem) and quantum fluctuations typically prevent the system from forming of the long-range order in one dimension. However, the approximate GSCF theory even in extreme conditions of one-dimensionality reproduces qualitatively and in some cases quantitatively the general ground state features of the Hubbard model. The GSCF results qualitatively well reproduce the exact Bethe-ansatz [35, 36] except in close vicinity of half filling. In contrast to the exact theory, the

GSCF chemical potential versus n in the metallic phase is a non monotonous function in the ground state and at very low temperatures. At finite temperatures, it becomes monotonous. The GSCF theory predicts the enhancement of antiferromagnetic correlations with temperature in the metallic state and properly describes the observed instability in the behavior of $\mu^{(+)}$ for ARPES data in $\text{Nd}_{2-x}\text{Ce}_x\text{CuO}_4$ [64, 152].

15.6 Bottom up Approach

Over the past few years, there has been a renewed increase in interest in electron instabilities in nanoclusters, assembled clusters of correlated materials in various topologies for synthesizing new nanomaterials with unique electronic and magnetic properties [18, 19, 153]. The bottom-up approach is motivated by a number of recent advances, both in experiments and theoretical studies of imperfect ceramic cuprates and inhomogeneous nanostructures. Bethe-ansatz for a given number of particles provides a closed form solution only for the spatially homogeneous state in one dimension [34, 107] that is difficult to analyze [29–32, 154] without having to resort to various approximations. In spite of significant progress in the understanding of the physics, there is still skepticism about the applicability of the Bethe-ansatz approach for studies of inhomogeneous systems, especially in the thermodynamic limit [155]. Perturbation theory is also usually inadequate, while numerical methods have serious limitations. A QMC approximation is not a reliable alternative, since the convergence at low temperatures due to the familiar sign problem is in general rather poor and the resulting phase diagram often leads to some controversy [156]. In contrast, exact calculations in small clusters [20–26, 150, 157] give an appealing alternative for the detection of possible phase separations and spatial inhomogeneities, especially at finite temperatures. As far as the authors are aware, an exact analysis of level crossing instabilities (degeneracies) in the canonical ground state eigenvalues and corresponding competing average energies at finite temperature for a general on-site interaction U and electron concentrations have not been attempted in small or moderate size clusters [158–162]. Exact computations of electron instabilities in various cluster geometries at the nanoscale level can be vital to the understanding of the role of thermal and quantum fluctuations for large pairing gaps and a transition temperature T_c in the correlated nanoclusters, nanomaterials, and corresponding “large” inhomogeneous systems [48–56].

Although our approach for “large” systems is only approximate, this class of clusters in an ensemble displays a common behavior, which we believe is generic for large thermodynamic systems. Our results for typical bipartite and frustrated (nonbipartite) cluster geometries have successfully mapped out scenarios in which many body local effects are sufficient to describe a spin-charge separation and pairing pseudogaps at the nanoscale level. Spatial

microscopic inhomogeneities have been observed in a number of STM probes in doped HTSCs. The coexistence of HTSC and inhomogeneity could be partially responsible for possible electron pair propagation and a high value of the superconducting transition temperature in an ensemble of clusters. The local inhomogeneities in HTSCs related to frustrated electronic phase separation are commonly expected in strongly correlated systems [163]. These ideas for inhomogeneous superconductivity have been further developed in the context of stripes [164]. It is important to distinguish these and related scenarios for possible pairing and superconductivity enhanced by local inhomogeneities for $U > 0$ from the homogeneous and coherent (metallic) superconductivity in perfect crystalline structures in conventional superconductors with electron attraction, $U < 0$. There is evidence suggesting that “inhomogeneities” at the nanoscale level, in the so-called stripes, surrounded by essentially neutral correlated Mott-Hubbard-like antiferromagnetic insulators [133,165], play a decisive role for the electron pairing and the origin of superconductivity at the atomic scale in HTSCs [112,134].

Besides the existence of charge pairing, inhomogeneities of possible electronic nature can exist in the form of spatially separated magnetic phases in cuprates and manganites under doping [94,95]. The magnetic inhomogeneities seen in other transition metal oxides at the nanoscale level [139–142,144,145], can be crucial for the spin pairing instabilities, origin of ferromagnetism and ferroelectricity in the spin and charge subsystems [38,143]. A phase separation of the ferromagnetic clusters embedded in an insulating matrix is believed to be essential to the CMR in manganese oxides. At sufficiently low temperatures, the spin redistribution in an ensemble of clusters can produce inhomogeneities in the ground state and at finite temperatures [49]. The non monotonous behavior of the chemical potential versus electron concentration found in the generalized self-consistent approximation [41] also suggests possible electron instabilities and inhomogeneities near half filling. From this perspective, exact studies at $T \geq 0$ of electron charge and spin instabilities at various $U \geq 0$, inter-site couplings $t' = c$, and various cluster topologies can give important clues for the understanding of charge/spin inhomogeneities and local deformations for the mechanism of pairings and magnetism in “large” concentrated systems whenever correlations are local. The flux criterion for pairing instabilities has been used for the Hubbard model and for two-dimensional and three-dimensional clusters. The conditions for the existence of pairing of carriers with charge $2e$ have been obtained in [56]. However, QMC calculations [166] in two dimensions show the absence of off diagonal long-range order. This negative outcome can be due two reasons: first, because the temperature at which the calculations were performed was not low enough, and second, because the parameter range in the Hamiltonian was apparently not suitable.

15.6.1 The Cluster Formalism

In nanoparticles, electrons and holes are limited to small regions and an effect known as quantum confinement yields a discrete spectrum and energy-level spacings between filled and empty states that can modify the thermodynamics. One can naively think that the standard tools for the description of phase transitions in clusters are not applicable and novel approaches are needed. However, it is important to realize that phase transitions and corresponding temperature driven crossovers in the grand canonical approach can be well-defined and classified without the use of the thermodynamic limit. There is also strong support with regard to the effectiveness of the grand canonical approach for studies of magnetism in small clusters due to recent experimental observation that average magnetization of a ferromagnetic cluster is a property of the ensemble of isolated clusters but not of the individual cluster [18, 19].

The sign of the hopping amplitude $t_{\mathbf{r}}(\delta)$ between the nearest neighbor sites in (15.1) leads, in the absence of electron-hole symmetry, to essential changes of electronic structure. For the planar (checkerboard) bipartite geometry, it is defined by choosing $t_{\mathbf{r}}(\delta) = t$ when δ connects sites within a square plaquette and $t_{\mathbf{r}}(\delta) = c$ when δ connects sites between two neighboring square plaquettes. In the so-called nonbipartite clusters such as tetrahedron, we consider coupling, $t_{\mathbf{r}} = \pm 1$. In addition, for the distorted pyramid or octahedron, we take the coupling parameter between the apical site and the atoms in the base equal to $t_{\mathbf{r}}(\delta) = c$. Numerical calculations of energy levels provide exact analytical expression for the grand canonical potential Ω_U in an external magnetic field (h)

$$\Omega_U = -T \ln \sum_{n, N} e^{-\frac{E_n - \mu N - h s^z}{T}}, \quad (15.21)$$

where N and s^z are the number of particles and the projection of the spin. The (first order) responses for first derivatives of the thermodynamic potential with respect to the chemical potential and applied magnetic field are as follows:

$$\chi_c(\mu) = \frac{\partial \langle N \rangle}{\partial \mu}, \quad \chi_s(\mu) = \frac{\partial \langle s^z \rangle}{\partial h} \quad (15.22)$$

The analytical expressions, derived for the average total number of electrons $\langle N \rangle$ and magnetization $\langle s \rangle$, are calculated exactly, without taking the thermodynamic limit.

15.7 General Methodology

Systems with correlated electrons in clusters display a rich variety of physical phenomena and properties similar to large thermodynamic systems [79–87] (phase separation, nanostripes, local gap and incoherent pairing, charge and

spin pseudogaps, etc.). The realization of these properties in clusters and bulk systems depends on interaction strength U , doping, temperature, the detailed type of crystal lattice, and sign of coupling t for the bipartite, and frustrated finite-size Hubbard lattice with implemented periodical boundary conditions in form of a rigorous theorem was formulated by Nagaoka at large U [38]. One can take a further step and consider a bulk system as a collection of many such decoupled clusters, which do not interact directly, but form a system in thermodynamic equilibrium [52]. Thus, one can consider a collection of such “molecules” or clusters, not at fixed average number of electrons per each cluster, but as a grand canonical ensemble, for fixed chemical potential μ at equilibrium. The distinction in local charge and spin density of states or corresponding susceptibilities having different pseudogaps can be considered as a signature of spin-charge separation. The electrons can undergo spin-charge separation due to level crossings driven by U or temperature, so that the collective excitations of electron charge and spin have different symmetries. Although this approach for large systems is only approximate, it nonetheless gives very important clues for the understanding of large systems whenever correlations are local. These results are directly applicable to nanosystems that usually contain many clusters, rather well separated and isolated from each other but nevertheless being in thermodynamic equilibrium with the possibility of having inhomogeneities for different numbers of electrons per cluster.

The main results regarding the cluster approach have been obtained in [48–57] (see also references therein). In this section, we provide a detailed description of the general methodology and the criteria for the charge and spin pairing instabilities in the canonical and grand canonical frameworks. We also discuss the rigorous conditions necessary for the occurrence of the quantum critical points, coherent pairings, and spontaneous transitions in the ground state and corresponding critical temperatures of crossovers for various phases and boundaries formulated recently in [48, 49]. As a basic building block of the CuO_2 planes in the HTSCs, various clusters can be used as a block reference to build up larger superblocks in 2d of desirable sizes by applying Cluster Perturbation Theory (CPT) [167], non-perturbative Real-Space Renormalization Group (RSRG) [168], Contractor Renormalization Group (CORE) [169], or Dynamical Cluster Approximation (DCA) for embedded 4-site clusters coupled to an uncorrelated bath.

15.7.1 The Canonical Charge and Spin Gaps

One can classify the charge and spin order parameters by analogy with phase transitions in the thermodynamic limit [49, 51–55, 57]. The opening of the gap is a local correlation effect, and clearly does not follow from long-range order, as exemplified here. In the ground state, the calculated differences in the canonical energy levels between configurations with various numbers of electron charge and spin determine the energy gaps for electron charge and

spin excitations. For given temperature T and interaction strength U , we have calculated the energy differences $\mu_+ = E(N+1) - E(N)$ and $\mu_- = E(N) - E(N-1)$ for the average canonical energies $E(N)$ by adding or subtracting one electron (charge) in the cluster for a given spin S (i.e., the difference between the excitation energies upon addition and removal of one particle). Thus, the charge excitation gap at finite temperature can be written as an energy difference, $\Delta^c(T) = \mu_+ - \mu_- = E(N+1) + E(N-1) - 2E(N)$. The physical meaning of $\Delta^c(T)$ by definition, involves three different states, and should not be interpreted as a property of only one of these states. Such a parameter, referred to a pairing energy, has been proposed earlier for the ground state ($T = 0$) energy and odd-even mass differences in nuclear physics [170, 171]. This gap parameter in [172] can capture the even and odd (parity) differences, which are measurable quantities in small grains. A similar gap parameter at $T = 0$ and half filling has been used also to study the Mott-Hubbard gap in [27] and pairing instabilities in [158–162] are analogous to the inter-configuration energy gap for the crossover between different many body ground state ionic configurations in solids. For example, the charge gap is simply equivalent to the excitation energy of the “reaction” between different cluster configurations d^N for fixed N

$$d^N + d^N \rightarrow d^{N+1} + d^{N-1}, \quad (15.23)$$

i.e., the difference in an canonical energies of ionization and affinity for many body cluster configurations in an ensemble. Equation (15.23) can be referred to as an analog of HOMO-LUMO like gap in [173] for interacting electrons. In the zeroth order approximation (atomic limit), when hopping term is neglected, we obtain electron correlations for pure ionic configurations. The inter-configuration fluctuations in cluster model can be considered as a next approximation for accounting the effect of potential screening in nanoclusters and nanostructured materials. Notice that expression (15.23) describes also a transition from a homogeneous d^N state into an inhomogeneous state, with different electron distribution, d^{N-1} and d^{N+1} . In an ensemble of clusters with two configurations d^{N+1} and d^{N-1} competing in energy for stability, a narrow resonance type increase in the electronic density of states is expected near quantum critical point, μ_P .

The main drawback of this quantum gas model is the absence of a direct coupling between clusters that prohibits quantum (hopping) transfer of electrons from one cluster to another. This drawback can be partially corrected by applying the canonical and grand canonical distributions. For example, in grand canonical approach, the clusters in contact with thermal reservoir are allowing the particle number to fluctuate. This is equivalent to the indirect transfer of electrons from cluster to cluster via a thermal excitations, which in fact implies that building block clusters do interact thermally. As we mentioned earlier, this approach can be considered appropriate for a qualitative understanding of concentrated bulk materials and large thermodynamic systems. However, the configurational change in the ensemble of isolated clusters

is supposedly due to the possible spontaneous fluctuations in the electron numbers and electron redistribution via a charge reservoir.

The spin gap in the canonical ensemble can be treated correspondingly. We calculate a spin gap as the difference in the average energies between the two cluster configurations with the various spin S states, $\Delta^s(T) = E(S+1) - E(S)$, for $E(S)$ being the average canonical energy in the spin sector at fixed N [34]. The redistribution of charge in excess electron/hole inhomogeneities or spin up/spin down in domains for an ensemble of clusters depends on the interaction strength, sign, and magnitude of the hopping term [49] (see Sect. 15.8). Regrettably, the canonical method discussed in this section becomes impractical at high-temperature since the number of states in the charge and spin sectors that need to be considered increases rapidly for $T \geq \Delta^s, \Delta^c$.

15.7.2 Quantum Critical Points: Level Crossings

The quantum phase transition of level crossing arises at a critical crossing point called quantum critical point (QCP). Close to the QCP, the system is characterized by a diverging correlation length, which is responsible for the weak or strong singular behavior of different physical observables [174]. The behavior of correlation functions, however, is not necessarily related to the behavior of long-range correlations. This behavior seems particularly interesting when a quantum phase transition (QPT) is associated with drastic modifications of the many body ground state in ultrasmall particles. There is no minimal critical length below which a quantum critical behavior disappears. Earlier in [52], we have found that the sharp QPTs can also occur in small Hubbard clusters driven by interaction strength, inter-site coupling, chemical potential, longitudinal, and transverse magnetic fields. Our recent observation is that the new type of the ground state QPT can occur in the bipartite and nonbipartite Hubbard clusters from the level-crossings induced by the cluster geometry [49]. The QCP found from exact calculations of the ground state energy gaps in Sect. 15.7.1 is defined from the condition $\Delta^c(U, c) \equiv 0$ at various electron counts N which yield, U_c , and c . Correspondingly, one can calculate the critical parameters for the onset of sharp transitions in the spin degrees from $\Delta^s(U, c) \equiv 0$. The vanishing gaps at in the ground state can be directly linked to QCP for the *onset of pair formation*. For instance, the QCP, U_c , separates electron–electron pairing from Mott-Hubbard-like electron–hole pairing, and also opposite spin pairing from the spin liquid state [49]. The QCP turns out to be a useful entity for the analysis of the phase diagram at zero temperature. The QCPs also can be derived within a grand canonical approach using exact calculations of charge and spin susceptibilities in the limit $T \rightarrow 0$.

15.7.3 Symmetry Breaking

As exemplified here, the symmetry-breaking transitions are still possible in finite-size systems and can be either smooth (second order) or sharp (first order). For example, the ferromagnetic phase, has a net magnetization (or a non zero total spin), the magnitude of which represents a spontaneously broken symmetry below the Curie point. Thus variables such as total spin or chirality are examples of natural order parameters. The transition observed in the underdoped regime in high T_c superconductors is unusual, in that it is not a mean-field (BCS-like quasi-particle) transition with a single transition temperature as seen in other conventional superconductors. Rather, there is a region in the phase diagram above the superconducting temperature T_c (where phase coherence and superconductivity begin) but below a characteristic temperature T^* where a “pseudogap” appears in the spectrum of electronic excitations. The exact calculations of phase diagrams at overdoped regime above T^* in clusters also clearly establish the presence of incoherent preformed electron pairs with spontaneous charge separation at different doping level.

Following the well-known results on U(1) phase symmetry for superconducting order [175–179] and SU(2) symmetry for rotational invariance, we can express the spin and pseudospin through electron operators as follows [34]

$$\begin{aligned}\hat{s}_{jx} &\equiv \frac{1}{2} \left(c_{j\uparrow}^+ c_{j\downarrow} + c_{j\downarrow}^+ c_{j\uparrow} \right), & \hat{s}_{jy} &\equiv \frac{1}{2i} \left(c_{j\uparrow}^+ c_{j\downarrow} - c_{j\downarrow}^+ c_{j\uparrow} \right), \\ \hat{s}_{jz} &\equiv \frac{1}{2} \left(c_{j\uparrow}^+ c_{j\uparrow} - c_{j\downarrow}^+ c_{j\downarrow} \right),\end{aligned}\quad (15.24)$$

which satisfy to the usual commutation relations. Averaging over all lattice sites, j gives the lattice spin components $\hat{S}_x, \hat{S}_y, \hat{S}_z$. Correspondingly, we can write for \mathbf{L} operators

$$\begin{aligned}\hat{L}_{jx} &\equiv \frac{(-1)^j}{2} \left(c_{j\uparrow}^+ c_{j\downarrow} + c_{j\downarrow}^+ c_{j\uparrow} \right), & \hat{L}_{jy} &\equiv \frac{(-1)^j}{2i} \left(c_{j\uparrow}^+ c_{j\downarrow} - c_{j\downarrow}^+ c_{j\uparrow} \right), \\ \hat{L}_{jz} &\equiv \frac{1}{2} \left(c_{j\uparrow}^+ c_{j\uparrow} + c_{j\downarrow}^+ c_{j\downarrow} - 1 \right),\end{aligned}\quad (15.25)$$

where $L_j^+ = c_{j\uparrow}^+ c_{j\downarrow}$ and $L_j^- = c_{j\downarrow}^+ c_{j\uparrow}$. Summation over all lattice (local) sites, j gives the total lattice spin $\hat{S}_x, \hat{S}_y, \hat{S}_z$ and total pseudospin $\hat{L}_x, \hat{L}_y, \hat{L}_z$ components. One can calculate expectation values for global operators L and S to discuss the symmetry at relatively low temperatures. At half filling, the state with $\langle L \rangle = 0$ and $\langle S \rangle = 0$ with $\Delta^{c,s} > 0$ corresponds to a Mott-Hubbard AF. However, away from half filling, for a certain range of U , the expectation values for $\langle L \rangle$ and $\langle S \rangle$ do not vanish. This is a sign of possible symmetry breaking with $\langle \hat{L} \rangle \neq 0$ with the formation of electron pairs. It is therefore important to establish whether T^* is just a crossover temperature arising from fluctuations in the order parameter that establishes superconductivity at $T \leq T_c$, or whether it marks a phase transition where symmetry is spontaneously broken [180]. In the spin pseudogap state, the time-reversal symmetry

is spontaneously broken in the overdoped region above T^* , which therefore corresponds to a phase transition. Sometimes ($U < 0$), a broken symmetry for orbital degree, $\langle L \rangle \neq 0$, can be accompanied with symmetry breaking for spin $\langle S \rangle \neq 0$ which is a sign for the coexistence of electron pairing and ferromagnetism [43, 181] at half filling.

The Nagaoka type ferromagnetism in a finite-size periodical lattice for one hole off half filling is an example of a symmetry-breaking transition in two and three dimension [38, 182, 183]. This symmetry, referred to as a “high spin symmetry,” is broken in the ferromagnetic phase with $\langle \hat{S} \rangle \neq 0$ because of the formation of a nonzero magnetic moment. It also depends on the sign of the hopping term t and the cluster topology, whether it is a bipartite or nonbipartite structure. For example, squares or cubes are the building blocks, or prototypes, of solids with bipartite lattices, whereas triangles, tetrahedrons, and octahedrons without electron-hole symmetry may be regarded as primitive units of typical frustrated systems (triangular, pyrochlore, perovskite). Exact studies of various cluster topologies can thus be very useful for the understanding of nanoparticles and respective bulk systems. For strongly correlated electrons in atomic limit $U/t \rightarrow \infty$ with one hole $n = 1$ or electron $n = -1$ off half filling, (where $n = N - N_e$) the ground state is predicted to be metallic, independent of the lattice structure. Conclusions about a stable spin configuration in the ground state are summarized as follows: (Case A) In bipartite lattices for both $tn < 0$ and $tn > 0$ and nonbipartite lattice for $tn < 0$, the ground state is a fully polarized ferromagnetic state; (Case B) In nonbipartite lattices with $tn > 0$, the ground state with one spin-flip is stable as opposed to the fully polarized ferromagnetic state. In the low spin state, the electron pairing implies the coupling of opposite spin, which can be considered as changing the spin and charge statistics from fermion (triplet) to boson (singlet), which eliminates some of the nodes in the spin part, transferring them to the radial charge part by converting holes into bosons. Correspondingly, in the high-spin limit, the state with one reversed spin can be considered as changing the statistics of electrons charge (holes) from fermion (triplet) to boson (singlet), which eliminates some of the nodes in the N -electron wave function, transferring them to the spin part. Generally speaking, this effect depends on the number of particles, the dimensionality, topology, and connectivity of the lattice, the interaction strength, the number of electrons, and lattice sites as well as the sign of the parameter t . Whether a fully or a partially polarized ferromagnetic phase persists up to a finite hole density is controversial and has been the subject of numerous investigations. The Nagaoka theorem for one extra hole off half filling predicts ferromagnetism either with maximum (saturated) or nonmaximum (unsaturated) spin at large U limit. One can flip a spin and find that the stable polarized spin state with respect to energy of nonmaximum spin leads to a negative spin (pairing) gap. Accordingly, the unstable saturated state, existing in frustrated lattice geometries with particular sign of hopping ($t > 0$) [38], implies a positive spin gap. The metallic nature of saturated Nagaoka state has also been disputed [48]. The validity of the

theorem for finite hole density, nonzero temperature, next nearest neighbor hopping, etc. is discussed in [49]. The small clusters for one extra hole or electron also exhibit the maximum spin at large U similar to saturated Nagaoka theorem. However, at relatively small U , we find a dominant instability that leads to electron pairing and opposite spin pairs rigidity in the metallic state. The stability of clusters with regard to inter-site coupling was considered in [48]. As it is shown in Sect. 15.7.7 for case B, unsaturated Nagaoka state at large U can be associated with low-spin, BCS-like coherent quantum state with minimal spin (singlet) and electron charge pairing [49].

Degeneracies in the energy levels of quantum systems have close relationships with symmetries. The Hubbard model exhibits two types of degeneracies: (a) a permanent degeneracy, such as (twofold) orbital or spin Kramers' degeneracies, refers to energy levels that stay degenerate for all U values; (b) crossing level degeneracies or corresponding quantum critical points (see Sect. 15.7.2) that occur at specific U values or intra-site coupling c . We observe that level crossings in clusters can not be explained in terms of U independent symmetries. In particular, for an ensemble of bipartite clusters, we obtain non trivial crossing degeneracies for charge and spin degrees from the conditions $\Delta^c(U_c) = 0$ and $\Delta^s(U_c) = 0$ for finite critical U_c values, respectively. In contrast, a spurious avoiding of level crossings occur for all U in frustrated clusters. An oscillatory behavior of the gap in planar square clusters manifests multiple level crossings for charge and spin degrees. Moreover, we found that the crossing level degeneracies [184, 185] can also occur through cluster deformation or by changing the sign of the hopping term t [49] in nonbipartite clusters. Thus, one can interpret these phenomena as spontaneously broken symmetries by considering the partial condensation of electron components in charge or spin sectors separately. Both charge and spin symmetries of different nature can be treated on an equal footing in this formalism. This greatly improves our physical intuition concerning complicated many body physics, which simplifies the use of approximate schemes for thermodynamically large inhomogeneous systems. One can consider the formalism to describe superconductivity, ferromagnetism, and other magnetic correlations using the same formal language for the energy gaps that we adopted for the above mentioned clusters. Similarly, the spontaneous ferromagnetism in insulating Mott-Hubbard-like state for charge degrees with $\Delta_c > 0$ is manifested in spin degrees at low temperatures by $\Delta^s < 0$. The gap picture gives an interesting interpretation of partial bosonization of electrons. For example, at half filling, the charge and spin pairing gaps at low temperatures are both positive and describe a stable Mott-Hubbard-like antiferromagnetic insulator (Sect. 15.7.4).

15.7.4 The Charge and Spin Instabilities

Many cooperative phenomena, invoked in the approximate treatments of the "large" concentrated systems, are also seen in the exact analysis of pairing

instabilities in the canonical ensemble of the small clusters in the thermodynamic equilibrium [49, 51–55, 57]. The sign and magnitude of the gap is also important in identifying the regions for the electron charge and spin instabilities, such as the electron–electron $\Delta^c < 0$, electron–hole $\Delta^c > 0$ pairings in the charge sector or the parallel $\Delta^s < 0$ and opposite $\Delta^s > 0$ spin pairings in the spin sector. The positive gap characterizes a smooth transition, and the negative gap is a signature for instability and tendency toward spontaneous phase transition and clustering. For example, the negative sign of charge gap implies phase (charge) separation (i.e., *segregation*) of the clusters on hole-rich (charge neutral) and hole-poor regions. The quantum mixing of the closely degenerate hole-poor d^{N-1} and hole-rich d^{N+1} clusters for one hole off the half filling case, instead of causing global phase separation, provides a stable spatial inhomogeneous medium that allows the pair charge to fluctuate. Moreover, in case of more than one order parameter, the key question is the exact relationship between the sign and magnitudes of the charge Δ^c gap and its corresponding spin Δ^s counterpart in various bipartite and frustrated cluster topologies. For example, instead of a full phase separation at $\Delta^c < 0$, the local inhomogeneities in the clusters with opposite spin $\Delta^s > 0$ (rigidity) can provoke electron redistribution and stability of various charge configurations. The positive charge gap, defining excitonic type electron-hole pairing, is indicative of the stability of the Mott-Hubbard insulating phase, while negative charge gap shows instability for electron binding into pairs and metallic behavior (see Sect. 15.7.3).

The inhomogeneities favored by the negative gaps are essential for providing a spontaneous redistribution of the electron charge or spin. The inhomogeneities in the charge distribution imply different electron configurations, close in energy, in an unstable ensemble of clusters. The negative spin gap describes a possible parallel spin pair binding instability. This picture implies the spontaneous ferromagnetism and phase (spin) separation into the domains in accordance with the Nagaoka theorem [38] (see Sect. 15.7.3). For the negative charge and spin gaps, one can introduce corresponding critical temperatures $T_c^P(\mu)$ and $T_s^F(\mu)$ versus chemical potential for boundaries or level crossing points between various phases derived from the condition that corresponding gaps disappear, i.e., $\Delta^{c,s}(T, \mu) = 0$. At these level crossing points, the change of sign for the order parameter is signaling the onset of a quantum phase transition between different phases in canonical ensemble. We define a phase transition also in the grand canonical ensemble in Sect. 15.7.5 as a weak singularity in susceptibility.

The obtained true first order phase transitions in finite systems with negative charge or spin gaps manifest the existence of a stable inhomogeneous phase. As a function of density (electron concentration), the electron pairs at low temperatures can undergo a gas-liquid type phase separation in charge and spin sectors. For example, arising inhomogeneities in the charge distribution for $\Delta^c < 0$ and $\Delta^s = 0$ imply the stable static heterostructure with ions

of different electron configurations of close energies as in the mixed valence compounds (see Sect. 15.7.6). These inhomogeneities in charge distribution in the absence of spin rigidity ($\Delta^s = 0$) can lead, in repulsive Hubbard clusters, to spontaneous phase transitions into a ferroelectric phase. The experimental and theoretical evidence suggests that electron charge pairing can cause both superconductivity and ferroelectricity, which is also a possible explanation for spontaneous polarization and “spontaneous symmetry breaking” at the nanometer-scale. When this happens, these particles that are made out of metal atoms no longer behave as if they are metallic.

These results are consistent with the nucleation of the “negative” charge gap in cuprates above [84]. At low temperatures, the dynamic picture for pair fluctuations between different electron configurations $\Delta^c < 0$ and $\Delta^s > 0$ is possible at relatively low temperatures in a spatially inhomogeneous, coherent state ($\Delta^s \equiv -\Delta^c$, see also Sect. 15.8.1). This condition resembles the global phase coherency also is consistent with the observation of nonlocal superconductivity at low excitation. At low energies and at higher energies, holes are still localized in an inhomogeneous “nanostripe” pattern [79]. In contrast, the negative spin gap describes the possible parallel spin pair binding instability. This picture of spin instability, that can be applied to manganites, implies spontaneous ferromagnetism and phase (spin) separation into domains in accordance with the Nagaoka theorem. For the negative gaps, one can introduce the critical temperatures $T_c^P(\mu)$ and $T_s^F(\mu)$ versus chemical potential for boundaries between various phases derived from the condition that the corresponding gaps disappear, i.e., $\Delta^{c,s}(T, \mu) = 0$.

15.7.5 The Charge and Spin Susceptibility Peaks

Conventional phase transitions occur at finite temperature, when the thermal fluctuations lead to a change in the physical state of a system. The susceptibility as an order parameter usually diverges in the thermodynamic limit, when approaching a critical temperature. For finite-size clusters, the electron density order parameter is related to charge susceptibility (compressibility), i.e., $\frac{\partial N}{\partial \mu}$. For a system undergoing a magnetic phase transition, this corresponds to magnetic susceptibility, i.e., $\frac{\partial s}{\partial h}$. The (second order) charge and spin fluctuation responses to electron or hole doping level (i.e., chemical potential μ) and an applied magnetic field (h), result in weak *saddle point* singularities (*crossovers*), which display clearly identifiable, prominent peaks in corresponding charge and spin density of states [52]. The energy difference in terms of μ between two consecutive susceptibility peaks serves as a natural order parameter in a grand canonical approach. The double peak structure in the density of states versus μ has a small, but nonzero weight inside the gaps at the infinitesimal temperature, $T \rightarrow 0$. Similar peak features in the atomic limit have recently been found also within the Greens function and equations of motion formalism in one-dimension [186]. We provide a comprehensive and systematic analysis of the model by considering all the

relevant response and correlation functions as well as thermodynamic quantities in the whole parameters space. To distinguish this from the canonical gaps at finite temperatures, we call them *pseudogaps* in the grand canonical approach. The opening of such distinct and separated pseudogap regions for the spin and charge degrees at the various fillings in μ space is indicative of the corresponding spin-charge separation. The crossover temperatures and phase boundaries for various transitions can be found by monitoring maxima and minima in charge and spin susceptibilities. We define the critical temperatures T_c and T^* in equilibrium as the temperatures at which the distances between the charge or spin susceptibility peaks or corresponding pseudogaps vanish (see Sect. 15.9.1). Notice that according to the given definition, the energy pseudogaps obtained in the grand canonical method are positive. This is a key difference from the canonical gaps, which can undergo an instability at $\Delta^{c,s} < 0$. For example, the Mott-Hubbard-like crossover temperatures for the metal-insulator transition is defined as a critical temperature at which two peaks merge into one with $\mu = \mu_+ = \mu_-$ and $\chi'_c(\mu) = 0$ and $\chi''_c(\mu) = 0$, i.e., as the temperature corresponding to a point of inflexion in $\chi_c(\mu)$, and similarly, one can define the Néel-like antiferromagnetic crossover temperature from the peak positions in spin susceptibility, χ_s [51, 52].

15.7.6 Charge and Spin Inhomogeneities

In condensed matter physics, two different paradigms for collectivized and localized electrons are typically applied. The first is the momentum or reciprocal space (k-space) picture, wherein electrons are delocalized throughout the material. Understanding these two separate paradigms is essential for a complete understanding of the physics of condensed matter, but rarely has it been as necessary to apply both these pictures as it has been to get insight into the electronic structure of the high-temperature superconductors (HTSCs) [115]. In this section, we review recent developments in the understanding of the relationship between the r-space and k-space electronic spectroscopies used to explore high-temperature superconductivity. The second is a local picture, in which one visualizes the quantum states of electrons in atomic orbitals or at impurity atoms in real space.

The origin of the local spatial inhomogeneities can be either electronic or structural, and can be driven by local lattice distortions and nonuniform electron concentration, due to doping irregularities. The charge distribution in superconducting cuprates is inhomogeneous both on macroscopic and microscopic scales: hole-rich clusters coexist with (hole-poor) neutral half filled insulating antiferromagnetic clusters. The characteristic energy scale responsible for the formation of the inhomogeneity is much larger than the superconducting energy scale for the formation of superconducting gap. It is therefore important to understand the nature of superconductivity at the atomic scale, since the description based on the BCS solution is applied only for strictly homogeneous systems. The developed cluster approach allows to

study the effects of intrinsic inhomogeneities on electron pairing instabilities and possible superconductivity. Moreover, these ideas for inhomogeneous superconductivity can be used to study phase separation and inhomogeneities in the context of nanostripes with frustrated structures, such as pyramids, tetrahedrons, octahedrons, etc.

The developed grand canonical approach can be applied to the understanding of the electron fluctuations and the spatial inhomogeneities at relatively high-temperature to model the behavior of the concentrated bipartite and frustrated structures. An ensemble of bipartite clusters at small and moderate U exhibits typical inhomogeneous behavior in its charge distribution. A normalized probability ω_N for the electron distribution in the grand canonical ensemble (as a function of temperature T) for various electron numbers N is the following

$$\omega_N = \sum_n e^{-\frac{E_{nN} - \mu N}{T}} \bigg/ \sum_{n,N} e^{-\frac{E_{nN} - \mu N}{T}}. \tag{15.26}$$

The calculated probabilities of electrons in competing configurations are shown in Fig. 15.8 for the ensemble of square plaquettes at $U = 4$. At low tem-

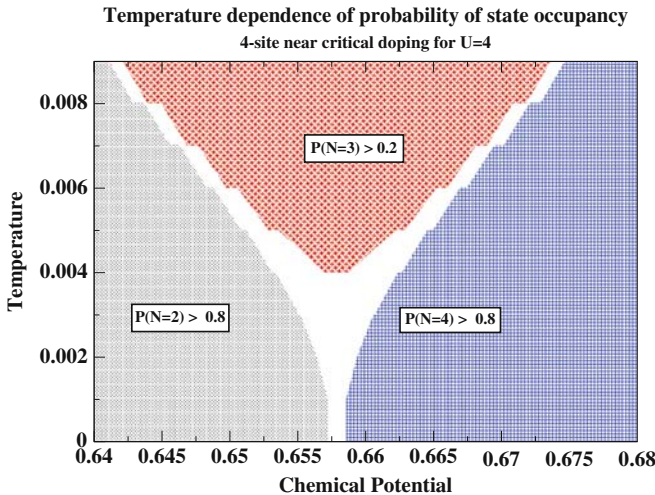


Fig. 15.8. Thermal occupation probabilities versus μ in grand canonical ensemble of 4-site clusters in the vicinity of quantum critical point μ_P for $\langle N \rangle \approx 3$ at $U = 4.0$. The phase separation at relatively low temperatures below $T \leq 0.0075$ manifests a significant suppression of $\langle N \rangle = 3$ clusters close to optimal doping $\mu_P = 6.557$. In this area, $\langle N \rangle = 2$ and $\langle N \rangle = 4$ clusters share equal weight probabilities, while at higher temperatures, $\langle N \rangle \approx 3$ also becomes thermodynamically stable. In equilibrium, the grand canonical value μ_P at optimal doping at $T = 0$ reproduces the result $\mu_P = (\mu_+ + \mu_-)/2$ for canonical approach, where μ_P is a quantum critical point for electron pairing, metallization and possible superconductivity

peratures and electron concentration close to μ_P , the clusters with $\langle N \rangle = 2$ and $\langle N \rangle = 4$ have equal probabilities, $\omega_2 = \omega_4 \approx 0.5$. In some circumstances, electron configurations in equilibrium can have close energies for the clusters in contact with a particle reservoir. This picture shows a mixture of ungapped and partially gapped states. As temperature increases, the ω_3 of $\langle N \rangle = 3$ clusters with unpaired spin gradually increases, while the probability of finding spin paired, hole-rich and hole-poor clusters decreases. In the canonical ensemble inhomogeneous superconductivity of locally generated pairs at rather low temperatures occurs due to thermal charge fluctuations (Sect.15.9.1). Interestingly, the thermal fluctuations in the canonical ensemble display “checkerboard” patterns, nanophase inhomogeneities, and nucleation of pseudogaps driven by temperature seen recently in nanometer and atomic scale measurements in HTSCs [79–82,84]. Microscopic spatial inhomogeneities and incoherent pairing pseudogaps in nanophases measured by STM correlate remarkably well with our predictions using an ensemble of square plaquettes and 2×4 nanoclusters [54]. Thus, exact calculations provide a conceptual solution for emergencies of (local) spatial inhomogeneities in microscopic canonical and the grand canonical ensembles.

15.7.7 The Coherent Charge and Spin Pairings

A considerable effort has been devoted to understand superconducting correlations experimentally and theoretically for the discrete excitation spectrum in nanoscale superconducting metallic grains, small size clusters, and nanoparticles [18, 19]. A key question in any such description is how results of the BCS theory for Cooper pairs are modified in finite systems. The behavior of such clusters near crossing degeneracies in a quantum coherent phase with minimal spin at low temperatures is somewhat similar to conventional the BCS superconductivity with the unique energy gap. The positive opposite spin pairing gap describes singlet coupling of spin pairs. On the other hand, the negative charge pairing manifests electron (charge) pairing in the charge sector. The obtained gaps can be used to characterize separate (partial) condensation (bosonization) of the charge and spin degrees. Indeed at rather low temperatures, the calculated positive spin gap Δ^s in grand canonical method becomes equal to the amplitude of a negative charge gap Δ^c derived in canonical method, $\Delta^s = |\Delta^c|$. Such behavior is similar to the existence of a single gap in the conventional BCS state. Following the Cooper pairs analogy in the BCS theory, one can call such an opposite spin (singlet) coupling and electron charge pairing as a spin coherent electron pairing in [49]. However, unlike the BCS theory, the charge gap differs from the spin pseudogap as temperature increases. For example, vanishing of a double peak structure in zero (field) magnetic spin susceptibility gives a critical temperature T_s^P , at which the spin pseudogap disappears. The canonical charge gap disappears at a critical temperature T_c^P given by $\Delta^c(T_c^P) = 0$.

The BCS-like coherent behavior and possible superconductivity of electrons with opposite spin pairing occur at rather low temperatures, below T_s^P (see Sect. 15.7.5), while electron charge pairing with negative gap, similar to what is observed in ferroelectricity in attractive Hubbard model [57], can be established at relatively higher temperatures, such as $T_s^P < T < T_c^P$. It is clear that consecutive pairing of electron charge and spin provides a new mechanism for reconciliation, condensation and full bosonization of electron charge and spin in the so-called coherent pairing state [49]. The positive spin gap calculated in the grand canonical approach implies a homogeneous spatial distribution for (opposite) spin. At relatively low temperatures, the homogeneous spatial (opposite spin) distribution dominates in the quantum coherent state. In contrast, the negative charge gap in the absence of coherency above T_c^P shows incoherent electron pairing with local (charge) inhomogeneities. Thus, the coherent pairing here is equivalent to the condition that much below T_s^P the spin gap Δ^s is equal to the magnitude of the charge gap, $|\Delta^c|$, while the phase coherence of pairs imply that these gaps have opposite sign, $\Delta^c = -\Delta^s$ (i.e., the charge and spin gaps of the equal magnitude but with the phase difference, η). The spin gap rigidity provides a full Bose-Einstein condensation of paired electron charge and opposite spins (singlets). We find a close analogy between the coherent electron pairing in clusters and real space singlet Cooper pairs in the BCS theory. Therefore, as in conventional superconductors at $T_c = 0$ the electron pairing and the phase coherence here occur simultaneously. The gaps in charge and spin sectors provide conditions for phase coherent transfer of singlet electron pairs across the local inhomogeneities. The coherent pairing also is somewhat similar to singlet pairs in resonance valence bond states or local inter-configuration fluctuations in mixed valence states [108, 187–189]. Unlike unconventional superconductors, the electron pairing and the onset of phase coherence in the BCS theory occur at the same temperature, i.e. T_c . Spontaneous phase separations and inhomogeneities for spin and charge degrees strongly depend on both, the Coulomb repulsion U and cluster topology. For instance in bipartite geometries, the charge separation leads to coherent pairing at small and moderate U , while Nagaoka-type ferromagnetic instabilities for spins occur at large U . In frustrated geometries, spontaneous transitions can lead to coherent pairing and saturated ferromagnetism for all U depending on the sign of the hopping term t , which can strongly affect the energy spectrum.

15.8 Ground State Properties

15.8.1 Bipartite Clusters

In the following paragraphs, we summarize the results for electron instabilities and phases shown in Fig. 15.9 for square 2×2 palquetts in canonical and grand canonical ensembles with average one hole off half filling at infinitesimal $T \rightarrow 0$ [49]. The vanishing of gaps at quantum critical points, $U_c = 4.584$

and $U_F = 18.583$, indicates energy level crossings and electron instabilities for charge and spin, respectively. The charge Δ^c and spin Δ^s gaps versus U in an ensemble of square clusters at $\langle N \rangle \approx 3$ exhibit (at infinitesimal temperature, $T \rightarrow 0$) the following phases. Inhomogeneous charge at $0 < U \leq U_c$ (Phase A): Charge and spin pairing gaps of equal amplitude $\Delta^s \equiv \Delta^P = -\Delta^c$ in the state with total spin $S = 0$ and (pseudospin) orbital momentum $L = 1$ describe Bose-Einstein condensation of electrons similar to BCS-like coherent pairing with a single energy gap, while neutral hole-poor (half filled) state with $S = 0$ and $L = 0$ corresponds to AF nanostripes. The electron pairing in Fig. 15.9 at $U < U_c$ takes place in the charge sector (canonical ensemble), while opposite spin pairing gap is derived in the grand canonical scheme [48]. Thus, electron charge and spin can be treated as two *interacting - homogeneous* canonical and *inhomogeneous* grand canonical entities (clusters). nanostripes. Homogeneous charge and spin at $U_c < U < U_F$ (Phase B): Mott-Hubbard-like insulator with $\Delta^c > 0$ and gapless $S = \frac{1}{2}$ and $L = \frac{1}{2}$ excitations describes a spin liquid behavior with $\Delta^s = 0$. Thus, the quantum spin liquids, observed in exact solution, are actually behave as charge insulators ($\Delta^c > 0$) with local magnetic moments that exhibit no magnetic order or spin excitation (energy) gap at zero temperature. Inhomogeneous spin at $U \geq U_F$ (Phase C): Parallel (triplet) spin pairing ($\Delta^s < 0$) displays $S = \frac{3}{2}$ the saturated ferromagnetism and $L = \frac{1}{2}$ at in Mott-Hubbard insulator for a positive charge gap, $\Delta^c > 0$. Notice, that incoherent opposite spin pairing $|\Delta^s| < \Delta^c$, different from the charge pairing at $U < U_c$, suggests spin-charge separation with low energy spin fluctuations at $U < U_F$. Notice, negative (*inhomogeneous*) spin gap in Fig. 15.9 at $U > U_F$ describes spin separation in canonical ensemble framework, while positive (*homogeneous*) charge gap in grand canonical scheme implies (*inhomogeneous*) electron-hole pairing.

In Figure 12 we have carried out exact numerical diagonalization and calculations of the charge gap and pairing in 8-site, 2×4 planar clusters, to illustrate similar effects on the earlier described properties of the 4-site clusters. Square clusters at weak and strong couplings share common important features with 2×4 ladders and other bipartite clusters [38]. The negative charge gap is equal to the opposite spin pairing gap, $\Delta^s = -\Delta^c$. This very similar to coherent BCS-like pairing in the ground state. For charge degrees at weak coupling, this gives an indication toward phase separation (i.e., *segregation*) on hole-rich (charge neutral) and hole-poor clusters. In contrast, at strong coupling, a negative spin pairing gap for parallel spins and positive charge gap reveal a ferromagnetic instability in accordance with the Nagaoka theorem. In large bipartite clusters at intermediate U , electrons behave differently from square clusters. For example, in 2×4 ladders, we found an oscillatory behavior of charge gap as a function of U [54]. The vanishing of the charge gap manifests the multiple level crossing degeneracy in a charge sector, which is indicative of a possible electron instability in bipartite clusters at moderate U . We shall see that the negative charge gap taking place also in the normal state (above T_c)

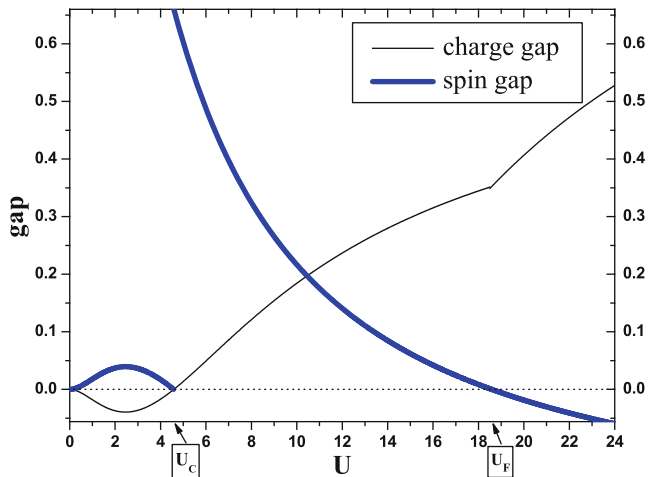


Fig. 15.9. Charge Δ^c and spin Δ^s gaps versus U in an ensemble of squares at $\langle N \rangle \approx 3$ and $T = 0$. Phase A: Charge and spin pairing gaps of equal amplitude at $U \leq U_c$ describe Bose-Einstein condensation of electrons similar to BCS-like coherent pairing with a single energy gap. Phase B: Mott-Hubbard-like insulator at $U_c < U < U_F$ leads to $S = \frac{1}{2}$ spin liquid behavior. Phase C: Parallel (triplet) spin pairing ($\Delta^s < 0$) at $U > U_F$ displays $S = \frac{3}{2}$ saturated ferromagnetism (see Sect. 15.8.1)

is a key point for the understanding of the mechanism of phase separation and unconventional superconductivity.

Exact diagonalization studies of energy gaps have been interpreted as favoring two or more critical points, which depend on the cluster size and cluster topology. The mechanisms of phase separation and various instabilities that we discuss in this work are somewhat different from what is usually discussed in the literature, such as in [95]. For example, in this many electron problem, there is a phase separation into two different many electron charge and spin states near half filling. At zero and low temperatures, due to the above instability, we find spontaneous transitions into states with local inhomogeneities in our cluster approach. In what follows, we identify this phenomenon as a spontaneous phase separation instability. In fact, such instabilities in a multi-parameter space are found to be rather common in various cluster topologies. In addition to this, we observe magnetic instabilities wherein there is competition, for example, between high and low spin states. Such phenomena appear to be generic and applicable to spin and charge degrees and are related to phase separation in the HTSCs and CMRs [56].

15.8.2 Tetrahedrons

For comparison with small bipartite clusters in Sect. 15.8.1, we consider here a minimal 4-site nonbipartite structure. A tetrahedron has a topology equivalent

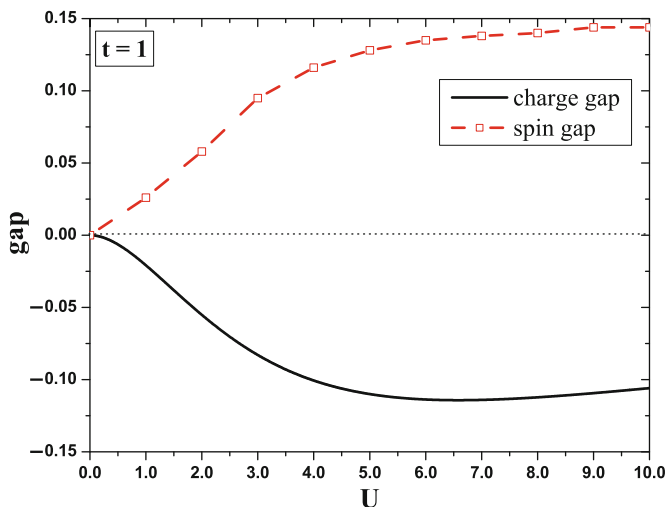


Fig. 15.10. Charge Δ^c and spin Δ^s gaps versus U in an ensemble of tetrahedrons at $t = 1$, $\langle N \rangle \approx 3$ and $T = 0.001$. Negative charge gap $\Delta^c < 0$ implies charge phase separation, while the positive, opposite spin pairing gap of equal amplitude $\Delta^s \equiv -\Delta^c$ describes Bose-Einstein condensation of electrons similar to BCS-like coherent pairing with a single (unique) energy gap. This negative charge gap, describing phase separation in tetrahedron clusters, is analogous to Phase A in square clusters [49]. The spin gap has been calculated using grand canonical approach

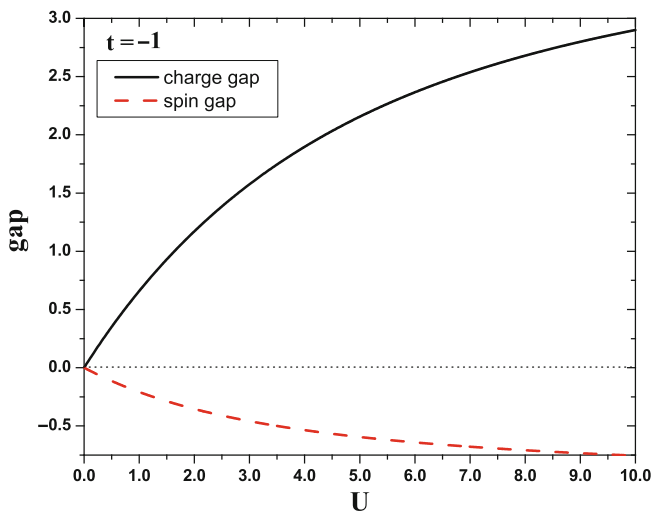


Fig. 15.11. The charge $\Delta^c > 0$ and parallel (triplet) spin $\Delta^s < 0$ gaps versus U in an ensemble of tetrahedrons at $t = -1$, $\langle N \rangle = 3$ and $T = 0.001$. The positive charge gap for all U describes Mott-Hubbard-like insulating behavior analogous to Phase C in square clusters [49]. The negative spin gap ($\Delta^s < 0$), coexisting with the charge pairing gap displays $S = \frac{3}{2}$ Nagaoka saturated ferromagnetism at all $U > 0$ (see Sect. 15.8.2). The charge gap has been calculated using grand canonical approach

to that of a square with the next nearest neighbor coupling ($t' = t$) and may be regarded as a primitive unit of a typical frustrated system. Nonbipartite systems, without electron-hole symmetry, exhibit the pairing instability that depends on the sign of t . Notice that the sign of t also leads to essential changes in the electronic structure. The tetrahedral clusters show pairing instabilities for charge degrees at $t = 1$ and spin degrees at $t = -1$ that optimize the negative charge and spin ($\Delta^c < 0, \Delta^s < 0$) gaps and corresponding condensation temperatures, $T_c^P(\mu)$ and $T_s^F(\mu)$ [190]. The negative gap in the canonical approach displays electron pairing $\Delta^P = |\Delta^c|$ instability for all U in this clusters. Figure 15.11 illustrates the charge and spin gaps at small and moderate U . The negative charge gap in Fig. 15.10 is indicative of the inhomogeneous charge redistribution and phase separation of electron charge into hole-rich (charged) and hole-poor (neutral) cluster configurations [52]. The phase diagram for $t = 1$ is similar to Phase A in Sect. 15.8.1, but unlike in square clusters valid for all U values applied for all U values. In contrast, the positive spin gap in the grand canonical approach $\Delta^s > 0$ corresponds to an uniform opposite spin distribution in Fig. 15.10. This BCS-like picture for charge and positive spin gaps of equal amplitude $\Delta^s \equiv \Delta^P = -\Delta^c$ at $\langle N \rangle \approx 3$, in analogy with the square clusters, will be called coherent pairing (CP) [49]. In equilibrium, the spin singlet background ($\chi_s > 0$) stabilizes phase separation of paired electron charge in a quantum CP phase. Fig. 15.10 illustrates the charge Δ^c and spin Δ^s gaps in tetrahedral clusters at $\langle N \rangle \approx 3, T \rightarrow 0$. The unique gap, $\Delta^s \equiv \Delta^P$ at $T = 0$, in Fig. 15.10 is consistent with the existence of a single quasiparticle energy gap in the BCS theory for $U < 0$ [57]. Positive spin gap for all U provides pair rigidity in response to a magnetic field and temperature (see Sect. 15.9.1). Notice that the coherent pairing exists also at large U where Nagaoka theorem for nonbipartite clusters with a specific sign of t can be applied. The stability of minimal spin $S = 0$ (singlet) state in tetrahedron at $t = 1$ is consistent with stability of nonmaximum (unsaturated) spin in the Nagaoka problem. This result shows that Nagaoka instability toward spin flip at large U in frustrated lattices with $t = 1$ can be associated with the BCS-like coherent pairing valid for general U .

The negative spin gap Δ^s in Fig. 15.11 shows the canonical energy differences between $S = \frac{3}{2}$ and $S = \frac{1}{2}$ electron configurations. Correspondingly, the positive charge gap $\Delta^c > 0$ in a stable Mott-Hubbard-like state is derived using grand canonical energies [52]. As in bipartite square clusters, grand canonical positive charge gap $\Delta^c > 0$ is different (incoherent) from the parallel spin pairing gap, $\Delta^s < 0$. Thus, the phase diagram for $t = -1$ with $\Delta^c \neq -\Delta^s$ is similar to the Phase C in Sect. 15.8.1, but applied for all U values in the phase diagram. The negative spin gap for all couplings implies parallel spin pairing and Nagaoka-like saturated ferromagnetism with maximum spin in the entire range of U . Spin-charge separation is considered to be one of the key properties of the correlated electrons that distinguishes $t = -1$ from $t = 1$. Such behavior at $t = -1$ is accompanied by spin-charge separation and formation

of the magnetic (spatial) inhomogeneities or domain structures [52] in a wide range of parameters. The main features discussed for 2×2 clusters are reproduced also for larger bipartite clusters. For example, Fig. 15.12 illustrates the variation of the charge gap versus U in linked 2×4 two-dimensional clusters. Oscillatory behavior at small U manifests the quantum critical points and corresponding phase transitions.

15.8.3 Square Pyramids

From the early days of high-temperature superconductivity, the question of a possible role of apical sites in p-type superconductors has been controversial. Here, we try to draw a closer connection to HTSCs perovskites and consider an ensemble of square pyramids or octahedron of perovskite structure [53, 56]. There is no significant influence of localized electron charge of the apical site on electron pairing and possible superconductivity associated with CuO_2 planes in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. When excess apex does not exist, i.e., $\delta = 0$, this system is an insulator. However, when excess apex oxygen is introduced, hole carriers are supplied into the CuO_2 planes, and the material shows superconductivity [191].

Figure 15.13 shows the charge gap at fixed $U = 3$ and $\langle N \rangle \approx 4$ under the variation of the coupling term c between the plane and the apex atoms. This picture gives surprisingly plausible evidence for understanding the detrimental role of excess electrons on charge pairing for possible distortions of pyramidal crystalline structure in perovskites. In Fig. 15.13, the strong distortion of the pyramid structure for $c = 0$ (with reduced coordination number) reproduces a charge pairing gap in planar square geometries. At $\langle N \rangle \approx 4$, the electron is localized, and there is no charge transfer from apex atom in an ensemble of pyramid clusters at $c = 0$. The negative charge gap, identical to the spin gap, exists only for $c \leq c_0$, where $c_0 = 0.35$ (in units of t) is a quantum critical point for level crossing degeneracy. Calculated electron distribution, as a function of c , shows that electron charge residing on the apical site does not contribute to the pairing whenever c is less than c_0 . Thus, we find a direct correlation between the size of the energy gap characterizing the superconducting state and a modulation of the atomic positions in the seen in high T_c superconductors [127–132] (see Sect. 15.3.4). The coupling in the pyramidal structure at $c < c_0$ for $\langle N \rangle \approx 4$ leads to a charge pairing instability with negative charge and positive spin gaps of equal amplitude as seen in square clusters at $\langle N \rangle \approx 3$ in Sect. 15.8.1. In contrast, at $c > c_0$, the induced charge gap driven by c leads to the electron hole pairing and transition into an insulating Mott-Hubbard-like behavior with $\Delta^c > 0$. Thus, the inter-site coupling c in square pyramids can be beneficial or detrimental for the electron pairing or ferromagnetism. A similar approach can be applied to understand the detrimental effect of the next nearest coupling term on pairing for the copper atom surrounded by oxygens in planar face centered square (fcs) geometry [49] or octahedral clusters associated with the perovskites [56]. Our results strongly support the

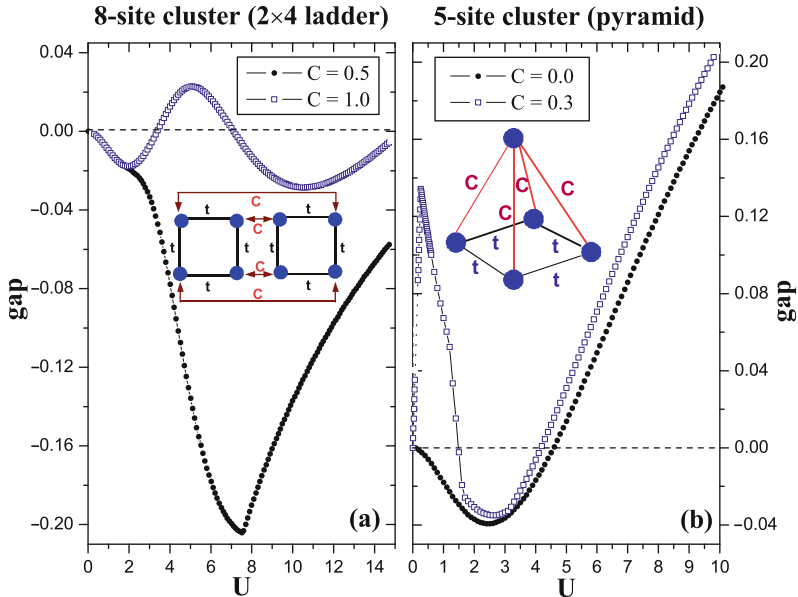


Fig. 15.12. The zero temperature charge gap at one hole off half filling in square pyramid and linked 2×4 clusters (at $T = 0$) as a function of the interaction strength U at various couplings c between the squares and the square to apex atom respectively. The negative in oscillatory behavior of charge gap displays at small and moderate U a charge pairing instability in deformed pyramidal or octahedral structures. This also underlines the role of vertex coupling and the quasi two-dimensional character of pairing which may be related to HTSC perovskites

idea that pyramids, fcs, octahedrons are all examples of nonbipartite clusters where particle-hole symmetry is absent (broken). In frustrated octahedrons the coupling to the vertex, which breaks particle-hole symmetry, has a similar detrimental effect on charge pairing. We have found a range of parameters and conditions necessary for Bose-Einstein condensation and possible superconductivity. At higher values of U , Nagaoka-type magnetic transitions are also observed for the appropriate sign of ($t < 0$).

15.9 Phase T - μ Diagram

15.9.1 Tetrahedrons at $t = 1$

The charge and spin susceptibility peaks in clusters, reminiscent of the singularities in infinite systems, display an extremely rich phase diagram at finite temperatures. The realization of a high transition temperature superconductivity in clusters and bulk systems depends on the interaction strength U ,

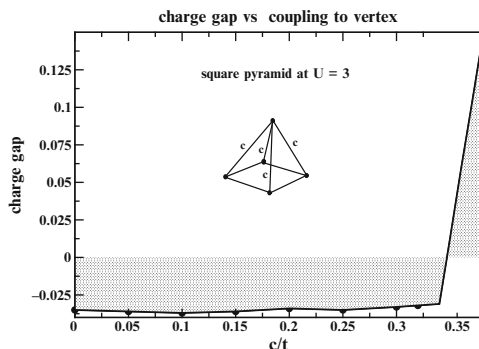


Fig. 15.13. Charge gap Δ^c versus coupling c between the apex atom and the four base atoms in deformed square pyramid ($t = 1$) for one hole of half filling, $\langle N \rangle = 4$, $U = 3$ and $T = 0.01$. Charge and spin pairing gaps of equal amplitude $\Delta^s \equiv \Delta^p = -\Delta^c$ at $c \leq 0.35$ imply coherent pairing, while $\Delta^c > 0$ and $\Delta^s < 0$ at $c \geq 0.35$ correspond to a ferromagnetic insulator for $S = \frac{1}{2}$. The negative gap associated with superconducting pairing in CuO_2 planes shows modulation with the distance of the copper atoms and the apical oxygen that anticorrelates with the coupling c

doping, the detailed nature of the crystal structure. The critical temperatures for various pairing instabilities in frustrated clusters also strongly depend on the sign of the hopping (t) term. Figure 15.14 for $t = 1$ illustrates a number of nanophases, defined in [52, 54], for the ensemble of tetrahedral and square (see inset) clusters at ($U = 4$) found earlier in tetrahedral and bipartite 2×2 and 2×4 clusters at moderate $U = 4$ values [49]. This diagram quite well captures the essential electron charge and spin pairing instabilities at finite temperatures. The curve $\mu_+(T)$ below T_c^P signifies the onset of charge pair condensation. The calculated susceptibility peaks in Fig. 15.14 correspond to the pseudogap crossover temperature T^* , which describes the transition from AF into spin liquid behavior. As temperature is lowered below T^* , a spin pseudogap is opened up first, as seen in nuclear magnetic resonance (NMR) experiments [54], followed by the gradual disappearance of the spin excitations, consistent with the suppression of low-energy excitations in the HTSCs probed by STM and ARPES [81–86]. In Fig. 15.13, curve T^* is just an equilibrium temperature arising from fluctuations in the order parameter (spin gap) that establishes (spin singlet) electron pairing. In the pseudogap state, the symmetry is spontaneously broken below $\mu_+(T)$, which therefore corresponds to a phase transition (spinodal region). The opposite spin CP phase, with fully gapped collective excitations, begins to form at $T \leq T_s^P$, and spin pairing rigidity gradually grows upon lowering of the temperature. As temperature decreases, both charge and spin pseudogaps merge into one gap at zero temperature. Therefore, at sufficiently low temperatures, this leads to the BCS-like coherent coupling of electron charge to bosonic (spin singlet)

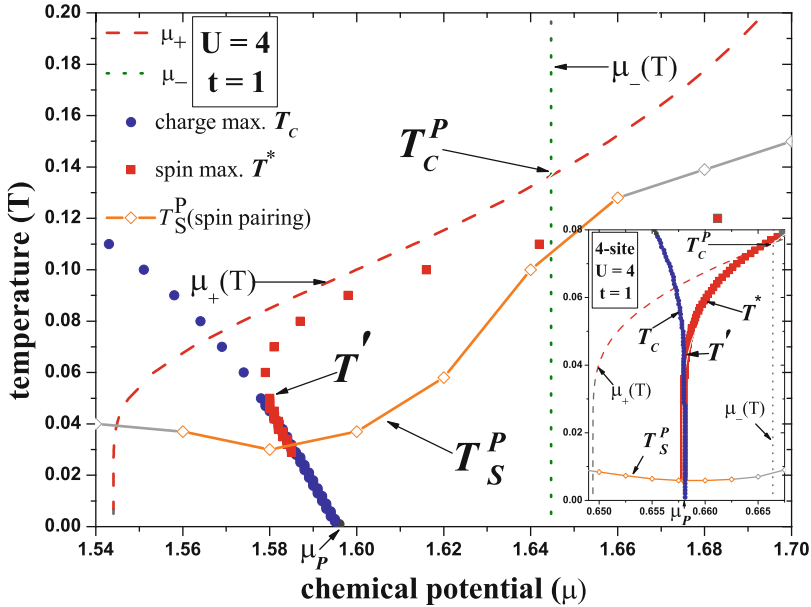


Fig. 15.14. The T - μ phase diagram of tetrahedrons without electron-hole symmetry at optimally doped $\langle N \rangle \approx 3$ regime near $\mu_P = 1.998$ at $U = 4$ and $t = 1$ illustrates the condensation of electron charge and onset of phase separation for charge degrees below T_C^P . The incoherent phase of preformed pairs with unpaired opposite spins exists above T_S^P . Below T_S^P , the paired spin and charge coexist in a coherent pairing phase. The spin and charge susceptibility peaks, denoted by T^* and T_c , define spin and charge pseudogaps calculated in the grand canonical ensemble, while phase boundaries $\mu_+(T)$ and $\mu_-(T)$ are evaluated in the canonical ensemble. Charge and spin peaks reconcile at $T \sim T'$, while χ^c peak below T_S^P signifies metallic (charge) spin liquid state (see inset for square cluster and [54]). The negative charge gap in canonical ensemble equal to the positive spin gap supports idea of spontaneous transition into the coherent pairing state below T_S^P

excitations (see Sect. 15.8.2). However, the spin gap is more fragile, and as temperature increases, it vanishes at T_S^P , while charge pseudogap is survived until higher temperature, T_C^P .

The charge inhomogeneities [79, 80] in hole-rich and charge neutral *spinodal* regions between μ_+ and μ_- are similar to those found in the ensemble of squares and resemble important features seen in the HTSCs. Pairing and transfer of holes are a consequence of the existence of an inhomogeneous background. In the absence of direct contact between clusters, the inhomogeneities in the grand canonical approach establish a transfer of paired electrons via this (thermal) bath media. Figure 15.14 shows the presence of bosonic modes below $\mu_+(T)$ and T_S^P for paired electron charge and opposite spin, respectively. This picture suggests condensation of electron charge and spin at various crossover temperatures, while condensation in the BCS theory occurs at a unique T_C value. The coherent pairing of holes here is a consequence of the

existence of homogeneous opposite spin pairing background, consistent with the STM measurements [112]. This led us to conclude that T_s^P can be relevant to the superconducting condensation temperature in the HTSCs. In the absence of spin pairing above T_s^P , the pair fluctuations between the two lowest energy states becomes incoherent. The temperature driven spin decoupling above T_s^P resembles an incoherent pairing (IP) phase seen in the HTSCs [80–82, 84]. Instead of becoming superconducting, the charged pairs without spin rigidity above T_s^P , coexist with hole-poor nanoclusters in (density) nonuniform IP state similar to a ferroelectric phase [57]. This result is interesting for the understanding of the possible nature of ferroelectricity as an electron pairing without coherency. This result also shows a relationship between incoherent pairing behavior in high temperature superconductors and some inhomogeneous materials that exhibit the ferroelectric effect at rather high temperatures. The unpaired weak moment, induced by a field above T_s^P , agrees with the observation of competing dormant magnetic states in the HTSCs [82]. The charge susceptibility peak positions, denoted by T_c , in Fig. 15.13 at equilibrium, are arising from fluctuations in the order parameter (charge gap) that establishes charge pairing. The coinciding χ^s and χ^c peaks in the vicinity of T' show full reconciliation of charge and spin degrees seen in the HTSCs above T_s^P . Above this temperature, we find strong reconciliation of spin and charge susceptibilities in bipartite clusters, compared to frustrated clusters (see inset in Fig. 15.14). In the absence of electron-hole symmetry in the tetrahedrons, the reentrant phenomenon can be observed at low temperatures [49]. In Fig. 15.14, as temperature increases near optimal doping $\mu \leq \mu_P$, clusters undergo a transition from a CP phase to a Mott-Hubbard-like behavior. Notice, the charge and spin pairing do not disappear in the overdoped regime for $\mu \geq \mu_P$ but are governed predominantly by the physics of antiferromagnets at half filling. In contrast, in the underdoped regime at low temperatures, the charge pairing pseudogap gradually approaches the spin pseudogap as in the conventional BCS theory. Notice that this diagram exhibits the main phases in HTSCs, including the spin glass phase $\Delta^P = 0$ and $\Delta^s > 0$ in the overdoped region at $\mu > \mu_P$ and proximity to Mott-Hubbard insulator (non Fermi liquid) in the underdoped region $\mu < \mu_P$. Figure 15.14 indicates also the temperature driven transition into the Mott-Hubbard-like dielectric with a positive charge gap ($\mu_+ > \mu_-$) above T_c^P and formation of coherent pairs at low temperatures. Notice that χ_s peak terminates at T_s^P , while χ_c peak reaches at zero temperature, μ_P , the QCP inside the superconducting dome.

15.10 Conclusion

Electron pairing, ferromagnetism and ferroelectricity in bulk materials, nanostructures, and nanoparticles are displaying a high sensitivity to the local electronic correlations and geometric structures. Many collective phenomena,

observed from approximate treatments of “large” systems, we also found in our exact solutions in nanoclusters, which do interact thermally. It is very appealing that both the intrinsic properties and their experimental manifestations (e.g., tunneling studies and spectroscopy) involve a wide range of fundamental issues, for example, phase transitions and symmetry breaking, superconductivity and superfluidity, etc. Exact solution of Hubbard model in different topologies provides important insights into several many-body problems in condensed-matter physics. One can draw a simple analogy between the quarks in nuclear physics and a quantum gas of clusters in condensed matter. The discovery that building block of nucleons and other strongly interacting particles are fractionally charged quarks has had a profound influence across the high energy physics. Similarly quantum gas of nanoclusters as building blocks of large inhomogeneous systems can describe the complexity of many body phenomena in condensed matter. The inhomogeneities and pair modulation found in STM experiments and exact calculations of clusters are promising for the description of nanostripe (geometric) patterns in high- T_c cuprates and magnetic domain structures in manganites. Of course this approach, being perfectly suitable for small nanoclusters (for which it is exact), has certain limitations when applied to bulk solids since coupling between building (cluster) blocks, in principle, has to be taken into account. The ensemble of weakly interacting quantum gas of clusters, can be a especially useful for understanding complex phenomena in inhomogeneous concentrated magnetic and superconducting materials, where coupling between clusters c is much less of the relevant energy scale within individual clusters, t . The small nanoclusters capture the magnetic and electron pairing instabilities in nanostructures and respective bulk materials and therefore can be useful for the prediction of electron coherent and incoherent pairings, ferroelectricity in nanoclusters, nanostructures, and possible superconductivity in nanoparticles [18, 19, 57]. There is no reason to expect that correlated electrons in solids should comply with approximate Landau Fermi liquid or mean field BCS paradigms. The quest for new paradigms beyond the familiar approximate theories fueled by the unexpected discoveries of exotic phenomena in high T_c superconductivity [192]. The exact calculation of Hubbard model in real space has thus emerged as an appealing alternative to study many body effects in large concentrated inhomogeneous systems. The principal conclusion is that the exact solution for optimal inhomogeneities mimicked in small clusters can target essential features relevant to existing inhomogeneities in materials at the nanoscale level. We found charge and opposite spin pairing gaps of equal amplitude in the ground state similar to the coherent pairing in conventional BCS theory. However, separate Bose-Einstein condensation of electron charge and spin degrees with two consecutive transitions into coherent pairing suggests a mechanism of full Bose-Einstein condensation different from the prediction of the BCS-like condensation scenario with a unique critical temperature. This picture with two temperature scales is consistent with the existence of two different energy scales for electron charge and spin pairing

condensation temperatures in the HTSCs [79–82, 84]. The analysis of phase diagrams in small clusters provides interesting insight into electron condensation, magnetism, and ferroelectricity at finite temperatures. Exact theory displays a number of inhomogeneous, coherent and incoherent nanophases seen recently by STM in numerous nanomaterials, assembled nanoclusters, and ultra-cold fermionic atoms [153,192]. For example, we found that these properties are strongly depend on the sign of the hopping term and cluster geometry rather than the Coulomb interaction. Thus, exactly solvable clusters may offer some important answers on long standing problems in condensed matter physics. This principal disentanglement of the nontrivial physics from complex structures has reduced the mystery of HTSC to a tangible problem of local charge and spin density formation.

Acknowledgements

The original work along the line presented in this manuscript was done in collaboration with Jim Davenport, Kalum Palanadage, Tun Wang, You-Ling, Ling-Yan Chen, Chung-Lei Lu. One of us (ANK) wants to thank his coauthors, Avak Saakyan, Petros Ovnianian, Nerses Ananikian, Lev Ananikyan, and Vahagn Pogosyan in Yerevan Physics Institute, where this work started. We would like also to acknowledge Daniil Khomskii, Valery Pokrovsky, Klim Kugel, Edward Rezayi, Jose Rodriguez, Guo-Meng Zhao and Alexander Balatsky for helpful discussions.

References

1. M. He, A. Simon, V. Duppel, Z. Anorg. Allg. Chem. **630**, 535 (2004)
2. J. Kröger, N. Néel, L. Limot, J. Phys.: Condens. Matter **20**, 223001 (2008)
3. P. Jena, B.K. Rao, R.M. Nieminen, Solid State Commun. **59**, 509 (1986)
4. C. Bréchnignac, Ph. Cahuzac, F. Carlier, M. de Frutos, Phys. Rev. Lett. **64**, 2893 (1990)
5. K. Sattler, J. Mühlbach, O. Echt, P. Pfau, E. Recknagel, Phys. Rev. Lett. **47**, 160 (1981)
6. S. Iijima et al., Nature (London) **354**, 56 (1991)
7. R. Moro, X. Xu, S. Yin, W.A. de Heer, Science **300**, 1265 (2003)
8. J. Hemberger et al., Nature (London) **434**, 364 (2005)
9. J.A. Alonso, M.J. López, L.M. Molina, F. Duque, A. Mañanes, Nanotechnology **13**, 253 (2002)
10. S. Heinze et al., Science **288**, 1805 (2000)
11. W.P. Halperin, Rev. Mod. Phys. **58**, 533 (1986)
12. D.C. Ralph, C.T. Black, M. Tinkham, Phys. Rev. Lett. **74**, 3241 (1995)
13. D.C. Ralph, C.T. Black, M. Tinkham, Phys. Rev. Lett. **78**, 4087 (1997)
14. S. Guéron, M.M. Deshmukh, E.B. Myers, D.C. Ralph, Phys. Rev. Lett. **83**, 4148 (1999)
15. R. Desmicht, G. Faini, V. Cros, A. Fert, F. Petroff, A. Vaurès, Appl. Phys. Lett. **72**, 386 (1998)

16. V.Z. Kresin, Y.N. Ovchinnikov, Phys. Rev. B **74**, 024514 (2006)
17. C.T. Black, Ph.D. Thesis, Harvard University, 1996
18. R. Moro, S. Yin, X. Xu, W.A. de Heer, Phys. Rev. Lett. **93**, 086803 (2004)
19. X. Xu, S. Yin, R. Moro, W.A. de Heer, Phys. Rev. Lett. **95**, 237209 (2005)
20. H. Shiba, P.A. Pincus, Phys. Rev. B **5**, 1966 (1972)
21. L.M. Falicov, R.H. Victora, Phys. Rev. B **30**, 1695 (1984)
22. A.N. Kocharian, J.H. Sebold, Phys. Rev. B **53**, 12804 (1996)
23. L. Tan, J. Calaway, Phys. Rev. B **46**, 5499 (1992)
24. J. Calaway, Phys. Rev. B **35**, 8723 (1987)
25. R. Schumann, Ann. Phys. **11**, 49 (2002)
26. R. Schumann, Ann. Phys. **17**, 221 (2008)
27. E.H. Lieb, F.Y. Wu, Phys. Rev. Lett. **21**, 192 (1968)
28. L.E. Lieb and D.C. Mattis, Phys. Rev. **125**, 164 (1962)
29. M. Takahashi, Prog. Theor. Phys. **42**, 67 (1972)
30. M. Takahashi, Prog. Theor. Phys. **52**, 103 (1974)
31. M. Takahashi, M. Shiroishi Phys. Rev. B **65**, 165104 (2002)
32. N. Kawakami, T. Usuki, A. Okiji, Phys. Lett. A **137**, 287 (1989)
33. J. Carmelo, P. Horsch, P.A. Bares, A.A. Ovchinnikov Phys. Rev. B **44**, 9967 (1991)
34. C. Yang, A.N. Kocharian, Y.L. Chiang, J. Phys.: Condens. Matter B **12**, 7433 (2000)
35. Zachary N.C. Ha, *Quantum many-body systems in one dimension, Series on advances in statistical mechanics*, vol. 12, (World Scientific, Singapore, 1996)
36. T. Usuki, N. Kawakami, A. Okiji, J. Phys. Soc. Japan **59**, 1357 (1990)
37. N. Kawakami, S.-K. Yang, Phys. Rev. **44**, 7844 (1991)
38. Y. Nagaoka, Phys. Rev. **147**, 392 (1966)
39. S. Sarker, C. Jayaprakash, H.R. Krishnamurthy, W. Wenzel, Phys. Rev. B **43**, 8775 (1991)
40. C. Jayaprakash, H.R. Krishnamurthy, S. Sarker, Phys. Rev. B **40**, 2610–2613 (1989)
41. A.N. Kocharian, C. Yang, Y.L. Chiang, T.Y. Chou, Internat. J. Mod. Phys. B **17**, 5749 (2003)
42. A.N. Kocharian, C. Yang, Y.L. Chiang, Phys. Rev. B **59**, 7458 (1999)
43. A.N. Kocharian, C. Yang, Y.L. Chiang, Physica C **341–348**, 253 (2000)
44. A.N. Kocharian, C. Yang, Y.L. Chiang, T.Y. Chou, Internat. J. of Mod. Phys. B **17**, 3363 (2003)
45. A.N. Kocharian, C. Yang, Y.L. Chiang, T.Y. Chou, Internat. J. of Mod. Phys. B **19**, 2225 (2005)
46. E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994)
47. K.W. Binder, D.W. Herrmann, *Monte Carlo Simulations in Statistical Physics, Springer Series in Solid-State Sciences*, vol. 80, (Springer, Berlin, 1992)
48. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, Phys. Lett. A **373**, 1074 (2009)
49. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, Phys. Rev. B **78**, 075431 (2008)
50. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, arXiv:cond-mat.str-el/0510609v1 (2005) (unpublished)
51. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, J. Magn. Magn. Mater. **300**, e585 (2006)

52. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, Phys. Rev. B **74**, 024511 (2006)
53. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, Phys. Lett. A **364**, 57 (2007)
54. G.W. Fernando, A.N. Kocharian, K. Palandage, T. Wang, J.W. Davenport, Phys. Rev. B **75**, 085109 (2007)
55. K. Palandage, G.W. Fernando, A.N. Kocharian, J.W. Davenport, J. Comput.-Aided. Mater. Des. **14**, 103 (2007)
56. G.W. Fernando, A.N. Kocharian, K. Palandage, J.W. Davenport, Phys. Rev. B **80**, 014525 (2009)
57. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, Ultramicroscopy **109**, 1066 (2009)
58. I. Bloch et al., Rev. Mod. Phys. **80**, 885 (2008)
59. A.M.C. Souza, H. Herrmann, Phys. Rev. B **77**, 085416 (2008)
60. H. Shiba, M. Ogata, Prog. Theor. Phys. Supplement **108** 265 (1992)
61. M. Ogata, T. Sugiyama, H. Shiba, Phys. Rev. B **43**, 8401 (1991)
62. J.D. Denlinger, G.-H. Gweon, J.W. Allen, C.G. Olson, J. Marcus, C. Schlenker, and L.-S. Hsu, Phys. Rev. Lett. **82**, 2540 (1999)
63. G.-H. Gweon, J.D. Denlinger, J.W. Allen, C.G. Olson, H. Höchst, J. Marcus, and C. Schlenker, Phys. Rev. Lett. **85**, 3985 (2000)
64. J.W. Allen et al., in *The Hubbard model*, ed. by R.Z. Baeriswyl et al. (Plenum, New York, 1995), p. 357
65. K.B. Blagoev, K.S. Bedell, Phys. Rev. Lett. **79**, 1106 (1997)
66. M. Yamanaka, M. Oshikawa, I. Affleck, Phys. Rev. Lett. **79**, 1110 (1997)
67. F.D.M. Haldane, J. Phys. C **14**, 2585 (1981)
68. J. Voit, Rep. Prog. Phys. **58**, 977 (1995), see references therein.
69. T. Lorenz et al., Nature (London) **418**, 614 (2002)
70. P.W. Anderson, Science **288**, 480 (2000)
71. P.W. Anderson, Science **235**, 1196 (1987)
72. A. Salem, G. Jakob, H. Adrian, Physica C: Superconductivity **415**, 62 (2004)
73. P.W. Anderson, Phys. Rev. Lett. **67**, 2092 (1991)
74. S.M. Hayden et al., Nature (London) **429**, 531 (2004)
75. S. Rabello, Q. Si, Europhys. Lett. **60**, 882 (2002)
76. Q. Si, Physica C **341–348**, 1519 (2000)
77. Ch. Renner et al., Nature (London) **416**, 518 (2002)
78. T. Lottermoser, M. Fiebig, Phys. Rev. B **70**, 220407 (2004)
79. Y. Kohsaka et al., Science **315**, 1380 (2007)
80. T. Valla et al., Science **314**, 1914 (2006)
81. A.C. Bódi, R. Laiho, E. Lähderanta, Physica C **411**, 107 (2004)
82. H.E. Mohottala et al., Nature Materials **5**, 377 (2006)
83. M.C. Boyer et al., Nat. Phys. **3**, 802 (2007)
84. K.K. Gomes et al., Nature (London) **447**, 569 (2007)
85. A.N. Pasupathy et al., Science **320**, 196 (2008)
86. Vershinin, S. Misra, Y. Abe, S. Ono, Y. Ando and A. Yazdani, Physica C **408–410**, 764 (2004)
87. R.E. Cohen, Nature (London) **358**, 136 (2005)
88. J. Zhao et al., Nature Materials, **7**, 953 (2008)
89. S.A. Kivelson et al., Nature Materials **7**, 927 (2008)
90. G.V.M. Williams, J.L. Tallon, J.W. Loram, Phys. Rev. B **58**, 15053 (1998)

91. A. Damascelli, Z. Hussain, Z.-X. Shen, *Rev. Mod. Phys.* **75**, 473 (2003)
92. J.S. Schilling, S. Klotz, in *Physical properties of high temperature superconductors*, vol. 3, ed. by D.M. Ginzberg (World Scientific, Singapore, 1992), p. 59
93. Y. Zha, V. Barzykin, D. Pines, *Phys. Rev. B* **54**, 7561 (1996)
94. P.B. Visscher, *Phys. Rev. B* **10**, 943 (1974)
95. V.J. Emery, S.A. Kivelson, H.Q. Lin, *Phys. Rev. Lett.* **64**, 475 (1990)
96. J.W. Alldredge et al., *Nat. Phys.* **4**, 319 (2008)
97. Y. Kohsaka et al., *Science* **315**, 1380 (2007)
98. J. Lee et al., *Nature (London)* **442**, 546 (2006)
99. K.M. Lang et al., *Nature (London)* **415**, 412 (2002)
100. J.E. Hoffman et al., *Science* **266**, 455 (2002)
101. J.E. Hoffman et al., *Science* **295**, 466 (2002)
102. M.T. Woodside et al., *Science* **296**, 1098 (2002)
103. D. Wohlleben, J. Rohler, *J. Appl. Phys.* **55**, 1904 (1984)
104. A.N. Kocharian, *JETP Lett.* **45**, 99 (2009)
105. C. Felser, *J. Alloys and Compounds*, **262–263**, 87 (1997)
106. J. Goodenough, *Phys. Rev.* **100**, 564 (1955)
107. P. Schlottmann, *J. Appl. Phys.* **87**, 5022 (2000)
108. A.N. Kocharian, D.I. Khomskii, [*Zh. Eksp. Teor. Fiz.*, **71** (1976) 767] *Sov. Phys. JETP* **44**, 404 (1976)
109. A.N. Kocharian, P.S. Ovnianian, *Sov. Phys. JETP* **47**, 326 (1978)
110. M.C. Boyer et al., *Nat. Phys.* **3**, 802 (2007)
111. W.D. Wise et al., *Nature Physics* **4**, 696 (2008)
112. E.W. Hudson et al., *Science* **285**, 88 (1999)
113. K. McElroy et al., *Nature (London)* **422**, 520 (2003)
114. Y. Murakami et al., *Nature (London)* **423**, 965 (2003)
115. K.M. Shen, J.C. Davis, *Materials Today* **11**, 14 (2009)
116. H. Ding, M.R. Norman, J.C. Campuzano, M. Randeria, A.F. Bellman, T. Yokoya, T. Takahashi, T. Mochiku, K. Kadowaki, *Phys. Rev. B* **54**, R9678 (1996)
117. A. Damascelli, Z. Hussain, Z.-X. Shen, *Rev. of Modern Phys.* **75**, 473 (2003)
118. A.G. Loeser et al., *Science* **273**, 325 (1996)
119. M. Kugler, O. Fischer, Ch. Renner, S. Ono, Y. Ando, *Phys. Rev. Lett.* **86**, 4911 (2001)
120. K. Tanaka et al., *Science* **314**, 1910 (2006)
121. W.S. Lee et al., *Nature (London)* **450**, 81 (2007)
122. N.-C. Yeh, C.-T. Chen, C.-C. Fu, P. Seneor, Z. Huang, C.U. Jung, J.Y. Kim, M.-S. Park, H.-J. Kim, S.-I. Lee, K. Yoshida, S. Tajima, G. Hammerl, J. Mannhart, *Physica C: Superconductivity* **367**, 174 (2002)
123. S. Hüfner, M.A. Hossain, A. Damascelli, G.A. Sawatzky, *Rep. Prog. Phys.* **71**, 062501 (2008)
124. J.Y.T. Wei, J. Ngai, P. Morales, *Proc. SPIE* **5932**, 593210 (2005)
125. Ch. Niedermayer, C. Bernhard, U. Binninger, H. Glöckler, J.L. Tallon, E.J. Ansaldo, J.I. Budnick, *Phys. Rev. Lett.* **71**, 1764 (1993)
126. X.J. Zhou et al., *Science* **286**, 268 (1999)
127. I. Martin, A.V. Balatsky, *Physica C (Amsterdam)* **357–360**, 46 (2001)
128. E. Abrahams, C.M. Varma, *Proc. Natl. Acad. Sci. USA* **97**, 5714 (2000)

129. J.A. Slezak, J. Lee, M. Wang, K. McElroy, K. Fujita, B.M. Andersen, P.J. Hirschfeld, H. Eisaki, S. Uchida, J.C. Davis, Proc. Natl. Acad. Sci. USA **105**, 3203 (2008)
130. M.R. Norman, PNAS **105**, 3173 (2008)
131. K. McElroy et al., Science **309**, 1048 (2005)
132. T.S. Nunner, B.M. Andersen, A. Melikyan, P.J. Hirschfeld, Phys. Rev. Lett. **95**, 177003 (2005)
133. E. Arrigoni, S.A. Kivelson, Phys. Rev. B **68**, 180503 (2003)
134. J. Eroles, G. Ortiz, A.V. Balatsky, A.R. Bishop, Inter. J. Mod. Phys. **15**, 2833 (2001)
135. J. Eroles, G. Ortiz, A.V. Balatsky, A.R. Bishop, Europhys. Lett. **50**, 540 (2000)
136. Q. Cui, C.-R. Hu, J.Y.T. Wei, K. Yang, Phys. Rev. B **73**, 214514 (2006)
137. P. Fulde, R. A. Ferrell, Phys. Rev. **135**, A550 (1964)
138. A.I. Larkin, Y.N. Ovchinnikov [Zh. Eksp. Teor. Fiz. 47, 1136 (1964)] Sov. Phys. JETP **20**, 762 (1965)
139. E. Dagotto, Science **309**, 257 (2005)
140. E. Dagotto, *Nanoscale phase separation and colossal magnetoresistance: physics of manganites and related compounds*, Springer Series in Solid-State Sciences, (Springer-Verlag Berlin Heidelberg New York 2003)
141. E.L. Nagaev, Physics-Uspechi **39**, 781 (1996)
142. E.L. Nagaev, Physica B **230–232**, 816 (1997)
143. J.B. Sokoloff, Phys. Rev. B **3**, 3826 (1971)
144. V. Kiryukhin, T.Y. Koo, H. Ishibashi, J.P. Hill, S-W. Cheong, Phys. Rev. B **67**, 064421 (2003)
145. L.N. Bulaevskii, C.D. Batista, M.V. Mostovoy, D.I. Khomskii, Phys. Rev. B **78**, 024402 (2008)
146. H.J. Schulz, Phys. Rev. Lett **65**, 2462 (1990)
147. H.J. Schulz, Phys. Rev. Lett **64**, 2831 (1990)
148. H.J. Schulz, Phys. Rev. Lett **64**, 1445 (1990)
149. P. Fazekas, E. Müller-Hartmann, Z. Phys. B **85**, 285 (1991)
150. J.E. Hirsch, Phys. Rev. B **31**, 4403 (1985)
151. B. Fine, T. Egami, Phys. Rev. B **77**, 014519 (2008)
152. P.W. Anderson, Phys. Rev. Lett. **63**, 1839 (1990)
153. S.Y. Wang, J.Z. Yu, H. Mizuseki, Q. Sun, C.Y. Wang, Y. Kawazoe, Phys. Rev. B **70**, 165413 (2004)
154. P. Schlottmann, Int J. Mod. Phys. B **11**, 355 (1997)
155. D. Braak, N. Andrei, Nucl. Phys. B **542**, 551 (1999)
156. A. ÓHare, F.V. Kusmartsev, K.I. Kugel, Phys. Rev. B **79**, 014439 (2009)
157. J.H. Hirsch, Phys. Rev. B **67**, 035103 (2003)
158. S. Belluci, M.Cini, P. Onorato, E. Perfetto, J. Phys.: Condens. Matter **18**, S2115 (2006)
159. W.-F. Tsai, S.A. Kivelson, Phys. Rev. B **73**, 214510 (2006)
160. S.R. White, S. Chakravarty, M.P. Gelfand, S.A. Kivelson, Phys. Rev. B **45**, 5062 (1992)
161. R.M. Fye, M.J. Martins, R.T. Scalettar, Phys. Rev. **42**, R6809 (1990)
162. N.E. Bickers, D.J. Scalapino, R.T. Scalettar, Int. J. Mod. Phys. B **1**, 687 (1987)
163. S.A. Kivelson, V.J. Emery, in *Strongly correlated electronic materials: the los alamos symposium 1993*, ed. by K.S. Bedell, Z. Wang, B.E. Meltzer, A.V. Balatsky, E. Abrahams (Addison-Wesley, Reading, MA, 1994)

164. E. Arrigoni, E. Fradkin, S.A. Kivelson, *Phys. Rev. B* **69**, 214519 (2004)
165. J.M. Tranquada et al., *Nature (London)* **375**, 561 (1995)
166. F.F. Assaad, W. Hanke, D.J. Scalapino, *Phys. Rev. Lett.* **71**, 1915 (1993)
167. D. Sénéchal, D. Perez, M. Pioro-Ladriere, *Phys. Rev. Lett.* **84**, 522 (2000)
168. J.-P. Malrieu, N. Guihery, *Phys. Rev. B* **63** 085110 (2001)
169. E. Altman, A. Auerbach, *Phys. Rev. B* **65**, 104508 (2002)
170. R.W. Richardson, *Phys. Lett.* **14**, 325 (1965)
171. A. Bohr, B.R. Mottelson, D. Pines, *Phys. Rev.* **110**, 936 (1958)
172. K.A. Matveev, A.I. Larkin, *Phys. Rev. Lett.* **78**, 3749 (1997)
173. C. Herring, in “*Exchange interactions among itinerant electrons*”, in *magnetism* vol. IV, ed. by G.T. Rado, H. Suhl (Academic, New York, 1966)
174. S. Sachdev, *Quantum phase transitions* (Cambridge University Press, Cambridge, 1999)
175. C.N. Yang, *Phys. Rev. Lett.* **63**, 2144 (1989)
176. C.N. Yang, *Phys. Lett. A* **161**, 292 (1991)
177. E. Lieb, *Phys. Rev. Lett.* **62**, 1201 (1989)
178. C.N. Yang, S.C. Zhang, *Mod. Phys. Lett. B* **4**, 759 (1990)
179. S.C. Zhang, *Phys. Rev. Lett.* **65**, 120 (1990)
180. A. Kaminski et al., *Nature (London)* **416**, 610 (2002)
181. A.N. Kocharian, G.W. Fernando, K. Palandage, J.W. Davenport, *Sol. State Phenomena* **152–153**, 583 (2009)
182. A.N. Kocharian, C. Yang, Y.L. Chiang, *Physica C* **341–348**, 253 (2000)
183. A.N. Kocharian, C. Yang, Y.L. Chiang, T.Y. Chou, *J. Phys. Soc. Jpn.* **71**, Suppl. 223 (2002)
184. K.R. Naqvi, *Phys. Rev.* **15**, 1807 (1977)
185. G.J. Hatton, *Phys. Rev.* **14**, 901 (1976)
186. F. Mancini, F.P. Mancini, *Phys. Rev. B* **77**, 061120 (2008)
187. P.W. Anderson, *Mater. Res. Bul.* **8**, 153 (1973)
188. P. Fazekas, P.W. Anderson, *Philos. Mag.* **30**, 432 (1974)
189. P.W. Anderson, *Science* **235**, 1196 (1987)
190. I.A. Sergienko, S.H. Curnoe, *Phys. Rev. B* **70**, 144522 (2004)
191. H. Kamimura, H. Ushio, S. Matsuno, T. Hamada, *Theory of copper oxide superconductors, Chap. VII: electronic structure of a CuO_5 pyramid in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+a}$* (Springer, Berlin, Heidelberg, 2005), p. 51
192. J.K. Jain, P.W. Anderson, *PNAS*, **106**, 9131 (2009)
193. J.K. Chin et al., *Nature (London)* **443**, 961 (2006)

Mechanical Properties of One-Dimensional Nanostructures

Gheorghe Stan and Robert F. Cook

Summary. The elastic mechanical properties of one-dimensional nanostructures are considered, with an emphasis on the use of contact-resonance atomic force microscopy methods to determine elastic moduli. Various methods used to determine elastic moduli of one-dimensional nanostructures are reviewed before detailed consideration of the experimental and analytical methods used in contact-resonance atomic force microscopy. As direct applications of contact-resonance atomic force microscopy on one-dimensional nanostructures, two measurement examples are shown here, for ZnO and Te nanowires. The variations of the elastic moduli of ZnO and Te nanowires with nanowire diameter are presented and interpreted in terms of core-shell models of nanowire structure. Based on combined theoretical, atomistic simulation, and experimental results, the importance of accurate and precise methods for measuring the mechanical properties of nanostructures, and how those methods need to be adjusted for one-dimensional nanostructures, is emphasized.

Key words: Atomic force microscopy, Contact resonance; Nanoscale elasticity, One-dimensional nanostructures.

16.1 Introduction

Nanostructures provide great potential for engineering and science applications: The ultra-small size leads to large surface area to volume ratios and consequent surface-induced changes in material properties from those observed in bulk analogs (in cases in which they exist - sometimes surface effects stabilize structures or phases that have no bulk analog). The ultra-small size also often leads to structures that contain no defects, such that component properties are very different from those with larger scales. By definition, such nanostructures must be of limited dimensionality, such that at least one of the dimensions is nanoscale, and include zero-dimensional nanostructures (nanoparticles, quantum dots), one-dimensional (1D) nanostructures (nanotubes (NTs), nanowires (NWs), nanobelts (NBs), nanorods (NRs)), and

two-dimensional nanostructures (sheets, membranes, plates). These reductions in dimensionality lead to further changes in material and component properties from the related three-dimensional bulk analogs.

Attention here is focused on the mechanical properties - specifically the elastic properties - of 1D nanostructures. Early considerations of molecular-scale carbon NTs recognized the extremely large axial stiffness and failure loads that such structures could exhibit, corresponding to extremely large effective moduli and strengths [1]. Composite materials incorporating carbon NTs would then exhibit extremely large strength- and stiffness-to-weight ratios, promising a revolution in structural materials, which appears to be approaching fruition [2]. A series of articles on carbon NTs (see journal issue containing [1]) focused primarily on processing of NTs so as to achieve the desired and defect-free structures that would enable such composites and other applications, although an indirect measurement of axial modulus did confirm the multi-TPa values predicted by simulation [3]. A more recent series of articles (see journal issue containing reference [4]) considering NTs and NWs more broadly highlighted the diversity of materials that could be formed into supra-molecular 1D nanostructures, the wide range of properties that they could exhibit, and the direct measurement of the elastic properties of isolated 1D nanostructures.

This chapter will consider measurement of elastic moduli of 1D nanostructures, with an emphasis on contact-resonance atomic force microscopy (CR-AFM) methods and the results obtained by these methods on several 1D nanostructure materials. Section 16.2 provides a review of the methods available for determining moduli of 1D nanostructures. This is followed by sections describing CR-AFM in detail and the extensions to CR-AFM required for its use in determining the elastic moduli of 1D nanostructures. Experimental measurements by CR-AFM of the elastic moduli of ZnO and Te NWs are then described in Sect. 16.5, followed by a section considering the physical and theoretical interpretations of the distinct size effects observed in the moduli measurements. The chapter concludes with a brief discussion of the importance of mechanical property measurements of 1D nanostructures in applications.

16.2 Mechanical Property Measurements of One-Dimensional Nanostructures

In this section, a short review of various techniques dedicated to quantitative measurement of the elastic moduli of 1D nanostructures will be provided. The pros and cons of each technique will be highlighted. Table 16.1 summarizes elastic modulus values measured on various 1D nanostructures by different techniques.

Table 16.1. Elastic modulus of 1D nanostructures measured by various techniques. The following abbreviations are used: MWCNTs – multi-walled carbon NTs; SWCNTs – single-walled carbon NTs; NWs – nanowires; NTs – nanotubes; NBs – nanobelts; NRs – nanorods. Superscript (a) indicates 1D nanostructures for which size-dependent elastic moduli were observed in measurements and superscript (b) refers to 1D nanostructures reported with no size-dependent elastic modulus

1D nanostructures (references)	E (GPa)	1D nanostructures (references)	E (GPa)
<i>Electromechanical resonance technique</i>			
MWCNTs [5]	100–1,000 ^a	MWCNTs [6]	20–35 ^b
SiC–SiO _x NWs [7]	50–70 ^a	ZnO NBs [8]	40–55 ^b
SWCNTs bundles [9]	76 ± 4 ^b	ZnO NWs [10]	135–220 ^a
MWCNTs [11]	126–937 ^b	SiO ₂ NWs [12]	47 ± 10 ^b
<i>Tensile loading technique</i>			
MWCNTs [13]	270–950 ^b	Pd NWs [14]	99.4 ± 6.6 ^b
MWCNTs [15]	315 ± 11 ^b	MWCNTs ropes [16]	450 ± 230 ^b
ZnO NWs [17]	97 ± 18 ^b	ZnO NWs [18]	140–160 ^a
<i>Three-point bending technique</i>			
MWCNTs [19]	650–1,220 ^b	MoS ₂ NT ropes [20]	120 ^b
Au NWs [21]	70 ± 11 ^b	SiO ₂ NWs [22]	76 ± 7 ^b
ZnS NBs [23]	52 ± 7 ^b	ZnO NBs [24]	100–160 ^a
Ge NWs [25]	112 ± 43 ^b	Ag NWs [26]	102 ± 23 ^b
GaN NWs [27]	220–320 ^a	CuO NWs [28]	70–300 ^a
LaB ₆ NWs [29]	467 ± 16 ^b	ZnO NWs [30]	133 ± 15 ^b
<i>Beam-bending technique</i>			
SiC NRs [31]	610–660 ^b	MWCNTs [31]	1,280 ± 590 ^b
ZnO NWs [32]	15–47 ^a	TaSi ₂ NWs [33]	170–485 ^a
<i>Nanoindentation technique</i>			
Ag NWs [34]	88 ± 5 ^b	ZnO NWs [35]	80–115 ^b
GaN NWs [35]	40–60 ^b	ZnS NBs [36]	36 ± 4 ^b
ZnO NBs [37]	31 ± 1 ^b	Au NWs [38]	110 ± 10 ^b
LaB ₆ NWs [29]	439 ± 26 ^b	ZnO NWs [39]	117–232 ^a
<i>Contact modulation AFM technique</i>			
Ag NWs [40]	65–140 ^a	ZnO NWs [41]	100–200 ^a
SnO ₂ NBs [42]	154 ± 18 ^b	MWCNTs [43]	21 ± 7 ^b
Te NWs [44]	45–85 ^a	AlN NTs [45]	250–400 ^b
Ag NWs [46]	70–160 ^a	Pb NWs [40]	15–30 ^a

16.2.1 Electric Field-Induced Mechanical Resonance of One-Dimensional Nanostructures

Electric field-induced mechanical resonance was introduced in 1999 by Poncharal et al. [5] for elastic bending-modulus measurements of carbon NTs and

used subsequently on various 1D nanostructures (single-walled and multi-walled carbon NTs [6, 9, 11], silicon carbide-silicon NWs [7], ZnO NBs [8], SiO₂ NWs [12], ZnO NWs [10], etc.). In this technique, an electromechanical stage positions the 1D nanostructures to be tested and controls the applied voltage inside a transmission electron microscope (TEM) or scanning electron microscope (SEM). 1D nanostructures are synthesized or attached at the end of an electrode and an ac voltage with tunable frequency is applied across the nanostructures with a counter electrode. The induced electrostatic force vibrates the nanostructures and mechanical resonances are produced when the frequency of the electric field matches the natural harmonics of the nanostructures. After the resonance frequencies are identified for a selected nanostructure, the dimensions of that nanostructure are determined from scan images. Given that the mechanical vibrations induced in a specimen are determined by its dimensions and elastic properties, classical elasticity theory [47] then uses the resonances of the investigated specimen (tube, cylindrical rod, etc.) to calculate the bending modulus.

The advantage of this technique resides in the fact that the investigated nanostructure can be directly imaged by TEM and its crystallographic structure observed. Thus, surface defects or changes in the internal structural crystallography can be correlated with the measured elastic modulus. The disadvantage is that the tests are limited to 1D nanostructures permanently attached to a substrate or a probe and extensive manipulation would be required to recover the tested specimen for subsequent applications. For 1D nanostructures integrated into a device, detection of their resonant electromechanical oscillations could be performed, for example, by a capacitive technique [48] or modulated scanning force microscopy [49]. However, as in any clamped configuration, boundary conditions assignment at the clamp ends can sometimes be ambiguous, in which case errors are introduced in the elastic modulus calculation [50].

16.2.2 Axial Tensile Loading of One-Dimensional Nanostructures

A direct method for simultaneously observing and measuring specimen deformation is that of a tensile test. By loading the specimen along its axis, the response of the specimen to the applied tension is quantified by elastic modulus, yield strength, ultimate tensile strength, etc. Tensile tests are the most standardized tests used for mechanical characterization of macroscale specimens. However, implementation of tensile tests at the nanoscale faces immense challenges due to the precarious manipulation and positioning of nanostructures into a testing device.

The first tensile testing of a 1D nanostructure was performed by Yu et al. [13] in 2000 by observing the axial stress-strain response of multi-walled carbon NTs inside a SEM. Individual multi-walled NTs were placed between two opposing AFM probes mounted on a multi-axis tensile stage. The tensile deformation of the NTs had been made observable by using AFM probes of

different compliance: one stiff and one compliant. The stiff AFM probe was driven outward and, in this way, the NT was pulled along its axis. The cantilever deflection of the compliant probe was used to determine the applied force on the NT. The length change of the NT was video recorded by SEM imaging. From the stress versus strain curves obtained, the elastic modulus of the outermost layers was estimated and a fracture mechanism of multi-walled carbon NTs under tension established.

A high-resolution control for tensile tests of 1D nanostructures was accomplished by Zhu and Espinosa [14] through of a micro-electromechanical system for in situ electron microscopy. In addition to a loading stage, the testing device included an actuator and a capacitive load sensor fabricated by surface micromachining. Using this device, tensile tests were performed, simultaneously measuring the applied load with nanoNewton resolution and specimen deformation with sub-nanometer resolution. Tensile loading methods involving less manipulation of the 1D nanostructures to be tested were proposed on vertically as-grown NWs. In this configuration, one end of the specimen is attached to the substrate and the other end to an AFM probe. A piezoelectric nanomanipulator was used to actuate either the substrate [51] or the AFM probe [17].

Owing to their simple measurement concept, tensile tests provide the most direct measurements of the mechanical response of 1D nanostructures. However, the results of such tensile tests can be still affected by uncontrollable twisting of the specimen during mounting or applied shear loads arising from imprecise alignment of the specimen along the tensile axis.

16.2.3 Three-Point Bending Test of Bridge-Suspended One-Dimensional Nanostructures

Used as a standard technique to measure the flexural properties of macroscopic specimens, the three-point bending test was first adopted at the nanoscale in 1999 by Salvétat et al. [19, 52] to measure the bending modulus of individual and bundles of carbon NTs that were bridge-suspended over holes in a porous alumina membrane. After their dispersion on the perforated membrane, the NTs were held over the holes (about 200 nm in diameter) by the adhesion formed between the unsuspended parts of the NTs and the substrate. A concentrated load was applied in the middle of a suspended NT by an AFM probe and the deflection of the beam measured. Elastic beam-theory [47] was used to analyze the deflection versus load curves and calculate the bending modulus of the specimens tested.

In the case of individual suspended NTs, the bending modulus coincides with the axial Young's modulus of the NTs [19], whereas for NT bundles bending as well as the shear between NTs has to be considered [52]. In this last case, the axial Young's modulus is obtained solely from measurements on small diameter and long ropes for which the intertube shearing is negligible. In addition, the intertube shear modulus is extrapolated from

measurements performed on larger cross-section bundles. The axial Young's modulus of individual carbon NTs (single-walled [19] as well as multi-walled [52]) was determined to be around 1 TPa and the intertube shear modulus about 1 GPa. The uncertainty of this type of measurement is due mainly to inaccurate knowledge of the specimens' dimensions: suspended length and cross-sectional area.

The accuracy of three-point bending tests on 1D nanostructures can be improved by performing successive load-deflection measurements along the suspended length of the structure [23, 24]. In this way, the uncertainty introduced by the inexact positioning of the AFM probe at the midpoint of the filament is eliminated. Also, a self-consistent interpretation of the load-deflection curves measured at different contact points along the bridge validates the assumption made for the boundary conditions of the system: either clamped-clamped or free-free ends. In a slightly different configuration, three-point bending tests on bridge-clamped 1D nanostructures were performed by moving the AFM probe laterally [21] with respect to the filament rather than normally. In this configuration, the clamping points were defined better and wire-substrate friction eliminated.

16.2.4 Beam-Bending of One-End-Clamped One-Dimensional Nanostructures

One easy way of observing the bending behavior of 1D nanostructures is to fix one of end of specimen rigidly to a base (clamped or pinned) and apply a known force at the free end of the specimen. Such a set up was used by Wong et al. [31] in 1997 to quantify for the first time the bending modulus of some 1D nanostructures (carbon NTs and SiC NRs). The experiment demonstrated the unique potential of AFM to investigate the mechanical properties of 1D nanostructures. Since then an entire class of AFM-based techniques has emerged to improve and diversify nanomechanical property measurements and specimen accessibility.

In the original experiment, silicon carbide NRs and multi-walled carbon NTs were dispersed on the flat surface of a MoS₂ single crystal (to reduce friction) and clamped at one end to this substrate by the deposited SiO₂ pads. The experiment was performed in water to further reduce the friction between the 1D nanostructures tested and the substrate during bending. In individual AFM scan lines, lateral force versus displacement curves were acquired as the positioning of the applied force was varied along a filament toward the free end. Because the lateral force was uncalibrated, the derivative of the force versus displacement rather than force versus displacement itself was then fitted with the corresponding formula from elastic beam theory to determine the bending modulus of the 1D nanostructure tested. This procedure also eliminates the unknown contribution of the friction force between a filament and the substrate and does not require exact knowledge of the pinning location.

A different experimental configuration for lateral bending tests on 1D nanostructures was introduced by Song et al. [32]. They observed the lateral bending of an AFM probe scanned in contact with vertically aligned NWs. The benefit of such an arrangement is that it does not require any manipulation of the nanostructures, so they can be tested directly on the substrate on which they were synthesized. Although easy to implement, the main disadvantage of this approach is that the diameter and the length of the NWs measured are inaccurately determined. To address this issue, the stress distribution along a vertical NW, which is bent laterally by an AFM probe, was investigated [53] by correlating in situ SEM observations with finite element simulations.

16.2.5 Instrumented Indentation of One-Dimensional Nanostructures

Instrumented indentation testing (IIT) is the most reliable technique used to measure mechanical properties such as elastic modulus and hardness of materials at the micro-scale. The technique consists of applying a prescribed load to an indenter in contact with the specimen tested. During IIT, the applied load and penetration depth are simultaneously recorded and a load-unload force-displacement curve is obtained. The hardness is simply taken as the load divided by the contact area at the maximum loading point, whereas the elastic modulus is calculated from the initial slope of the unloading portion of the force-displacement curve. A specific tip-shape calibration is made prior to measurements, in terms of the projected contact area of the indenter as a function of contact depth, usually using a fused silica standard.

With loads in the milli- to micro-Newton range and spatial resolution in the nanometer to micrometer range, the current IIT devices, or “nanoindenters,” are suitable for measuring mechanical properties at the micro-scale. To probe entities of even smaller dimensions, such as 1D nanostructures, impediments such as lack of precise contact-point determination, adequate experimental procedures for preserving the integrity of the system tested, and verified measurement interpretation have to be overcome. Despite these limitations, the use of nanoindenters has been extended for testing mechanical properties of various 1D nanostructures simply supported on substrates [34, 54].

Nominally, the analysis of IIT measurements follows the basic methodology developed by Oliver and Pharr [55] for indentation of a flat material surface. However, in the case of indentation of 1D nanostructure such as NWs or NTs, contact geometry corrections need to be considered. As an example, the ellipticity of the contact has to be considered for a round tip indenting cylindrical NWs or NTs [56]. Furthermore, due to the large sensing depths that are characteristic in indentation tests, mechanical contributions from both interface contacts has to be considered; from the indenter-NW interface as well as from the NW-substrate interface [35].

Besides the earlier described measurement geometry, that of a probe radially indenting a 1D nanostructure on a substrate, nanoindentation tests have been also performed on vertically aligned 1D nanostructures (e.g., ZnO NWs [39,57]) to investigate their buckling behavior and estimate the elastic modulus.

16.2.6 Contact Modulation AFM-Based Techniques

A distinct class of nanomechanical property measurement methods is that of AFM-based techniques, in which the dynamics of the AFM cantilever is used to access the mechanical response of the material tested. Endowed with nanoscale spatial resolution and nanoNewton force control, techniques such as CR-AFM (which includes atomic force acoustic microscopy (AFAM) [58] and ultrasonic AFM [59]), ultrasonic force microscopy [60], heterodyne force microscopy [61], passive overtone microscopy [62], and torsional harmonic dynamic force microscopy [63] can be used to quantify the mechanical properties of a probe-sample contact. By adding mechanical property measurement capability to regular AFM nano-scale topography scanning, such techniques can provide high-resolution, versatile, and complementary characterization of stand-alone nanostructures.

Ultrasonic force microscopy [60] relies on detecting cantilever deflections in response to an ultrasonic vibration that leads to periodic indentation of the sample at frequencies much greater than the cantilever resonance. For a given threshold value of the modulation amplitude, the tip detaches from the sample for short periods of time during each modulation cycle. These pull-off events are detected as distinct nonlinearities in the cantilever deflection and are used to probe the local mechanical properties of the sample: the derivative of the indentation-force versus the threshold modulation amplitude provides the effective stiffness of the contact.

In CR-AFM [58, 59], the applied load is large enough to overcome the adhesion force, and the tip-sample contact is placed in the linear regime. In contrast to ultrasonic force microscopy, the modulation in CR-AFM is of small amplitude and variable frequency. Considerations of cantilever dynamics and tip-sample contact mechanics are necessary to convert the measured contact resonance frequencies into contact stiffness and elastic modulus of the sample tested.

Applications of ultrasonic force microscopy and CR-AFM to measure the elastic modulus of 1D nanostructures were performed on SnO₂ NBs [42], ZnO NWs [41], multiwalled carbon NTs [43], Te NWs [44], and AlN NTs [45]. Measurements by CR-AFM have also been made on 1D nanostructures suspended over holes [40]. In this case, the CR-AFM (performed with electrostatic modulation) was used to provide accurate inputs for the three-point bending test. The measured contact stiffness was determined by the beam deformation rather than the contact deformation.

16.3 Contact-Resonance Atomic Force Microscopy

The concept of the CR-AFM technique is to detect and track the change in the resonance frequencies of an AFM cantilever-probe brought into contact with the material tested. By solving equations describing the cantilever dynamics and the tip-sample contact mechanics, the measured contact resonance frequencies are converted into the local elastic modulus of the material probed. In the measurement geometry, the AFM probe is well described as a clamped-spring coupled beam, clamped at the cantilever base and spring-coupled with the sample at the cantilever end. It is this spring-coupling between the AFM probe and the sample that tunes the resonances of the cantilever to the elastic properties of the sample. Therefore, there are two main steps in CR-AFM: (1) Based on a clamped-spring coupled beam model of the cantilever, determine the stiffness of the probe-sample contact from the measured contact resonance frequencies and (2) using a contact mechanics model that adequately describes the elastic deformation of the probe-sample contact, convert the contact stiffness into the elastic modulus of the sample.

The CR-AFM technique was introduced by Rabe et al. [58] as AFAM and Yamanaha et al. [59] as ultrasonic AFM (UAFM). In both versions, the AFM cantilever is mechanically vibrated by a small-amplitude oscillation over a kiloHertz to megaHertz frequency range to detect the shift in the cantilever resonances from air to contact. In UAFM, the vibration is imposed from the base of the cantilever whereas in AFAM the tip-sample contact is vibrated from beneath the sample. Regardless of the location of the vibrating source, the two techniques work in the same way and the same analysis is used to convert the measured contact resonance frequencies into elastic modulus. In the following, the term CR-AFM will be used to describe these techniques. The basic methodology of CR-AFM will be briefly reviewed in this section; a more detailed review of cantilever dynamics and practical aspects of this technique are available elsewhere [64, 65].

In the past few years, CR-AFM has been used successfully to quantify the elastic properties of a large variety of materials and structures: piezoelectric ceramics [66], metal films [67], diamond-like carbon films [68], glass-fiber-polymer matrix composites [69], clay minerals [70], polycrystalline materials [71], and 1D nanostructures (NWs [41, 44], NBs [42], and NTs [43, 45]). Nanoscale resolution has been demonstrated in CR-AFM measurements of elastic properties either in local point measurements [41, 44, 45, 66, 67, 72, 73] or elastic modulus mapping [68, 69, 71, 74].

Besides excellent spatial and measurement resolution, the suitability of using CR-AFM for modulus measurements is also promoted by the simple additional instrumentation required to augment a commercial AFM for CR-AFM. An example of such instrumentation is shown in Fig. 16.1, and illustrates the measurement sequence: First, the AFM probe is brought into contact with the sample through a piezoelectric scanner by means of the AFM controller, which maintains the prescribed applied load through its Z-

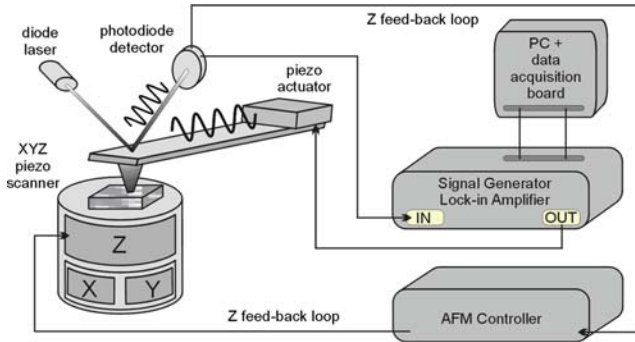


Fig. 16.1. A schematic diagram of an experimental setup for local point CR-AFM measurements. The internal signal generator of the lock-in amplifier is used to modulate the AFM cantilever brought into contact with the sample (Reprinted with permission from Stan and Cook [74]. Copyright 2008 Institute of Physics Publishing)

feedback loop. A small-amplitude mechanical vibration is then superimposed on the loaded AFM cantilever through a piezo-actuator oscillation. This vibration is transmitted through the AFM tip and excites acoustic waves in the sample. The cantilever's end conditions are modified by the tip-sample coupling and this change is directly observed in the resonance frequency of the vibrated cantilever. To detect the cantilever resonances, the frequency of the oscillating signal is swept over kiloHertz–megaHertz range and the unfiltered high-frequency response is collected from the AFM photodiode through lock-in amplifier detection.

16.3.1 Cantilever Dynamics in CR-AFM

In CR-AFM, the load applied at the tip-sample contact is chosen to be large enough to exceed adhesion contact forces, but small enough to prevent plastic deformation at the contact. Under these restrictions, the contact is deformed elastically and the simple model of such elastic tip-sample coupling is that of a linear spring. The idealized system, schematically shown in Fig. 16.2, consists of a beam clamped at one end and spring-coupled at the other. Two simplified versions are considered here: normal contact coupling only, as in Fig. 16.2a, and both normal and lateral contact couplings, as in Fig. 16.2b. The lateral coupling is imposed principally by the tilt of the cantilever with respect to the sample. Viscoelastic characteristics of the tip-sample contact are neglected.

As the vibrating cantilever is brought from air into contact, its resonance frequencies undergo distinct shifts. The contact resonance frequency of a given vibrational mode of the cantilever is in the range from the free (in air) resonance frequency of that mode to the free resonance frequency of the next mode, and depends on the mechanical properties of the contacted material

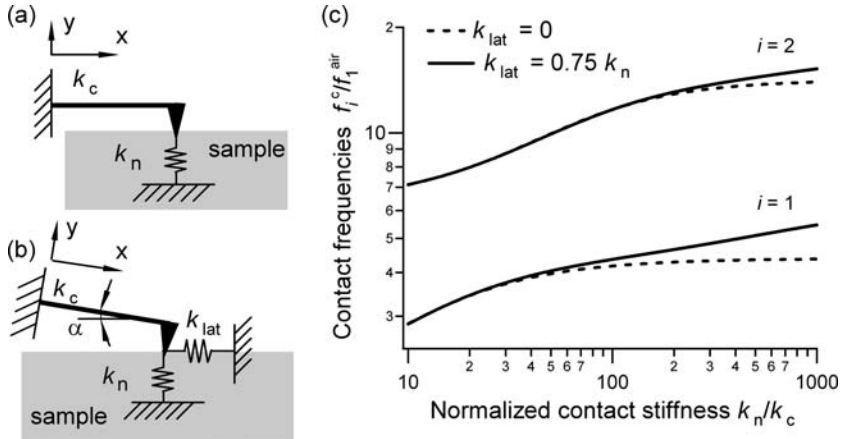


Fig. 16.2. Idealized AFM probe (cantilever and tip) spring-coupled to the sample: (a) only normal contact coupling is considered, (b) both normal and lateral contact couplings are considered, (c) the effect of a supplemental lateral contact coupling on the first two ($i = 1, 2$) contact resonance frequencies of the cantilever as a function of the normal contact stiffness. The frequencies are normalized to the first resonance frequency of the cantilever in air, f_1^{air} (Reprinted with permission from Stan and Cook [74]. Copyright 2008 Institute of Physics Publishing)

and the applied load. Once the contact resonance frequencies are measured, the contact stiffness is calculated by considering the cantilever dynamics in a clamped-spring coupled beam configuration: clamped at the base of the cantilever and spring coupled at the end of the cantilever. Thus, for the configuration shown in Fig. 16.2b, the wave equation for the flexural vibrations along the cantilever beam [58],

$$EI \frac{\partial^4 y(x, t)}{\partial x^4} + \rho A \frac{\partial^2 y(x, t)}{\partial t^2} = 0, \tag{16.1}$$

must be solved with the following boundary conditions at the fixed (clamped) end, $x = 0$, and at the spring-coupled end, $x = L$, of the cantilever:

$$\begin{cases} y(x, t)|_{x=0} = 0, \quad \frac{\partial y(x, t)}{\partial x} \Big|_{x=0} = 0 \\ EI \frac{\partial^2 y(x, t)}{\partial x^2} \Big|_{x=L} = -F_x h, \quad EI \frac{\partial^3 y(x, t)}{\partial x^3} \Big|_{x=L} = F_y \end{cases}, \tag{16.2}$$

where

$$\begin{cases} F_x = h (k_n \sin^2 \alpha + k_{lat} \cos^2 \alpha) \frac{\partial y(x, t)}{\partial x} \Big|_{x=L} + \sin \alpha \cos \alpha (k_{lat} - k_n) y(x, t) \Big|_{x=L}, \\ F_y = h \sin \alpha \cos \alpha (k_{lat} - k_n) \frac{\partial y(x, t)}{\partial x} \Big|_{x=L} + (k_n \cos^2 \alpha + k_{lat} \sin^2 \alpha) y(x, t) \Big|_{x=L}. \end{cases} \tag{16.3}$$

Here, E is the Young’s modulus along the long axis of the cantilever, I the moment of inertia of the cantilever cross-section, ρ the density of the cantilever

material, A the cross-sectional area of the cantilever, h the height of the AFM tip, and α is the tilt angle of the cantilever with respect to the sample surface. k_n and k_{lat} are normal and lateral contact stiffnesses, respectively. In terms of these parameters, the cantilever stiffness is given by $k_c = 3EI / L^3$, where L is the length of the cantilever, and the angular frequency of the flexural cantilever vibrations is $\omega = k^2 \sqrt{EI / \rho A}$, where k is the flexural wavenumber. The characteristic equation obtained from the above equations provides the dispersion relationships for the resonance frequencies of the clamped-coupled cantilever as a function of the contact spring-coupling stiffnesses:

$$\lambda^4 (1 + \cos \lambda \cosh \lambda) + \lambda^3 (\sin \lambda \cosh \lambda + \cos \lambda \sinh \lambda) C_1 + 2\lambda^2 (\sin \lambda \sinh \lambda) C_2 + \lambda (\sin \lambda \cosh \lambda - \cos \lambda \sinh \lambda) C_3 + (1 - \cos \lambda \cosh \lambda) (C_1 C_3 - C_2^2) = 0, \tag{16.4}$$

with, $C_1 = 3(h/L)^2 (k_n \sin^2 \alpha + k_{lat} \cos^2 \alpha) / k_c$, $C_2 = 3(h/L) \sin \alpha \cos \alpha (k_{lat} - k_n) / k_c$, $C_3 = 3(k_n \cos^2 \alpha + k_{lat} \sin^2 \alpha) / k_c$, and $\lambda = kL$. A detailed treatment, which also addresses the tip position along the cantilever can be found elsewhere [64].

In Fig. 16.2c are shown the first two resonance frequencies of the clamped-coupled cantilever as a function of the normal contact stiffness for two cases: with the lateral contact coupling neglected, as in Fig. 16.2a, and with $k_{lat} = 0.75 k_n$, as in Fig. 16.2b. The cantilever angle was taken as $\alpha = 11^\circ$ and the tip-height ratio taken as $h/L = 0.06$. A measurement example is shown in Fig. 16.3a with the first two resonance frequencies of a cantilever shifting from air to contact on Si(100). The measured contact resonances were then used

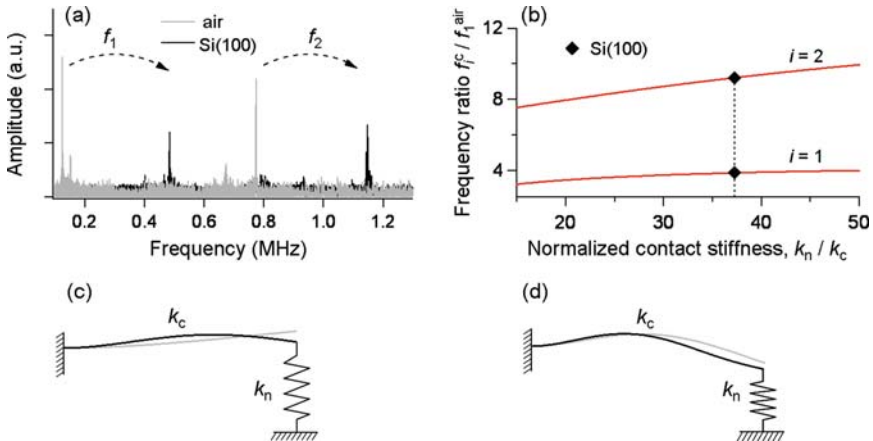


Fig. 16.3. (a) The shifts in the first and second resonance frequencies of a cantilever from air to contact on Si(100). (b) With the resonances shown in (a), the normal contact stiffness is calculated by using the clamped-spring coupled model for the cantilever. First and second (d) flexural resonance vibrations of the free and spring-coupled cantilever with the contact stiffness from (b)

in (16.4) to calculate the normalized contact stiffness as shown in Fig. 16.3b ($k_{\text{lat}} = 0$ and $\alpha = 0$ were assumed here). Figures 16.3c and 16.3d show how the first two flexural resonance vibrations of the cantilever are modified when the free end of the cantilever is attached to a spring of the same stiffness as the calculated contact stiffness. With the measurement of at least two resonance frequencies (usually the first two modes) the position of the tip along the cantilever is also unequivocally determined from a set of two modified equations (16.4) [58,64].

16.3.2 Contact Mechanics in CR-AFM

The conversion of the contact stiffness, calculated from the measured contact resonance frequencies, into the elastic modulus of the material probed is made by means of a contact mechanics model that adequately describes the tip-sample contact. In the case of neglect of adhesive contact forces, the Hertz model¹ [75] is the usual choice to characterize the elastic deformation of the contact formed between the spherical end of the AFM tip and a flat surface. For such contact geometry, the normal contact stiffness, k_n , along the direction normal to the surface, is given by the derivative of the normal force, F_n , with respect to the relative approach distance, δ_n , between the tip and sample in contact: $k_n = \partial F_n / \partial \delta_n$. For a sphere-on-flat Hertzian contact, $\delta_n = (9F_n^2 / 16R_T E^{*2})^{1/3}$, so the normal contact stiffness is simply

$$k_n = 2aE^*. \quad (16.5)$$

The contact radius, $a = (3F_n R_T / 4E^*)^{1/3}$, is determined by the normal load F_n , the tip radius R_T , and the reduced elastic modulus E^* of the contact. Equation (16.5) is valid for contact by any solid of revolution, including that of a flat punch indenting a flat surface [76], although in this case the contact radius is independent of the applied load. These two extreme cases, spherical tip on flat surface and flat punch on a flat surface, bracket the typical CR-AFM contact geometries. The reduced elastic modulus is a combination of the indentation moduli of the tip, M_T , and sample, M_S :

$$1/E^* = 1/M_T + 1/M_S. \quad (16.6)$$

For elastically isotropic materials, the indentation modulus is simply expressed in terms of the Young's modulus, E , and the Poisson's ratio, ν , $M = E / (1 - \nu^2)$. In the case of elastically anisotropic materials that still possess threefold rotation axis symmetry, the indentation modulus can be calculated numerically for a given indentation direction as a function of the elastic constants of the indented material [77]. Anisotropy corrections to the indentation modulus can be also evaluated from measurements [71].

¹ In Hertzian contact, the two bodies are homogenous, isotropic, and linearly elastic (obeying Hooke's law), the points of contact are assumed to be in the plane tangent to the parabolic surfaces of the two bodies, and the contact is frictionless.

The goal of CR-AFM measurements is to obtain the value of the indentation modulus of the sample tested, namely M_S . This can be accomplished by using (16.4)–(16.6) with the following provisions: even if the lateral contact stiffness is negligible, the cantilever stiffness, contact radius, and indentation modulus of the tip are still needed for calculation. With two of these parameters, contact radius and cantilever stiffness, not accurately measured in common experiments, the results of CR-AFM could be seriously erroneous. A work-around method is to perform successive CR-AFM measurements on the material to be tested and on a reference material, for which the indentation modulus is either known or can be calculated based on its elastic constants. By taking the ratio of the contact stiffnesses normalized to the cantilever stiffness (these are the quantities calculated from the measured contact resonances) for the test and reference materials, the indentation modulus of the sample can be expressed in terms of known and measured parameters:

$$1/M_S = \left(k_n^{(R)} / k_n^{(S)} \right)^m / M_R + \left[\left(k_n^{(R)} / k_n^{(S)} \right)^m - 1 \right] / M_T. \quad (16.7)$$

Here, the symbols S and R stand for test sample and reference, and m distinguishes between a flat tip, $m = 1$, and a spherical tip, $m = 3/2$, in contact with a planar surface. It has been shown experimentally that more accurate CR-AFM results are obtained by selecting a reference material with mechanical properties similar to that of the test material [73]. In Sect. 16.3.3, one reasonable explanation for this observation is provided.

Another spring-coupling that contributes to the dynamics of the system is that of the lateral contact stiffness, k_{lat} (see Fig. 16.2b), which causes mechanical vibrations to be induced in the surface plane of the sample. Two things contribute to this coupling: one is the tilt of the cantilever with respect to the sample surface and the other is the rocking of the tip in the cantilever–tip plane by the flexural waves that vibrate the cantilever. The lateral contact stiffness is defined as the change in the tangential contact force F_{lat} with the lateral displacement δ_{lat} : $k_{\text{lat}} = \partial F_{\text{lat}} / \partial \delta_{\text{lat}}$. For a Hertzian contact between a spherical tip and a flat surface, k_{lat} is simply expressed as [78]

$$k_{\text{lat}} = 8aG^*, \quad (16.9)$$

where G^* is the reduced shear modulus. In characterizing the elastic lateral deformation of the tip–flat sample contact, the reduced shear modulus is given by a combination of the shear indentation moduli of the tip, N_T , and sample, N_S [79]:

$$1/G^* = 1/N_T + 1/N_S. \quad (16.10)$$

For elastically isotropic materials, the shear indentation modulus is expressed in terms of the shear modulus, G , and Poisson’s ratio, $N = G/(2 - \nu)$. In the absence of independent CR-AFM measurements that provide k_{lat} , the ratio k_n / k_{lat} is approximated as $2(1 - \nu)/(2 - \nu)$, for $E_T \gg E_S$ [80]. No slip over the contact region is assumed in the derivation of both normal and lateral contact stiffnesses.

16.3.3 Precision and Accuracy in CR-AFM Measurements (Dual-Reference Calibration Method for CR-AFM)

To improve the precision and accuracy of CR-AFM measurements, two simple tasks can be pursued with a minimal effort: (1) Collect a large number of measurements in a stable tip-sample contact geometry and (2) use at least two references with respect to which the measurements on the test material are calibrated.

A method of enhancing the precision of CR-AFM is to perform measurements on a large grid of points over flat surfaces (e.g., 10×10 over $3 \mu\text{m}^2$). Also, as with many nanoscale contact interactions, CR-AFM measurements are sensitive to contact area changes. Thus, in addition to the use of smooth sample surfaces, care has to be taken to track any tip changes that occur and ensure that measurements are acquired with stable tip geometry. As an example, in [81], with a stable flattened tip, CR-AFM measurements were obtained in a sample order of MgF_2 , CaF_2 , Au, and Si, twice, starting and ending with the stiffest materials to ensure that no systematic errors based on the sample order were introduced into the analysis. The reproducibility and stability of the tip-surface contact resonance spectra were checked by acquiring at each grid node the first two contact resonance frequencies on each tested material. In Fig. 16.4 are shown the distributions of the first and second contact resonance peaks acquired on Au(111), $\text{CaF}_2(100)$, and $\text{MgF}_2(001)$ surfaces. A Gaussian fit of each histogram was used to determine the most probable contact resonance frequency and estimate the measurement uncertainty. As can be seen, single point measurements were occasionally outside of the average distributions, showing that inhomogeneous or contaminated probed regions, as well as slight variations in the contact geometry, can provide puzzling

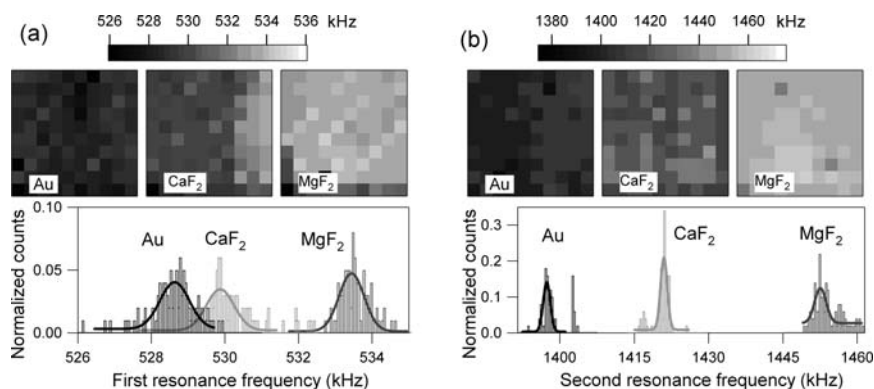


Fig. 16.4. (a) First and (b) second contact resonance frequencies measured over sampling arrays on Au, CaF_2 , and MgF_2 . In the bottom panels, the histograms obtained from the measured values in the top panels are shown (Reprinted with permission from Stan and Price [81]. Copyright 2006 American Institute of Physics)

results. This demonstrates that the solution for precise quantitative CR-AFM measurements is the use of a sampling array over the investigated area.

Improvement in the accuracy of CR-AFM is accomplished by the use of a dual-reference calibration method. The idea of this approach is to eliminate the uncertainty introduced by assuming the indentation modulus of the tip as a known parameter. Usually, as in (16.6), the tip is implicitly considered as a second reference, with the same indentation modulus assumed as for the bulk material of which it is composed. As the mechanical properties of silicon tips during scans can be altered by SiO₂ formation on the surface of the tip and the generation of amorphous regions induced as a result of nanodeformation, this assumption can sometimes not be true. The dual-reference method eliminates the need for any assumption of the indentation modulus of the tip and, instead, requires the use of two reference materials with known indentation moduli, M_{R1} and M_{R2} . The normal contact stiffnesses $k_n^{(S)}$, $k_n^{(R1)}$, and $k_n^{(R2)}$, measured with the same AFM probe on the sample tested and the two chosen references are used to calculate the indentation modulus of the sample [81]:

$$M_S = \frac{\left(k_n^{(R1)}/k_n^{(R2)}\right)^m - 1}{\left(k_n^{(R1)}/k_n^{(S)}\right)^m [1/M_{R2} - 1/M_{R1}] + \left(k_n^{(R1)}/k_n^{(R2)}\right)^m /M_{R1} - 1/M_{R2}}. \quad (16.11)$$

The validity of the dual reference method was checked on a group of four single crystals for which the indentation moduli were calculated from their elastic constants. By using as references the theoretical values for each pair of the four materials studied [Au(111), CaF₂(100), MgF₂(001), and Si(100)], the moduli of the other two materials were then calculated using (16.11). The results are shown in Fig. 16.5a. The deviations from the theoretical values for the calculated indentation moduli are less than 5% and the material indentation moduli determined with this dual-reference method are also insensitive to the assumed tip shape (flat or spherical). Part of this good agreement is also due to the fact that the averaged values from sampling arrays were used in calculations.

The uncertainty ΔM_S in the dual-reference method is only marginally affected by the mechanical properties of the tip, being primarily determined by the mechanical properties of the two chosen reference materials. In Fig. 16.5b are shown the overall uncertainties ΔM_S calculated from (16.11) by summing in quadrature [81]. As can be seen, the best measurement accuracy is obtained when the moduli of the two reference materials bracket that of the test material. The largest uncertainty, a worst case scenario, would be to have two references, each with similar mechanical properties but different from those of the tested material. Therefore, the calculation of ΔM_S constitutes a selection rule for the two reference materials that are appropriate for measuring a given test material.

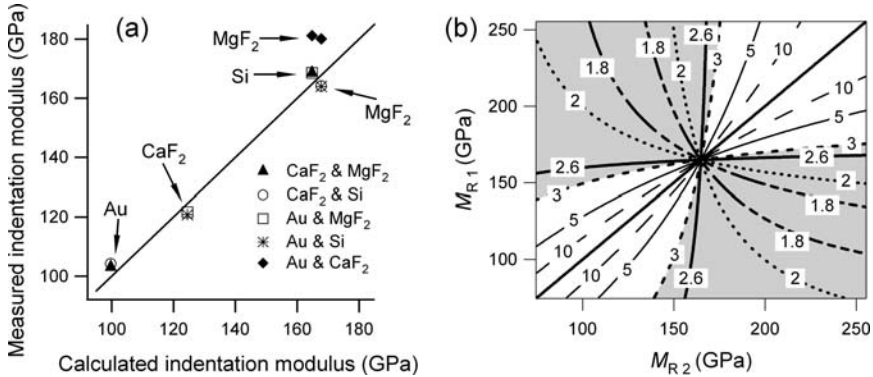


Fig. 16.5. (a) Results from CR-AFM dual-reference method: measured versus calculated values. For each measured material, the pairs of reference materials used are indicated in the legend. (b) Theoretical uncertainty ΔM_S in measuring the indentation modulus M_S of Si(100) with the dual reference method. Choosing two references that bracket Si results in an uncertainty of 3 GPa or less (gray areas) (Reprinted with permission from Stan and Price [81]. Copyright 2006 American Institute of Physics)

16.4 Contact-Resonance Atomic Force Microscopy Applied to Elastic Modulus Measurements of 1D Nanostructures

To assure the validity of elastic moduli of nanostructures measured by CR-AFM, the interpretation of measurements has to be amended through proper consideration of the contact geometry. This is because, for measurements on nanostructures, the contact geometry significantly deviates from the usual spherical-on-flat geometry. And, as in any other two-body contact problem, the applied load, the elastic properties, as well as the shapes of the two bodies forming the contact have to be considered for correct determination of the strain and stress distributions at the contact. Only by assuring the correct contact geometry can CR-AFM be advanced to provide elastic property measurement of nanostructures. The elimination of all the other contributions to the measurements except the elastic parameters allows discerning and analyzing possible variations in the elastic properties of nanostructures due to size effects.

The problem of contact geometry can be addressed either analytically or numerically. Using dedicated finite element analysis packages, a numerical solution can be conveniently generalized for any form and structure of the two objects in contact (tip and nanostructure). However, sometimes, the computation can be time and resource consuming. As a consequence, although it is limited to the case of two paraboloids of revolution pressed against each other in a Hertzian contact, the analytical solution is still useful for quick application

due to its simplicity. Thus, the Hertz solution for the elastic sphere-on-cylinder contact deformation was conveniently applied for contacts between an AFM tip and NTs [56, 82] and an AFM tip and NWs [41, 44]. In the following, the derivation of the analytical expressions of normal and lateral contact stiffnesses characterizing the Hertz sphere-on-cylinder contact will be reviewed for CR-AFM measurement interpretation on NWs of round cross-section.

16.4.1 Normal Contact Stiffness of the Tip–Nanowire Contact

The elastic deformation experienced by the contact formed between two solid paraboloids, in this particular case the spherical end of an AFM tip and a NW of circular cross-section (refer to Fig. 16.6), can be analyzed in the frame of Hertzian contact mechanics. The normal contact stiffness, k_n , of the contact characterizes the spring behavior of the contact under a load, F_n , that is applied normally at the contact point; normal here refers to the direction perpendicular to the contact plane. It can be shown that for points within the contact area, the relative distance, d , between points on the two unstressed surfaces can be related to the squares of their projections on the contact plane [75, 83]:

$$d = Ax^2 + By^2. \quad (16.12)$$

The distance d is measured along the direction of the applied load and the coordinates x and y are along and perpendicular to the long axis of the NW, respectively. The parabolic coefficients A and B are determined by the radii of curvature of tip, R_T , and NW, R_{NW} :

$$A = 1/2R_T, \quad B = 1/2R_T + 1/2R_{NW}. \quad (16.13)$$

The contact area formed is an ellipse of eccentricity $e = \sqrt{1 - b^2/a^2}$, where a and b are the minor and major semi-axes of the ellipse of contact. With

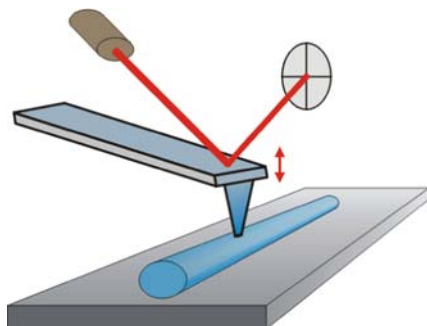


Fig. 16.6. The normal tip–NW contact stiffness can be determined from CR-AFM measurements on top of the NW. This contact stiffness can be then converted into the radial indentation modulus of the NW (Reprinted with permission from Stan et al. [41]. Copyright 2007 American Chemical Society)

known R_T and R_{NW} , the eccentricity e can be readily determined from

$$\frac{R_{NW}}{R_T + R_{NW}} = (1 - e^2) \frac{K(e) - E(e)}{E(e) - (1 - e^2) K(e)}, \quad (16.14)$$

where $K(e) = \int_0^{\pi/2} 1/\sqrt{1 - e^2 \sin^2 \theta} d\theta$ and $E(e) = \int_0^{\pi/2} \sqrt{1 - e^2 \sin^2 \theta} d\theta$ are the complete elliptical integrals of first and second order, respectively.

The relative displacement between the deformed tip and NW is given by

$$\delta_n = \left(9\gamma F_n^2 / 16R_T E^{*2}\right)^{1/3} \quad (16.15)$$

and the major semi-axis of the ellipse of contact by

$$a = \frac{K(e)}{\pi} (6F_n R_T / \gamma E^*)^{1/3}, \quad (16.16)$$

with the reduced elastic modulus E^* as defined by (16.6). The geometrical factor γ is $2(1 - e^2) K^3(e) (2R_{NW} + R_T) / \pi^2 E(e) R_{NW}$ and converges to 1 for a sphere-on-flat contact geometry. By taking the derivative of the expression (16.15) with respect to F_n , the normal contact stiffness becomes

$$k_{n, NW} = \pi a E^* / K(e). \quad (16.17)$$

In a simple experimental setup, successive CR-AFM measurements can be performed on top of the NW and on a flat reference, which could be the supporting substrate. The measured contact frequencies are then converted into corresponding contact stiffnesses: $k_{n, NW}$ given by (16.17) and $k_{n, ref}$ given by (16.5). From the ratio of these two contact stiffnesses, the indentation modulus of the NW, M_{NW} , can be readily extracted:

$$1/M_{NW} = \gamma^{-1/2} \left(k_{n, ref} / k_{n, NW}\right)^{3/2} (1/M_{ref} + 1/M_T) - 1/M_T. \quad (16.18)$$

Here, M_{NW} characterizes the elastic response of the NW along the direction of the applied load, and is thus the radial indentation modulus of the NW.

16.4.2 Lateral Contact Stiffness of the Tip–Nanowire Contact

The lateral contact stiffness, k_{lat} , can be observed in the spring response of the contact to a laterally applied force, F_{lat} , in the plane of contact. In the nonslip elastic regime, the lateral deformation of a Hertzian contact is proportional to the applied lateral force [78], and the lateral contact stiffness is independent of the lateral force but depends on the normal force through the size of the contact area. For the above described tip–NW contact and the direction of F_{lat} along the NW axis, the lateral contact stiffness has the same expression

$k_{\text{lat}} = 8aG^*$ as (16.9) with a given by (16.16) and the reduced shear modulus given by

$$G_{\text{T,NW}}^* = \frac{\pi}{4} \left[\frac{K(e) - \nu_{\text{T}}B(e)}{G_{\text{T}}} + \frac{K(e) - \nu_{\text{NW}}B(e)}{G_{\text{NW}}} \right]^{-1}, \quad (16.19)$$

where G_{T} , ν_{T} and G_{NW} , ν_{NW} the shear and Poisson's ratio of the tip and NW, respectively, and $B(e) = [E(e) - (1 - e^2)K(e)]/e^2$. By comparing (16.5) with (16.17), it can be observed that in the expression for the normal contact stiffness in the tip-on-NW case, the reduced elastic modulus has the same expression as for the tip-on-flat case, the difference between the normal contact stiffness formulas being in the geometrical prefactor. However, in establishing this correspondence for the lateral contact stiffness, the formulas are the same in both tip-on-flat and tip-on-NW cases, but the geometrical corrections in the tip-on-NW case cannot be factorized anymore (they are mixed into the reduced shear modulus (16.19) definition).

As in the case of normal contact stiffness measured by CR-AFM, a cantilever calibration can also be avoided in the case of lateral contact stiffness measurements by taking the ratio of the lateral stiffness measured on the sample tested (NW) and on a reference (flat substrate). Thus, the reduced shear modulus of the tip-NW contact can be expressed as a function of the reduced shear modulus of the tip-flat contact and the ratio of the lateral contact stiffnesses of these two contacts:

$$G_{\text{T, NW}}^* = \frac{\pi}{2k(e)} \left(\frac{\gamma E_{\text{T, NW}}^*}{E_{\text{T, ref}}^*} \right)^{1/3} \frac{k_{\text{lat, NW}}}{k_{\text{lat, ref}}} G_{\text{T, ref}}^* \quad (16.20)$$

In Fig. 16.7, the load dependences are shown for the contact radius, deformation along the loading direction, and normal and lateral contact stiffnesses of the tip-sample contact as a function of radius of curvature of the NW tested, $R_{\text{NW}} = 5, 20, 50$ nm, infinity (flat surface). The tip radius was assumed as $R_{\text{T}} = 15$ nm. Both tip and samples were assumed to be Si with $E_{\text{Si}} = 130$ GPa, $\nu_{\text{Si}} = 0.28$, and $G_{\text{Si}} = 79.4$ GPa. As can be seen, significant deviations in the contact stiffness are caused by the contact geometry as a function of NW radius. By omitting proper consideration of the contact geometry contribution to the measured contact stiffnesses, elastic modulus values will be erroneously calculated.

16.5 Elastic Moduli of ZnO and Te Nanowires Measured by CR-AFM

16.5.1 CR-AFM Measurements on ZnO Nanowires

Zinc oxide nanostructures hold tremendous promise for novel and versatile devices due to their intrinsic semiconducting and piezoelectric properties, as well as the diverse structural morphologies that can be achieved by

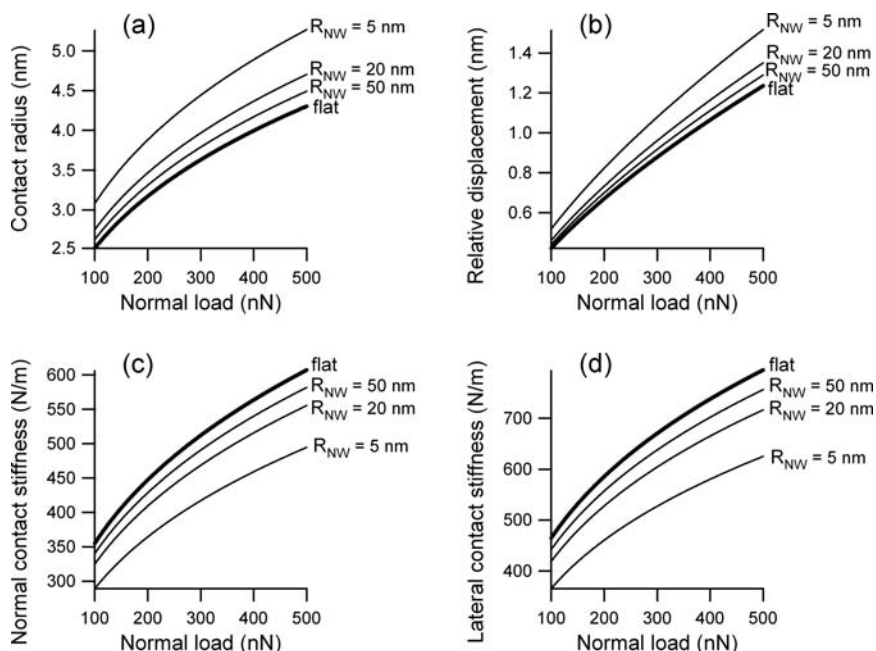


Fig. 16.7. Load dependences of the (a) contact radius, (b) normal deformation, (c) normal contact stiffness, and (d) lateral contact stiffness of the tip-on-flat contact as well as tip-on-NW contacts. The tip radius was assumed to be 15 nm and the nanowire radius 5, 20, and 50 nm, respectively. Both tip and samples are Si

current synthesis techniques. ZnO nanostructures can take the form of NWs, NBs, nanorings, and nanohelices, with demonstrated potential for novel nanoscale piezoelectric devices [84]. The anticipated performance of such devices depends, at least in part, on having knowledge of the structural and mechanical properties of ZnO nanostructures. The small dimensions of these nanostructures can, however, raise serious challenges for experimental investigations of properties relevant to device applications. Indeed, a quick inspection of Table 16.1 illustrates the conflicting results reported so far by different techniques on the elastic modulus of ZnO 1D nanostructures: The Young's modulus of ZnO NWs was found in the range of 30–250 GPa, with or without a clear dependence of the NW diameter. Unlike most of the techniques used to measure the mechanical properties of 1D nanostructures, CR-AFM doesn't require any specimen manipulation or special test geometries. It is therefore ideal for local elastic modulus measurements of simply substrate-supported 1D nanostructures and has been successfully used to quantify the size-dependence of the elastic modulus of ZnO NWs [41].

ZnO NWs were grown on Au catalyst layers deposited on SiO₂ substrates through thermal evaporation of commercial Zn powder at 550 °C in an N₂-O₂ gas mixture. The NWs were then cleaved from the substrate, suspended

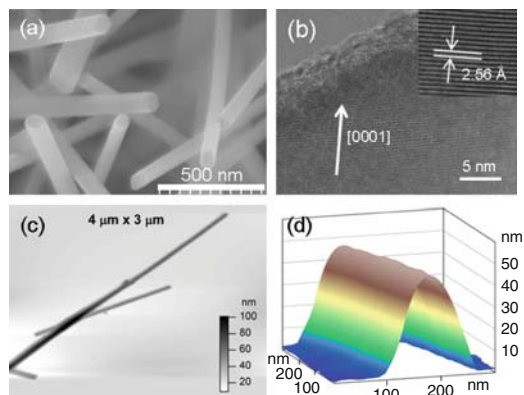


Fig. 16.8. Morphological characterization of ZnO NWs used for CR-AFM measurements: (a) SEM image of as-grown ZnO NWs. (b) HRTEM image of the outer layers of a ZnO nanowire showing the spatial periodicity of the (0002) atomic planes. (c) AFM image of laid-down ZnO NWs. (d) Detailed AFM profile of a 300 nm segment of a ZnO nanowire (Reprinted with permission from Stan et al. [41]. Copyright 2007 American Chemical Society)

in ethanol, ultrasonicated, and dispersed on clean Si(111) wafers. SEM and AFM images of individual NWs before and after their removal from the substrate are shown in Fig. 16.8a and c. High-resolution transmission electron microscopy (HRTEM) and electron diffraction analysis revealed that the ZnO NWs were single crystals grown along [0001] with an inter-planar spacing of 2.55 Å between the adjacent (0002) planes of ZnO wurtzite (see Fig. 16.8b). Due to the growth procedure and conditions, the NWs had rounded cross-sections (Fig. 16.8a and d) rather than hexagonal ones. From AFM images, the radius of the NW and that of the probe tip (approximated as a spherical surface) can be calculated by deconvoluting the contributions of the tip and the ZnO NW to the average cross-sectional AFM profile (Fig. 16.8d).

The diameter dependence of the radial elastic modulus of ZnO NWs was investigated by performing CR-AFM measurements on top of ZnO NWs of various diameters dispersed on a piece of Si(111) wafer. The measurements on each tested NW were bracketed by measurements, performed in the same experimental conditions, on the Si(111) substrate to assure a consistent calibration of the NWs elastic moduli. Moreover, no significant change was detected in the contact resonance frequencies measured on the substrate before and after testing the ZnO NW, which confirms that the contact geometry was preserved during measurements, and thus the results were reproducible. An example of this is shown in Fig. 16.9a for CR-AFM measurements performed to determine the elastic modulus of a 46 nm diameter ZnO NW. As explained in Sect. 16.3.3, sampling arrays were collected and the average values used to improve the precision of measurements. With the resonance frequencies identified, the normal contact stiffness $k_{n, \text{NW}}$ on each tested NW was

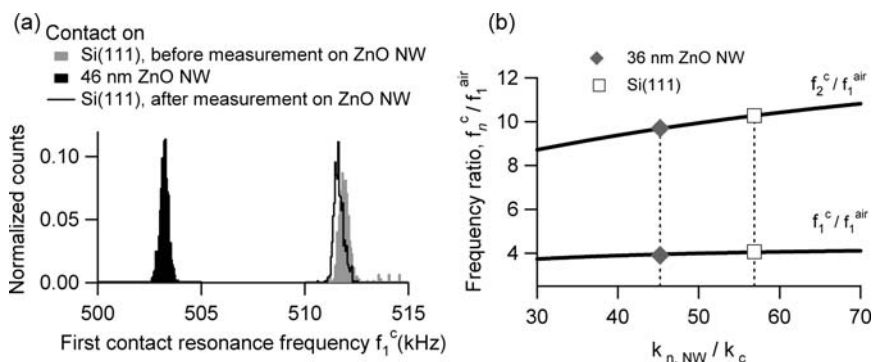


Fig. 16.9. (a) Histograms of the first contact resonance frequency measured on a 46 nm ZnO NW and on the Si(111) reference substrate. (b) The clamped-coupled beam model is used to convert the measured contact resonance frequencies into normalized contact stiffness (here shown for measurements made on a 36-nm diameter ZnO NW) (Reprinted with permission from Stan et al. [41]. Copyright 2007 American Chemical Society)

calculated using a clamped-spring coupled cantilever model [58]. In Fig. 16.9b, this procedure is shown for the measured contact resonance frequencies on a 36 nm diameter ZnO NW and Si(111) substrate as plotted on the theoretical curves of contact resonance frequencies versus contact stiffness. The contact resonance frequencies f_1^c and f_2^c were expressed in terms of the fundamental resonance frequency in air, f_1^{air} , and the contact stiffness $k_{n, \text{NW}}$ was normalized to the cantilever stiffness, k_c . At least two resonance frequencies (usually the first two) are used in this calculation to eliminate the uncertainty introduced by the position of the tip along the cantilever. Considering the indentation modulus of the reference material as known, the normal contact stiffnesses $k_{n, \text{NW}}$ and $k_{n, \text{ref}}$ measured on the NW probed and the Si(111) substrate are used in (16.18) to determine the radial indentation modulus of the NW.

With the same experimental setup, friction-type measurements can be used for determining the tangential shear modulus of the NW probed. Lateral CR-AFM measurements can be also used for shear modulus measurements, in which case an extra piezo-actuator mounted beneath the substrate provides the torsional oscillation of the cantilever probe in contact with the specimen [79]. Here, the friction-type measurements were simply performed with a nonvibrating tip dragged back and forth along the top surface of the NWs (see Fig. 16.10a), at a speed of 10 nm s^{-1} and under a constant normal force of 250 nN. The lateral contact stiffness itself is characterized in the no-slip linear elastic regime. In this regime, the variation of the lateral force is proportional to the lateral voltage signal V from the photodiode (Fig. 16.10a) and is recorded as a function of the displacement along the materials studied (Fig. 16.10b). Using a calibration factor, the slope of the initial part of friction

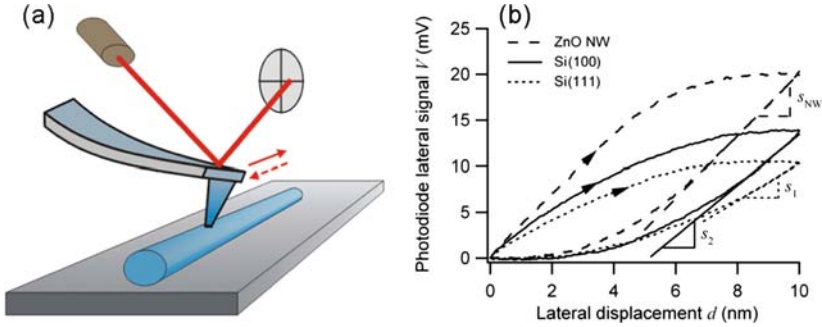


Fig. 16.10. (a) Friction-type measurements on top of a ZnO NW. (b) The initial slope of lateral force-displacement curves is used to calculate the tangential shear modulus. The arrows on the friction loops indicate the scan direction. As illustrated in (a), the lateral force is proportional to the lateral signal from the photodiode (Reprinted with permission from Stan et al. [41]. Copyright 2007 American Chemical Society)

loops (Fig. 16.10b) provides the lateral stiffness κ of the cantilever-tip-sample system. As measured, κ includes not only a contribution k_{lat} from the contact between the tip and the material but also a contribution κ_p that comes from twisting the AFM probe [85],

$$1/\kappa = 1/k_{\text{lat}} + 1/\kappa_p. \quad (16.21)$$

The tangential shear modulus G_{NW} can be extracted from the determined lateral contact stiffness k_{lat} by means of (16.9) and (16.19). In order to avoid the calibrations required for κ_p and the lateral signal of the photodiode detector in (16.21), additional measurements on two reference materials can be employed. Two sets of equations, (16.20) and (16.21), can be used to couple the measurement made on a ZnO NW and the two flat-surface reference materials (e.g., Si(100) and Si(111)). This allows the determination of G_{NW} in terms of the slopes of the linear parts of the friction loops measured on NW and the two reference materials (refer to Fig. 16.10b).

The results of CR-AFM and friction-type measurements on ZnO NWs of diameters in the 25–150 nm range are shown in Fig. 16.11. The radial Young's modulus E_{NW} was determined from the radial indentation modulus M_{NW} in the isotropic approximation, $E_{\text{NW}} = M_{\text{NW}}(1 - \nu^2)$, with the Poisson's ratio $\nu = 0.3$. Both the radial Young's modulus and the tangential shear modulus of the investigated ZnO NWs exhibit significant size dependence for wire diameters smaller than 80 nm. In the limit of very thick NWs, the values of bulk elastic moduli are approached. Here, the bulk limits, 110 GPa for the Young's modulus and 50 GPa for the shear modulus, are obtained from measurements on a 250×450 nm rectangular cross-section ZnO NB (the measurements are referred to as bulk in Fig. 16.11). This value for the bulk Young's modulus compares well with the experimental value of 112 GPa obtained from

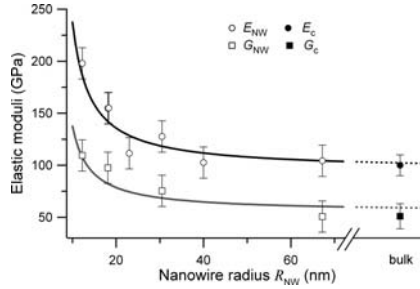


Fig. 16.11. Elastic moduli E_{NW} and G_{NW} as determined from CR-AFM and friction-type measurements on ZnO NWs of various diameters. The fitting curves were obtained by modeling the NW structure as composed of two coaxial cylindrical parts: a core made of bulk material and a shell of stiffer material relative to the bulk (Reprinted with permission from Stan et al. [41]. Copyright 2007 American Chemical Society)

nanoindentation measurements on bulk single-crystal ZnO [86], but is smaller than the theoretical value 140 GPa.

The size dependence of the elastic properties of ZnO NWs can be understood qualitatively by recognizing that the thinner NWs have a relatively larger number of atomic bonds in the surface and near-surface layers and that these bonds are shorter and stiffer than their bulk counterparts [87]. A more quantitative approach stems from modeling the NW [10] as a cylindrical core (bulk) and a shell (surface region) that have different radial elastic moduli E_{core} and E_{shell} , respectively. For a core-shell NW with a shell thickness s and a core radius $R_{core} = R_{NW} - s$, an approximate estimate of the radial Young's modulus E_{NW} can be gained from the analysis of strain under conditions of uniform radial stress, which yields

$$R_{NW}/E_{NW} = s/E_{shell} + R_{core}/E_{core}. \quad (16.22)$$

Also, the strain analysis for a NW subjected to a shear load distributed uniformly (for simplicity) on its lateral surface yields a similar equation for the shear modulus G_{NW}

$$R_{NW}/G_{NW} = s/G_{shell} + R_{core}/G_{core}. \quad (16.23)$$

In both (16.22) and (16.23), the Poisson's ratio of the core and shell parts was assumed to be the same.

Over the investigated ZnO NWs diameter range, the core-shell description provided by (16.22) was found to fit the Young's modulus measurements for the following parameters: $E_{core} = 95$ GPa, $E_{shell} = 190$ GPa, and $s = 12$ nm. The fit curve is plotted in Fig. 16.11 and shows that in the bulk limit, the radial Young's modulus approaches the value of E_{core} . The value $s = 12$ nm is consistent with estimations of the shell thickness based on HRTEM images

(Fig. 16.8b) which show a polycrystalline surface layer that is at least 5-nm thick. Also, this estimation of the shell thickness, $s = 12$ nm, is in order-of-magnitude agreement with an independent assessment based on a core-shell model for the axial Young's modulus of ZnO NWs [10]. Such an assessment, which places the shell thickness at 4.4 nm [10], is the result of extracting the effective axial Young's modulus of the NW from the sum of the flexural rigidities of a core and a shell that have different axial moduli. For both the axial Young's modulus [10] and the radial modulus [41], suitably designed core-shell models appear to satisfactorily capture the increase of the corresponding elastic modulus with decreasing NW diameter. However, in both of these core-shell models, the shell thickness parameter ends up being the radius of the smallest NW investigated, and the value E_{core} closely corresponds to the thinnest NW. In neither of the models does the Young's modulus saturate at a constant E_{shell} as the diameter is decreased past a certain value, which means that the fits only hold within the diameter range studied. The reason for this most probably lies in the fact that the models, while providing reasonable interpretation of the experimental data, are perhaps too simplified.

With the shell thickness determined from fitting the E_{NW} values, the same parameter s was consistent with the shear modulus data. Using $s = 12$ nm in (16.22), the remaining fit parameters were found to be $G_{\text{core}} = G_{\text{shell}}/2 = 55$ GPa. The corresponding fitting curve is also shown in Fig. 16.11. Over the investigated diameter range, the shear modulus exhibits similar diameter dependence as the Young's modulus does. Thus, the shear modulus measured on NWs thicker than 100 nm approaches the bulk limit value, G_{core} , and monotonically increases in the limit of thin NWs. The analysis here shows that the shell thickness $s = 12$ nm estimated from fitting the Young's modulus data is also consistent with a core-shell model that describes the shear modulus.

The investigated ZnO NWs have a polycrystalline shell (Fig. 16.8b) that cannot be in epitaxial relation with the core of the NW. This means that the stress created at the interface could effectively modify the elastic properties of the core. It is even conceivable that the polycrystalline formation compensates the surface effects, which otherwise could exert an even stronger effect on the elastic properties of NWs. With or without a polycrystalline shell, the strain experienced by the surface layers due to reconstruction and relaxation depends on the diameter of NWs, and thus, the concept of a shell with constant elastic modulus may only hold as an approximation.

16.5.2 CR-AFM Measurements on Te Nanowires

Over the past few years, the synthesis of tellurium (Te) nanostructures in the form of NWs, NBs, and NTs has received distinct attention [88]. Such synthesis efforts are aimed at exploiting various electrical, photoelectric, thermoelectric, and piezoelectric properties of Te [89] in developing new nanoscale electronic and optoelectronic devices. As is the case with any 1D nanostructure, accurate

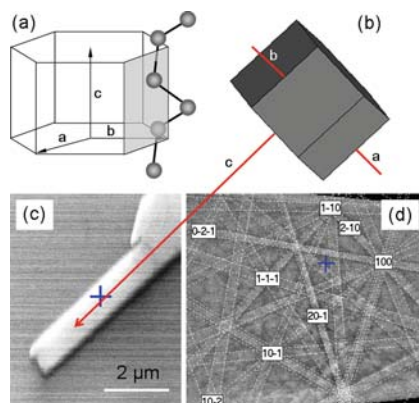


Fig. 16.12. (a) The crystallographic structure of Te is generated by attaching to each node of its hexagonal lattice three atoms along helical chains that are parallel to the c -axis ($[0001]$ direction). (b) and (c) The orientation of the Te hexagonal cell as found from EBSD made on the top facet of a nanowire. (d) EBSD pattern used to establish the crystallographic orientation shown in (c)

knowledge of the mechanical properties of those made of Te is required prior to integration into functional devices.

Te nanostructures were grown on Si(100) substrates by a thermal vapor-transport deposition method. At 180°C , the deposition mainly provided long ($2\text{--}10\ \mu\text{m}$) Te NWs of uniform cross-section. The morphology of the Te NWs so obtained was examined using a field emission scanning electron microscope (FESEM) equipped with an electron backscatter diffraction (EBSD) detector (see Fig. 16.12). EBSD patterns on the top facets of NWs (e.g., Fig. 16.12d) matched simulated patterns for Te structure and showed that the NWs growth direction was the c -axis; all the investigated top facets were found to be in the same $\{10\bar{1}0\}$ plane of the Te hexagonal lattice. Also, single-crystal structure, (0001) growth direction, and faceting of Te NWs were revealed in HRTEM. In Fig. 16.13 are shown such NWs at different magnifications: A FESEM image in Fig. 16.13a encompasses many Te 1D nanostructures; a higher magnification AFM image of a Te T -junction made by a NB and a NW is shown in Fig. 16.13b; and in Fig. 16.13c a 500-nm segment of the wire length from Fig. 16.13b is further detailed by AFM scanning. The wire cross-section as observed by AFM shown in Fig. 16.13c is rectangular, and this was the case for all the Te NWs investigated here. As the vertical profile of the walls is not resolved accurately by AFM, the observed rectangular cross-section could, in fact, originate from an elongated hexagonal section.

After Te NWs are localized on the substrate by tapping-mode AFM, CR-AFM measurements can be performed back and forth on top of each NW and the Si(100) substrate. Examples of such measurements are shown in Fig. 16.14. Figure 16.14a shows the shifts in the resonance frequencies of the AFM

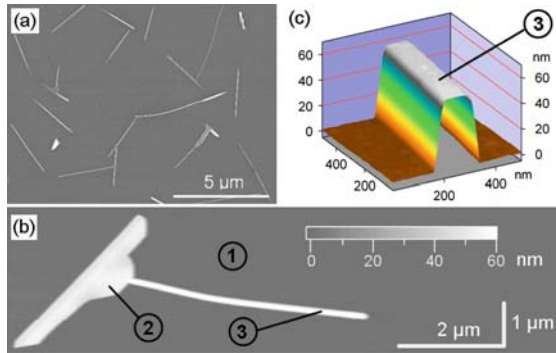


Fig. 16.13. (a) FESEM image of Te NWs on Si(100) substrate. (b) CR-AFM measurements are performed at locations 1, 2, and 3, to measure the contact stiffness on Si(100) substrate, Te NB, and Te NW, respectively. (c) A 500 nm segment of the investigated Te NW is shown in detail using AFM (Reprinted with permission from Stan et al. [44]. Copyright 2008 American Institute of Physics)

cantilever from air to contact on a 25-nm thick Te NW and Si(100) substrate, respectively. In Fig. 16.14b are shown CR-AFM measurements performed at the locations indicated in Fig. 16.13b, and clear differences are observed not only between Si(100) and Te NW but also between a Te NW and a NB of the same height. The contact resonance frequencies measured on each NW and the Si(100) substrate are used to calculate the normalized contact stiffness, $k_{n,NW}/k_c$, which is then converted into the indentation modulus by means of (16.7). In addition to the contact stiffness values, $M_{\text{Si}(100)} = 164.8 \text{ GPa}$ needs to be used in (16.7) for the indentation modulus of the reference substrate and the probe tip.

As the thickness of NWs decreased, the measured contact resonance frequencies were observed to shift to greater and greater values. In terms of the elastic response, this indicates an increase in the elastic modulus, as can be seen in Fig. 16.15a. Starting from values slightly greater than 45 GPa for NWs thicker than 100 nm, the increasing trend in the indentation modulus plateaued around 85 GPa for NWs with diameters less than 30 nm. A similar enhancement factor of 1.5–2 has been measured, by different techniques, for the size-dependent elastic modulus of other metallic (Ag and Pb [40, 46]) as well as nonmetallic (ZnO [10, 18, 41] and CuO [28]) NWs. However, the behavior exhibited by Te NWs differs from that reported for other types of NWs for which the elastic modulus was found to increase continuously with the reduction in diameter.

The observed size dependence in the mechanical properties of these nanostructures can be understood by considering a surface stiffness effect [10, 40, 41] and the large surface-to-volume ratio at the nanoscale [87]. It has been shown in atomistic simulations that a significant tensile strain accumulates in the near-surface layers of NWs as a result of the reduced atom coordination and

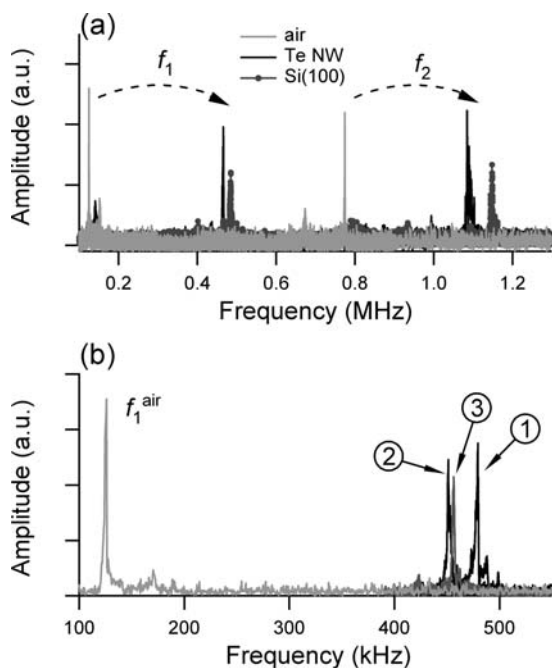


Fig. 16.14. (a) The shift in the resonance frequency of the first two vibrational CR-AFM modes from air to contact on a 25-nm thick Te NW and Si(100). (b) The shift in the first resonance frequency at each of the positions indicated in Fig. 16.13b: (1) on Si(100) substrate, (2) on Te NB, and (3) on Te NW. Although the NB and NW are of the same height, different CR-AFM frequencies were measured; this suggested a width-to-thickness dependence of the elastic modulus of the two nanostructures probed (Reprinted with permission from Stan et al. [44]. Copyright 2008 American Institute of Physics)

distorted symmetry [90, 91]. This determines, in turn, a progressive lattice contraction from outside to inside and places the inner core in compression. The accumulated strain leads thus to nonlinear elastic effects such that the surface is stiffened relative to the bulk. To quantify the contribution of surface stiffening to the observed axial Young's modulus of ZnO NWs, Chen et al. [10] considered a NW composed of a NW core made of bulk material and a stiffer NW shell of constant thickness, coaxial with the core. Following this interpretation, a linear variation of the shell thickness s as a function of the NW thickness t can be assumed over the elastic modulus size-dependent range in our measurements as follows [44]:

$$2s = t_{\min}(t_{\max} - t) / (t_{\max} - t_{\min}), \quad (16.21)$$

where t_{\max} and t_{\min} approximately delimit the size-dependent elastic modulus range: t_{\max} is the NW thickness above which the shell is of negligible thickness and t_{\min} is the NW thickness below which the measured elastic modulus is that

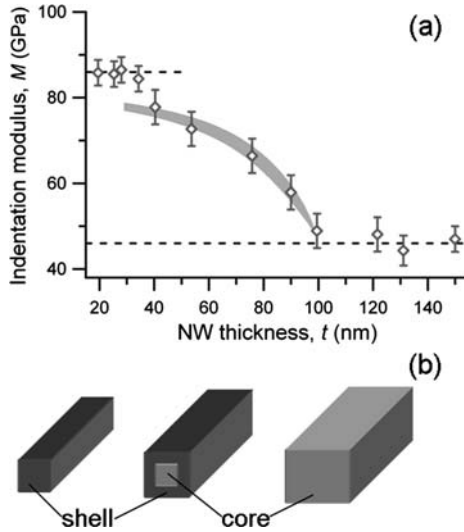


Fig. 16.15. (a) Indentation modulus of Te NWs as a function of thickness. The dashed lines indicate the maximum (very thin NWs) and minimum (very thick NWs) extreme values. Over the variable modulus range, the indentation modulus of the core–shell nanostructure was calculated within the gray band. (b) Schematic diagram of the core–shell structure from very thick (mostly core) to very thin NWs (mostly shell) (Reprinted with permission from Stan et al. [44]. Copyright 2008 American Institute of Physics)

of the stiffened shell; core–shell diagrams for the considered thickness range are shown in Fig. 16.15b. Owing to the simple contact geometry, spherical tip on flat top of Te NWs, the elastic contributions of the core and shell can be separated in the indentation modulus of the NW, similar to the case of an indented coated substrate [92]:

$$M_{\text{NW}} = I(\xi) M_{\text{shell}} + [1 - I(\xi)] M_{\text{core}}. \tag{16.22}$$

M_{core} and M_{shell} are the indentation moduli of the core and the shell, respectively, and $I(\xi) = (2/\pi) \arctan(\xi) + [(1-2\nu)\xi(1 + \xi^{-2}) - \xi/(1 + \xi^2)] / 2\pi(1 - \nu)$ is a weighting function of ξ and Poisson’s ratio ν (here taken to be the same for both core and shell); $\xi = s/a$ is the shell thickness s normalized to the contact radius a .

The above contact mechanics considerations are valid in interpreting the measured indentation modulus on Te NWs in the thickness range, where the shell is presumably thinner than the core. Thus, with $M_{\text{shell}} \cong 85$ GPa and $M_{\text{core}} \cong 46$ GPa determined from the two limiting regions, a good fit is provided by (16.22) in the transition region with fit parameters of $t_{\text{min}} = 29$ nm, $t_{\text{max}} = 100$ nm, and $a = 4.0$ nm. The fit given by the model is within the gray band in Fig. 16.15a as the Poisson’s ratio is varied from 0.1 (upper edge) to 0.4 (lower edge). The applicability of the model is restricted

from the domain of thin NWs in which the core is reduced in thickness so much that it cannot be considered as a substrate for the indented shell. The fit parameter $a = 4.0$ nm agrees well with the contact radius $a_{\text{Hertz}} = 4.2$ calculated with the Hertz model, $a_{\text{Hertz}} = (3F_n R_T / 4E^*)^{1/3}$, for the experimental parameters used: $F_n = 250$ nN, $E^* = 50$ GPa, $R_T = 20$ nm. As the sensing depth of CR-AFM is about three times the contact radius, the substrate contribution can be ruled out for all the NWs probed.

A schematic diagram of the core-shell partition of the NW cross-section is shown in Fig. 16.15b for different NW thicknesses. Very little variation in the elastic modulus is observed for NWs thinner than 30 nm, suggesting that when the shell extends over the entire NW thickness, the attendant lattice contraction and disorder is less of an energy penalty than that required to reach the equilibrium inter-atomic order of the bulk core. The average thickness of the shell for the Te NWs measured here, $t_{\text{min}}/4 = 7.2$ nm, is comparable with the values reported previously for ZnO NWs, 4.4 [10] and 12 nm [41], respectively. Here, a caveat regarding overinterpretation is in order: A different shell thickness is implied by the 12% reduction in indentation modulus observed in the measurements for a wide NB compared with a NW of the same thickness (locations 2 and 3 in Figs. 16.13b and 16.14b). This observation, consistent with the measured width-to-thickness dependence of the elastic modulus of ZnO NBs [93], suggests an aspect-ratio dependence for moduli, as well as a scale dependence for nanostructures.

Although the above interpretation is more intuitive than fundamental, it captures the size-dependent effects observed in elastic modulus measurements on Te NWs and provides a simple explanation for the surface stiffness enhancement in NWs of reduced diameters that must limit more physically based explanations drawn from theoretical models or simulations.

16.6 Surface Effects on the Mechanical Properties of 1D Nanostructures

The elastic moduli of bulk materials are independent of component size and are regarded as intrinsic material properties. This is because surface effects, although existent, are negligible in the case of macroscopic structures and the major contribution to modulus is the stiffness of atomic bonds within the volume of the material. Contrarily, surface effects become significant in the case of structures with at least one dimension constrained to the nanoscale, leading to large surface-to-volume ratios. Compared to bulk values, the elastic moduli of 1D nanostructures are modified by surface proximity through boundary effects (surface relaxation, surface stress, surface reconstruction, etc.) and surface morphology effects (surface roughness, surface oxidation, surface defects, etc.). It is suggested by theory [90, 91, 94, 95] and observed in experiments [10, 18, 40, 41, 44, 46] that mechanical properties of 1D nanostructures are size-dependent. There are, however, serious discrepancies when

theoretical results (from analytical models or atomistic simulations) are compared with experimental measurements. The difficulty in identifying general behavior in the size-dependence of the elastic moduli of 1D nanostructures stems from the fact that the surface effects mold the remaining bulk contribution to generate a collective structure response. The elastic modulus of 1D nanostructures is then no longer a material property, but a characteristic of a particular structure with dependence on material, size, and shape.

To explain the observed deviations of the elastic moduli of 1D nanostructures from bulk values, analytical approaches [10, 40, 46, 94, 95] have been proposed to incorporate surface contributions into existing continuum mechanics models and, in this way, recover the nanostructure properties from those of a bulk body with surfaces. These corrections consider surface stress and surface energy to be the key factors responsible for the mechanical behavior of 1D nanostructures. A state of minimum energy is achieved in the outer layers of 1D nanostructures through atomic lattice relaxation that progresses from the surface to the equilibrium bulk order. Reduced atomic coordination, surface reconstruction, and distorted symmetry are accommodated in this surface relaxation. As a result, significant surface stresses are developed in the outer layers and their elastic properties changed. With surface contributions included, the elastic modulus of 1D nanostructures is referred to as the apparent [40] or effective [10] elastic modulus and, ideally, should be observable in measurements.

Based on these considerations, NWs are modeled as composite structures composed of an inner part (core) and surface layers (shell). The core has the same elastic properties as the bulk material, whereas the shell is assumed to accommodate morphological structure changes arising from surface energy and surface stress effects. There is no defining cutoff depth to delimit the shell from the core, and in most cases the shell thickness is perceived in models as a fit parameter. The shell is assumed to have either (1) a constant [10, 41] or diameter-dependent thickness [44] and a constant Young's modulus or (2) a negligible thickness compared with the NW diameter and a constant surface stress [96]. Depending on the approach used, either the elastic Young's modulus or the surface stress of the shell is considered, more or less, as an adjustable parameter in fitting the diameter dependence of the elastic response of the NW. The predicted enhanced surface stiffness was asserted in interpretation of the measured size-dependent elastic modulus of NWs [10, 40, 41, 44, 46, 96].

In atomistic simulations, an interatomic potential, assumed to properly describe the properties of the material studied, is used to model the investigated structure. By allowing the energy of the system to decrease and the simulated structure to relax, the structure self-adjusts into equilibrium configurations. The observable quantities of interest (stress, strain, elastic modulus, etc.) are then calculated using the laws of thermodynamics. Notably, in atomistic simulations, the lattice relaxation occurs naturally in the surface layers, whereas in analytical models, the surface effects have to be more or less imposed. As revealed in some measurements (ZnO NWs [10], ZnO NWs [41],

Ag and Pb NWs [40], Ag NWs [46], Te NWs [44], and ZnO NWs [18]), atomistic simulations for mechanical properties of certain 1D nanostructures (Au NWs [95], ZnO NBs [90], Ni NWs [91], Ag NWs [97], and ZnO NWs [18]) indicated an enhanced stiffness of the surface layers compared to the bulk behavior. However, there are also theoretical simulations that predict more compliant behavior of other 1D nanostructures (Si NWs [98], Ni NWs [99], W NWs [100]) as their cross-section is reduced. Such size-dependent compliance was reported in experiments as well (Si cantilevers [101], GaN NWs [102]).

To date a well-argued correspondence between experiments and simulations is still missing. Even when experimental and atomistic simulation results were compared side by side for the same 1D nanostructures (Ag NWs [97]), the size-dependence of the elastic modulus was found over different diameter ranges. Thus, atomistic simulations predict that NWs with diameter thicker than about 10 nm have the same elastic modulus as their bulk counterparts and a size-dependent elastic modulus occurs only for NWs with diameters less than about 10 nm. On the other hand, in experiments, variations in the elastic modulus have been reported for NWs thinner than about 100 nm in diameter. There are also reports of good agreement between measurements and atomistic simulation results over adjacent NW diameter ranges (ZnO NWs [18]).

In establishing the missing link between theoretical simulations and measurements of elastic modulus of 1D nanostructures, caution should be taken on each side. On one hand, various techniques and measurements provided conflicting results: e.g., values in the range 30–200 GPa were reported for the elastic modulus of ZnO NWs, either size-dependent [10, 18, 41] or size-independent [8, 30, 35]. Are the NWs different in each experiment? Are the measurements faulty? Is one technique more adequate than the others for specific nanostructures tested? On the other hand, in simulations, the surface energy is mainly the only factor acknowledged to cause changes in the elastic modulus of NWs as a function of diameter. Is the chosen interatomic potential providing an accurate description of the existing surface stress? And to what extent should surface defects, contaminants, oxide layers, surface roughness, etc. be considered in the model? Did the computational experimental procedure affect the accuracy of simulation results? The observed discrepancies can be addressed by acknowledging the individual contributions of various surface effects both in experiments and atomistic simulations. In experiments, the contribution of one or another surface effects can be isolated in measurements on 1D nanostructures exhibiting a particular enhanced surface effect (e.g., elastic modulus measurements on NWs of the same diameter but of different surface roughness). In atomistic simulations, all the details of a given measured specimen have to be considered in the model (e.g., anisotropic and inhomogeneous crystallographic structure, surface chemistry, and morphology).

16.7 How Important Are the Mechanical Properties of 1D Nanostructures in Applications?

Accurate knowledge of the elastic properties of 1D nanostructures is critical to the development of devices and materials with new capabilities that rely on the special properties of such structures - and will probably be the most important mechanical properties to measure at the nanoscale: Nanorobots that act as tools, sensors, or actuators will make mechanical application of 1D nanostructures, carbon NTs in particular [103,104]. As maintaining mechanical integrity is a direct requirement for robot performance, the deformation of 1D nanostructural components will be completely elastic. Similarly, electrical actuation of mechanical deformation of 1D nanostructures, e.g., carbon NT bundles [105], to form nanoscale electrical switches will require cyclic elastic deformation. Piezoelectric and semiconducting ZnO NWs and NW arrays capable of mechanical to electrical transduction to generate electrical power are deformed elastically, whether by direct contact [106] or by vibration, as in an energy-harvesting device [107]. This is also true of the larger-scale piezoelectric ZnO NB applications [108,109].

At the largest scales, macroscopic advanced composite structures that incorporate 1D nanostructures, e.g., carbon NTs in direct mechanical application [2] or for electroacoustic transduction [110] will place 1D nanostructures into static or cyclic elastic deformation. In the former case, from a strictly empirical point of view, accurate measurements of nanostructure modulus are essential for making predictions of composite elastic properties as a function of nanostructure morphology, orientation, or volume fraction incorporated into the composite matrix [111]. Such measurements and corresponding measurements of composite modulus will also be critical in determining whether composite processing and fabrication methods are optimized. Poor adhesion between the nanostructures and the matrix, poor dispersion of the nanostructures, and lack of orientation control of the nanostructures will all lead to decreases in the composite modulus that can only be interpreted if the elastic properties of the nanostructure reinforcement are known. In addition, in many cases composite design optimization will be made such that constraints from other properties requirements (e.g., electrical, thermal) or other factors (e.g., cost, environmental impact) must be considered. Accurate knowledge of the nanostructure elastic properties is thus critical for multi-dimensional optimization.

In devices that do not have an explicit mechanical function, dimensional stability of 1D nanostructures is critical to device performance, and thus deformation of the structures during device operation will be elastic. Examples include semiconducting Si NWs used in electrical applications [112], semiconducting NWs used in optoelectronic [113,114] or photovoltaic [115] applications, transparent NWs for photonics applications [116], and surface-functionalized Si or ZnO NWs for chemical and biological sensors [117–119].

In these last sensor applications, adhesion of the nanostructure to external species is a critical element of sensor performance. Adhesive properties are likely to be almost as important as elastic properties in terms of the mechanical properties that will determine nanostructure performance in applications. This is true for both nano-scale applications (adhesion of external elements to nanorobot components) and macro-scale applications (adhesion of the matrix to nanostructures in composites). Here again, however, accurate knowledge of the elastic properties of the nanostructure under consideration is required to determine adhesive properties: many models of adhesion require knowledge of the effective modulus of the contacting bodies in order to extract adhesion energies from pull-off forces [120–122].

Finally, as alluded to in Sect. 16.6, measurement of the elastic properties of 1D nanostructures provides a great scientific opportunity with great application potential. Atomistic simulations can now model structures of sufficiently large scale and experiments can measure mechanical properties at sufficiently small scales that comparison between experiment and model is reasonable. Accurate and precise measurements of elastic modulus of 1D nanostructures can thus provide input into the development of accurate inter-atomic potentials and atomic-scale models. Modulus is a key parameter for comparison with experiment in such model development as it is an equilibrium parameter that is accessible from model outputs either explicitly or by thermodynamic relations. Once established by verification with elastic modulus measurements, such models can then be applied to predict the structure and properties of other nanostructures and guide new materials and structure development.

References

1. M.S. Dresselhaus, Carbon nanotubes-Introduction. *J. Mater. Res.* 13, 2355–2356 (1998)
2. C.R. Welch, W.F. Marcuson, I. Adiguzel, Will supermolecules and supercomputers lead to super construction materials? *Civil Eng.* 78, 42–52 (2008)
3. O. Lourie, H.D. Wagner, Evaluation of Young's modulus of carbon nanotubes by micro-Raman spectroscopy. *J. Mater. Res.* 13, 2418–2422 (1998)
4. R. Tenne, P.M. Ajayan, Z.L. Wang, Y. Li, P. Yang, Foreword. *J. Mater. Res.* 21, 2709–2710 (2006)
5. P. Poncharal, Z.L. Wang, D. Ugarte, W.A. de Heer, Electrostatic deflections and electromechanical resonances of carbon nanotubes. *Science* 283, 1513–1516 (1999)
6. L. Gao, Z.L. Wang, Z. Bai, W.A. de Heer, L. Dai, M. Gao, Nanomechanics of individual carbon nanotubes from pyrolytically grown arrays. *Phys. Rev. Lett.* 85, 622–625 (2000)
7. Z.L. Wang, Z.R. Dai, R.P. Gao, Z.G. Bai, J.L. Gole, Side-by-side silicon carbide–silica biaxial nanowires: Synthesis, structure, and mechanical properties. *Appl. Phys. Lett.* 77, 3349–3351 (2000)
8. X.D. Bai, P.X. Gao, Z.L. Wang, E.G. Wang, Dual-mode mechanical resonance of individual ZnO nanobelts. *Appl. Phys. Lett.* 82, 4806–4808 (2003)

9. P. Jaroenapibal, D.E. Luzzi, S. Evoya, S. Arepalli, Transmission-electron-microscopic studies of mechanical properties of single-walled carbon nanotube bundles. *Appl. Phys. Lett.* 85, 4328–4300 (2004)
10. C.Q. Chen, Y. Shi, Y.S. Zhang, J. Zhu, Y.J. Yan, Size dependence of Young's modulus in ZnO nanowires. *Phys. Rev. Lett.* 96, 075505 (2006)
11. X.L. Wei, Y. Liu, Q. Chen, M.S. Wang, L.M. Peng, The very low shear modulus of multi-walled carbon nanotubes determined simultaneously with the axial Young's modulus via in situ experiments. *Adv. Funct. Mater.* 18, 1555–1562 (2008)
12. D.A. Dikin, X. Chen, W. Ding, G. Wagner, R.S. Ruoff, Resonance vibration of amorphous SiO₂ nanowires driven by mechanical or electrical field excitation. *J. Appl. Phys.* 93, 226–230 (2003)
13. M.F. Yu, O. Lourie, M.J. Dyer, K. Moloni, T.F. Kelly, R.S. Ruoff, Strength and breaking mechanism of multiwalled carbon nanotubes under tensile load. *Science* 287, 637–640 (2000)
14. Y. Zhu, H.D. Espinosa, An electromechanical material testing system for in situ electron microscopy and applications. *Proc. Natl. Acad. Sci. USA* 102, 14503–14508 (2005)
15. H.D. Espinosa, Y. Zhu, N. Moldovan, Design and operation of a MEMS-based material testing system for nanomechanical characterization. *J. Microelectromech. Syst.* 16, 1219–1231 (2007)
16. Z.W. Pan, S.S. Xie, L. Lu, B.H. Chang, L.F. Sun, W.Y. Zhou, G. Wang, D.L. Zhang, Tensile tests of ropes of very long aligned multiwall carbon nanotubes. *Appl. Phys. Lett.* 74, 3152–3154 (1999)
17. S. Hoffmann, F. Östlund, J. Michler, H.J. Fan, M. Zacharias, S.H. Christiansen, C. Ballif, Fracture strength and Young's modulus of ZnO nanowires. *Nanotechnology* 18, 205503 (2007)
18. R. Agrawal, B. Peng, E.E. Gdoutos, H.D. Espinosa, Elasticity size effects in ZnO nanowires – a combined experimental-computational approach. *Nano Lett.* 8, 3668–3674 (2008)
19. J.P. Salvetat, A.J. Kulik, J.M. Bonard, G.A.D. Briggs, T. Stöckli, K. Méténier, S. Bonnamy, F. Béguin, N.A. Burnham, L. Forró, Elastic modulus of ordered and disordered multiwalled carbon nanotubes. *Adv. Mater.* 11, 161–165 (1999)
20. A. Kis, D. Mihailovic, M. Remskar, A. Mrzel, A. Jesih, I. Piwonski, A.J. Kulik, W. Benoit, L. Forró, Shear and Young's moduli for MoS₂ nanotube ropes. *Adv. Mater.* 15, 733–736 (2003)
21. B. Wu, A. Heidelberg, J.J. Boland, Mechanical properties of ultrahigh-strength gold nanowires. *Nat. Mater.* 4, 525–529 (2005)
22. H. Ni, X. Li, H. Gao, Elastic modulus of amorphous SiO₂ nanowires. *Appl. Phys. Lett.* 88, 043108 (2006)
23. Q. Xiong, N. Duarte, S. Tadigadapa, P.C. Eklund, Force-deflection spectroscopy: a new method to determine the Young's modulus of nanofilaments. *Nano Lett.* 6, 1904–1909 (2006)
24. W. Mai, Z.L. Wang, Quantifying the elastic deformation behavior of bridged nanobelts. *Appl. Phys. Lett.* 89, 073112 (2006)
25. L.T. Ngo, D. Almécija, J.E. Sader, B. Daly, N. Petkov, J.D. Holmes, D. Erts, J.J. Boland, Ultimate-strength Germanium nanowires. *Nano Lett.* 6, 2964–2968 (2006)
26. B. Wu, A. Heidelberg, J.J. Boland, J.E. Sader, X. Sun, Y. Li, Microstructure-hardened silver nanowires. *Nano Lett.* 6, 468–472 (2006)

27. Y. Chen, I. Stevenson, R. Pouy, L. Wang, D.M. McIlroy, T. Pounds, M.G. Norton, D.E. Aston, Mechanical elasticity of vapour-liquid-solid grown GaN nanowires. *Nanotechnology* 18, 135708 (2007)
28. E.P.S. Tan, Y. Zhu, T. Yu, L. Dai, C.H. Sow, V.B.C. Tan, C.T. Lim, Crystallinity and surface effects on Young's modulus of CuO nanowires. *Appl. Phys. Lett.* 90, 163112 (2007)
29. H. Zhang, J. Tang, L. Zhang, B. An, L.C. Qin, Atomic force microscopy measurement of the Young's modulus and hardness of single LaB₆ nanowires. *Appl. Phys. Lett.* 92, 173121 (2008)
30. B. Wen, J.E. Sader, J.J. Boland, Mechanical properties of ZnO nanowires. *Phys. Rev. Lett.* 101, 175502 (2008)
31. E.W. Wong, P.E. Sheehan, C.M. Lieber, Nanobeam mechanics: elasticity, strength, and toughness of nanorods and nanotubes. *Science* 277, 1971–1975 (1997)
32. J. Song, X. Wang, E. Riedo, Z.L. Wang, Elastic property of vertically aligned nanowires. *Nano Lett.* 5, 1954–1958 (2005)
33. Y.L. Chueh, L.J. Chou, J. Song, Z.L. Wang, Mechanical and magnetic properties of Ni-doped metallic TaSi₂ nanowires. *Nanotechnology* 18, 145604 (2007)
34. X. Li, H. Gao, C.J. Murphy, K.K. Caswell, Nanoindentation of silver nanowires. *Nano Lett.* 3, 1495–1498 (2003)
35. G. Feng, W.D. Nix, Y. Yoon, C.J. Lee, A study of the mechanical properties of nanowires using nanoindentation. *J. Appl. Phys.* 99, 074304 (2006)
36. X. Li, X. Wang, Q. Xiong, P.C. Eklund, Mechanical properties of ZnS nanobelts. *Nano Lett.* 5, 1982–1986 (2005)
37. H. Ni, X. Li, Young's modulus of ZnO nanobelts measured using atomic force microscopy and nanoindentation techniques. *Nanotechnology* 17, 3591–3597 (2006)
38. X. Li, P. Nardi, C.W. Baek, J.M. Kim, Y.K. Kim, Direct nanomechanical machining of gold nanowires using a nanoindenter and an atomic force microscope. *J. Micromech. Microeng.* 15, 551–556 (2005)
39. S.J. Young, L.W. Ji, S.J. Chang, T.H. Fang, T.J. Hsueh, T.H. Meen, I.C. Chen, Nanoscale mechanical characteristics of vertical ZnO nanowires grown on ZnO:Ga/glass templates. *Nanotechnology* 18, 225603 (2007)
40. S. Cuenot, C. Frétiigny, S. Demoustier-Champagne, B. Nysten, Surface tension effect on the mechanical properties of nanomaterials measured by atomic force microscopy. *Phys. Rev. B* 69, 165410 (2004)
41. G. Stan, C.V. Ciobanu, P.M. Parthangal, R.F. Cook, Diameter-dependent radial and tangential elastic moduli of ZnO nanowires. *Nano Lett.* 7, 3691–3697 (2007)
42. Y. Zheng, R.E. Geer, K. Dovidenko, M. Kopycinska-Müller, D.C. Hurley, Quantitative nanoscale modulus measurements and elastic imaging of SnO₂ nanobelts. *J. Appl. Phys.* 100, 124308 (2006)
43. L. Muthaswami, Y. Zheng, R. Vajtai, G. Shekawat, P. Ajayan, R.E. Geer, Variation of radial elasticity in multiwalled carbon nanotubes. *Nano Lett.* 7, 3891–3894 (2007)
44. G. Stan, S. Krylyuk, A.V. Davydov, M. Vaudin, L.A. Bendersky, R.F. Cook, Surface effects on the elastic modulus of Te nanowires. *Appl. Phys. Lett.* 92, 241908 (2008)
45. G. Stan, C.V. Ciobanu, T.P. Thayer, G.T. Wang, J.R. Creighton, K.P. Purushotham, L.A. Bendersky, R.F. Cook, Elastic modulus of faceted aluminum nitride nanotubes measured by contact resonance atomic force microscopy. *Nanotechnology* 20, 035706 (2009)

46. G.Y. Jing, H.L. Duan, X.M. Sun, Z.S. Zhang, J. Xu, Y.D. Li, J.X. Wang, D.P. Yu, Surface effects on elastic properties of silver nanowires: contact atomic-force microscopy. *Phys. Rev. B* 73, 235409 (2006)
47. J.M. Gere, S.P. Timoshenko, *Mechanics of Materials*, 2nd edn. (PWS-KENT, Boston, MA, 1984)
48. V. Sazonova, Y. Yaish, H. Üstünel, D. Roundy, T.A. Arias, P.L. McEuen, A tunable carbon nanotube electromechanical oscillator. *Nature (London)* 431, 284–287 (2004)
49. D. Garcia-Sanchez, A. San Paulo, M.J. Esplandiu, F. Perez-Murano, L. Forró, A. Aguasca, A. Bachtold, Mechanical detection of carbon nanotube resonator vibrations. *Phys. Rev. Lett.* 99, 085501 (2007)
50. D.J. Zeng, Q.S. Zheng, Resonant frequency-based method for measuring the Young's modulus of nanowires. *Phys. Rev. B* 76, 075417 (2007)
51. P.E. Marszalek, W.J. Greenleaf, H. Li, A.F. Oberhauser, J.M. Fernandez, Atomic force microscopy captures quantized plastic deformation in gold nanowires. *Proc. Natl. Acad. Sci. USA* 97, 6282–6286 (2000)
52. J.P. Salvetat, G.A.D. Briggs, J.M. Bonard, R.R. Bacsá, A.J. Kulik, T. Stöckli, N.A. Burnham, L. Forró, Elastic and shear moduli of single-walled carbon nanotube ropes. *Phys. Rev. Lett.* 82, 944–947 (1999)
53. S. Hoffmann, I. Utke, B. Moser, J. Michler, S.H. Christiansen, V. Schmidt, S. Senz, P. Werner, U. Gösele, C. Ballif, Measurement of the bending strength of vapor-liquid-solid grown silicon nanowires. *Nano Lett.* 6, 622–625 (2006)
54. S.X. Mao, M. Zhao, Z.L. Wang, Nanoscale mechanical behavior of individual semiconducting nanobelts. *Appl. Phys. Lett.* 83, 993–995 (2003)
55. W.C. Oliver, G.M. Pharr, An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *J. Mater. Res.* 7, 1564–1583 (1992)
56. M.F. Yu, T. Kowalewski, R.S. Ruoff, Investigation of the radial deformability of individual carbon nanotubes under controlled indentation force. *Phys. Rev. Lett.* 85, 1456–1459 (2000)
57. L.W. Ji, S.J. Young, T.H. Fang, C.H. Liu, Buckling characterization of vertical ZnO nanowires using nanoindentation. *Appl. Phys. Lett.* 90, 033109 (2007)
58. U. Rabe, K. Janser, W. Arnold, Vibrations of free and surface-coupled atomic force microscope cantilevers: Theory and experiment. (1996) *Rev. Sci. Instrum.* 67, 3281–3293 (1996)
59. K. Yamanaka, S. Nakano, Ultrasonic atomic force microscopy with overtone excitation of cantilever. *Jpn. J. Appl. Phys.* 35, 3787–3792 (1996)
60. K. Yamanaka, H. Ogiso, O. Kolosov, Ultrasonic force microscopy for nanometer resolution subsurface imaging. *Appl. Phys. Lett.* 64, 178–180 (1994)
61. M.T. Cuberes, H.E. Assender, G.A.D. Briggs, O.K. Kolosov, Heterodyne force microscopy of PMMA/rubber nanocomposites: nanomapping of viscoelastic response at ultrasonic frequencies. *J. Phys. D: Appl. Phys.* 33, 2347–2355 (2000)
62. T. Drobek, R.W. Stark, W.M. Heckl, Determination of shear stiffness based on thermal noise analysis in atomic force microscopy: Passive overtone microscopy. *Phys. Rev. B* 64, 045401 (2001)
63. O. Sahin, S. Magonov, C. Su, C.F. Quate, O. Solgaard, An atomic force microscope tip designed to measure time-varying nanomechanical forces. *Nature Nanotech.* 2, 507–514 (2007)

64. U. Rabe, Atomic force acoustic microscopy, in *Applied Scanning Probe Methods II*, ed. by B. Bhushan, H. Fuchs (Springer, Berlin Heidelberg New York, 2006), pp. 37–90
65. D.C. Hurley, Contact resonance force microscopy techniques for nanomechanical measurements, in *Applied Scanning Probe Methods XI*, ed. by B. Bhushan, H. Fuchs (Springer, Berlin Heidelberg New York, 2009), pp. 97–138
66. U. Rabe, M. Kopycinska, S. Hirsekorn, J. Muñoz Saldaña, G.A. Schneider, W. Arnold, High-resolution characterization of piezoelectric ceramics by ultrasonic scanning force microscopy techniques. *J. Phys. D: Appl. Phys.* 35, 2621–2635 (2002)
67. M. Kopycinska-Müller, R.H. Geiss, J. Müller, D.C. Hurley, Elastic-property measurements of ultrathin films using atomic force acoustic microscopy. *Nanotechnology* 16:703–709 (2005)
68. D. Passeri, A. Bettucci, M. Germano, M. Rossi, A. Alippi, V. Sessa, A. Fiori, E. Tamburri, M.L. Terranova, Local indentation modulus characterization of diamondlike carbon films by atomic force acoustic microscopy two contact resonance frequencies imaging technique. *Appl. Phys. Lett.* 88, 121910 (2006)
69. D.C. Hurley, M. Kopycinska-Müller, A.B. Kos, R.H. Geiss, Quantitative elastic-property measurements at the nanoscale with atomic force acoustic microscopy. *Adv. Eng. Mater.* 7, 713–718 (2005)
70. M. Prasad, M. Kopycinska, U. Rabe, W. Arnold, Measurement of Young's modulus of clay minerals using atomic force acoustic microscopy. *Geophys. Res. Lett.* 29, 1172 (2002)
71. A. Kumar, U. Rabe, S. Hirsekorn, W. Arnold, Elasticity mapping of precipitates in polycrystalline materials using atomic force acoustic microscopy. *Appl. Phys. Lett.* 92, 183106 (2008)
72. U. Rabe, S. Amelio, M. Kopycinska, S. Hirsekorn, M. Kempf, A.W. Göken, Imaging and measurement of local mechanical properties by atomic force acoustic microscopy. *Surf. Interface Anal.* 33, 65–70 (2002)
73. D.C. Hurley, K. Shen, N.M. Jennett, J.A. Turner, Atomic force acoustic microscopy methods to determine thin-film elastic properties. *J. Appl. Phys.* 94, 2347–2354 (2003)
74. G. Stan, R.F. Cook, Mapping the elastic properties of granular Au films by contact resonance atomic force microscopy. *Nanotechnology* 19, 235701 (2008)
75. K.L. Johnson, *Contact Mechanics* (Cambridge University, Cambridge, 1985)
76. I.N. Sneddon, The relation between load and penetration in the axisymmetric boussinesq problem for a punch of arbitrary profile. *Int. J. Eng. Sci.* 3, 47–57 (1965)
77. J.J. Vlassak, W.D. Nix, Indentation modulus of elastically anisotropic half spaces. *Phil. Mag. A* 67, 1045–1056 (1993)
78. R.D. Mindlin, Compliance of elastic bodies in contact. *J. Appl. Mech.* 16, 259–268 (1949)
79. D.C. Hurley, J.A. Turner, Measurement of Poisson's ratio with contact-resonance atomic force microscopy. *J. Appl. Phys.* 102, 033509 (2007)
80. P.E. Mazeran, J.L. Loubet, Force modulation with a scanning force microscope: an analysis. *Tribol. Lett.* 3, 125–132 (1997)
81. G. Stan, W. Price, Quantitative measurements of indentation moduli by atomic force acoustic microscopy using a dual reference method. *Rev. Sci. Instrum.* 77, 103707 (2006)

82. I. Palaci, S. Fedrigo, H. Brune, C. Klinke, M. Chen, E. Riedo, Radial elasticity of multiwalled carbon nanotubes. *Phys. Rev. Lett.* 94, 175502 (2005)
83. G.M.L. Gladwell, Contact problems in the classical theory of elasticity, Sijthoff & Noordhoff (Alphen aan den Rijn, The Netherlands, 1980)
84. X. Wang, J. Song, Z.L. Wang, Nanowire and nanobelt arrays of zinc oxide from synthesis to properties and to novel devices. *J. Mater. Chem.* 17, 711–720 (2007)
85. R.W. Carpick, D.F. Ogletree, M. Salmeron, Lateral stiffness: A new nanomechanical measurement for the determination of shear strengths with friction force microscopy. *Appl. Phys. Lett.* 70, 1548–1550 (1997)
86. S.O. Kucheyev, J.E. Bradby, J.S. Williams, S.M.V. Jagadish, Mechanical deformation of single-crystal ZnO. *Appl. Phys. Lett.* 80, 956–958 (2002)
87. C.Q. Sun, B.K. Tay, X.T. Zeng, S. Li, T.P. Chen, J. Zhou, H.L. Bai, E.Y. Jiang, Bond-order-bond-length-bond-strength (bond OLS) correlation mechanism for the shape-and-size dependence of a nanosolid. *J.Phys.: Condens. Matter* 14, 7781–7795 (2002)
88. Y. Xia, P. Yang, Y. Sun, Y. Wu, B. Mayers, B. Gates, Y. Yin, F. Kim, H. Yan One-dimensional nanostructures: synthesis, characterization, and applications. *Adv. Mater. (Weinheim, Ger.)* 15, 353–389 (2003)
89. A.A. Kudryavtsev, *The Chemistry and Technology of Selenium and Tellurium* (Collet's Ltd., London, 1974)
90. A.J. Kulkarni, M. Zhou, F.J. Ke, Orientation and size dependence of the elastic properties of zinc oxide nanobelts. *Nanotechnology* 16, 2749–2756 (2005)
91. Y. Wen, Y. Zhang, Z. Zhu, Size-dependent effects on equilibrium stress and strain in nickel nanowires. *Phys. Rev. B* 76, 125423 (2007)
92. H. Gao, C.H. Chiu, J. Lee, Elastic contact versus indentation modeling of multi-layered materials. *Int. J. Solids Struct.* 29, 2471–2492 (1992)
93. M. Lucas, W. Mai, R. Yang, Z.L. Wang, E. Riedo, Aspect ratio dependence of the elastic properties of ZnO nanobelts. *Nano Lett.* 7, 1314–1317 (2007)
94. R.E. Miller, V.B. Shenoy, Size-dependent elastic properties of nanosized structural elements. *Nanotechnology* 11, 139 (2000)
95. R. Dingreville, J.M. Qu, M. Cherkaoui, Surface free energy and its effect on the elastic behavior of nano-sized particles, wires and films, *J. Mech. Phys. Solids* 53, 1827 (2005)
96. J. He, C.M. Lilley, Surface effect on the elastic behavior of static bending nanowires. *Nano Lett.* 8, 1798–1802 (2008)
97. M.T. McDowell, A.M. Leach, K. Gall, On the elastic modulus of metallic nanowires. *Nano Lett.* 8, 3613–3618 (2008)
98. B. Lee, R.E. Rudd, First-principles calculation of mechanical properties of Si <001> nanowires and comparison to nanomechanical theory. *Phys. Rev. B* 75, 195328 (2007)
99. P.S. Branicio, J.P. Rino, Large deformation and amorphization of Ni nanowires under uniaxial strain: A molecular dynamics study. *Phys. Rev. B.* 62, 16950–16955 (2000)
100. P. Villain, P. Beauchamp, K.F. Badawi, P. Goudeau, P.O. Renault Atomistic calculation of size effects on elastic coefficients in nanometer-sized tungsten layers and wires. *Scr. Mater.* 50, 1247–1251 (2004)
101. X. Li, T. Ono, Y. Wang, M. Esashi, Ultrathin single-crystalline-silicon cantilever resonators: Fabrication technology and significant specimen size effect on Young's modulus. *Appl. Phys. Lett.* 83, 3081–3083 (2003)

102. C.Y. Nam, T.D. Jaroenapibal, D.E. Luzzi, S. Evoy, J.E. Fischer, Diameter-dependent electromechanical properties of GaN nanowires. *Nano Lett.* 6, 153–158 (2006)
103. L. Dong, B.J. Nelson, Tutorial-Robotics in the small, Part II: Nanorobotics. *IEEE Robot. Automat. Mag.* 14, 111–121 (2007)
104. L. Dong, A. Subramanian, B.L. Nelson, Carbon nanotubes for nanorobotics. *Nanotoday* 2, 12–20 (2007)
105. M.Y.A. Yousif, P. Lundgren, P. Ghavanini, B.S. Enoksson, CMOS considerations in nanoelectromechanical carbon nanotube-based switches. *Nanotechnology* 19, 285204 (2008)
106. Z.L. Wang, J. Song, Piezoelectric nanogenerators based on zinc oxide nanowire arrays. *Science* 312, 242–246 (2006)
107. X. Wang, J. Song, J. Liu, Z.L. Wang Direct-current nanogenerator driven by ultrasonic waves. *Science* 316, 102–105 (2007)
108. B.A. Buchine, W.L. Hughes, F.L. Degertekin, Z.L. Wang, Bulk acoustic resonator based on piezoelectric ZnO belts. *Nano Lett.* 6, 1155–1159 (2006)
109. R. Yang, Y. Qin, L. Dai, Z.L. Wang, Power generation with laterally packaged piezoelectric fine wires. *Nature Nanotech.* 4, 34–39 (2009)
110. L. Xiao, Z. Chen, C. Feng, L. Liu, Z.Q. Bai, Y. Wang, L. Qian, Y. Zhang, Q. Li, K. Jiang, S. Fan, Flexible, stretchable, transparent carbon nanotube thin film loudspeakers. *Nano Lett.* 8, 4539–4545 (2008)
111. M.F. Ashby, Y.J.M. Brechet, Designing hybrid materials. *Acta Mater.* 51, 5801–5821 (2003)
112. F. Patolsky, B.P. Timko, G. Yu, Y. Fang, A.B. Greytak, G. Zheng, C.M. Lieber, Detection, stimulation, and inhibition of neuronal signal with high-density nanowire transistor arrays. *Science* 313, 1100–1104 (2006)
113. X. Duan, Y. Huang, R. Agarwal, C.M. Lieber, Single-nanowire electrically driven lasers. *Nature* 421, 241–245 (2003)
114. Y. Li, F. Qian, J. Xiang, C.M. Lieber, Nanowire electronic and optoelectronic devices. *Materials Today* 9, 18–27 (2006)
115. B.T. Tian, X. Zheng, T.J. Kempa, Y. Fang, N. Yu, G. Yu, J. Huang, C.M. Lieber, Coaxial silicon nanowires as solar cells and nanoelectronic power sources. *Nature* 449, 885–889 (2007)
116. P.J. Pauzauskie, P. Yang, Nanowire photonics. *Mater. Today* 9, 36–45 (2006)
117. Y. Cui, Q. Wei, H. Park, C.M. Lieber, Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science* 293, 1289–1292 (2001)
118. F. Patolsky, C.M. Lieber, Nanowire nanosensors. *Mater. Today* 8(4), 20–28 (2005)
119. M. Willander, P. Klason, L.L. Yang, S.M. Al-Hilli, Q.X. Zhao, O. Nur, ZnO nanowires: chemical growth, electrodeposition, and application to intracellular nano-sensors. *Phys. Stat. Sol. (c)* 5, 3076–3083 (2008)
120. K.L. Johnson, K. Kendall, A.D. Roberts, Surface energy and the contact of elastic solids. *Proc. R. Soc. London A* 324, 301–313 (1971)
121. B.V. Derjaguin, V.M. Müller, Y.P. Toporov, Effect of contact deformation on the adhesion of particles. *J. Colloid Interf. Sci.* 53, 314–326 (1975)
122. D. Maugis, Adhesion of spheres: the JKR-DMT transition using a Dugdale model. *J. Colloid Interface Sci.* 150, 243–269 (1992)

Colossal Permittivity in Advanced Functional Heterogeneous Materials: The Relevance of the Local Measurements at Submicron Scale

Patrick Fiorenza, Raffaella Lo Nigro, and Vito Raineri

Summary. Recently, giant dielectric permittivities ($\epsilon' \sim 10^4$) have been found in several nonferroelectric materials such as $\text{CaCu}_3\text{Ti}_4\text{O}_{12}$ (CCTO) (Subramanian et al., J. Solid State Chem. 151:323, 2000; Homes et al., Science 293:673, 2001), doped-NiO (Wu et al., Phys. Rev. Lett. 89:217601, 2002) systems ($\text{Li}_x\text{Ti}_y\text{Ni}_{1-x-y}\text{O}$, $\text{Li}_x\text{Si}_y\text{Ni}_{1-x-y}\text{O}$, $\text{K}_x\text{Ti}_y\text{Ni}_{1-x-y}\text{O}$), CuO, (Lin et al., Phys. Rev. B 72:014103, 2005; Sarkar et al., App. Phys. Lett. 92:142901, 2008) etc., and most important, the high ϵ' values of these materials are almost independent over a wide range of temperature. This is one of the most intriguing features for their implementations in microelectronics devices, and as a consequence, these materials have been subjected to extensive research. Here, an introduction to such materials and to the methods for their dielectric characterization is given. So far, the crucial question is whether the large dielectric response is an intrinsic property of new class of crystals or an extrinsic property originated by a combination of the structural properties and other features such as defects and inhomogeneities. Preliminary, this peculiar dielectric behavior has been explained in powder ceramics by the internal barrier layer capacitor (IBLC) model, that is the presence of semiconducting domains surrounded by thin insulating regions within the crystal microstructure. It has been considered the most appropriate model and it has been generally accepted to explain the giant response of these materials. However, it could not be transferred to single crystals and thin films. In this scenario, scanning probe-based methods (like STM, KPFM, C-AFM, SIM etc) represent the most powerful instrument to understand the colossal permittivity-related physical phenomena, by investigations at nanoscale, clarifying the local effects responsible of the rising of macroscopic giant dielectric responses. Scanning probe microscopy investigations showed the relevance of inhomogeneity within single crystal, polycrystalline ceramics, and thin films. In particular, they are powerful tools to point out the presence of few nanometer wide internal barrier layers and of electrical domains, which are not recognisable with standard macroscopic electric characterization techniques.

Key words: Colossal permittivity, Heterogeneous materials, Scanning probe microscopy.

Abbreviations

AFM	Atomic force microscopy
BLCs	Barrier layer capacitance elements
C-AFM	Conductive atomic force microscopy
CCTO	$\text{CaCu}_3\text{Ti}_4\text{O}_{12}$
EF-TEM	Energy filtered transmission electron microscopy
G/Gs	Grain/grains
GB/GBs	Grain boundary/grain boundaries
IBLC	Internal barrier layer capacitor
IS	Impedance spectroscopy
KPFM	Kelvin probe force microscopy
LDOS	Local density of states
MW	Maxwell Wagner
SBLC	Surface barrier layer capacitor
SIM	Scanning impedance microscopy
SPM	Scanning probe microscopy
SPS	Scanning probe spectroscopy
STM	Scanning tunneling microscopy
TEM	Transmission electron microscopy

17.1 Introduction

Microelectronics is driven by an almost insatiable appetite for smaller and faster devices. In memory devices based on capacitive components, such as static and dynamic random access memories, the materials dielectric constant ϵ' will ultimately decide the degree of miniaturization. In other applications, such as hybrid and Rf circuits, the 95% of chip area consists of capacitors. Their capacitive densities are directly related to the permittivities of the used dielectrics. Materials with colossal permittivities ($>1,000$) can increase the capacitive density of several orders of magnitude, thus drastically reducing the capacitor chip area.

Recently, giant permittivity ($\epsilon' \sim 10^4$) have been measured in several non-ferroelectric materials such as $\text{CaCu}_3\text{Ti}_4\text{O}_{12}$ (CCTO), doped NiO systems ($\text{Li}_x\text{Ti}_y\text{Ni}_{1-x-y}\text{O}$, $\text{Li}_x\text{Si}_y\text{Ni}_{1-x-y}\text{O}$, $\text{K}_x\text{Ti}_y\text{Ni}_{1-x-y}\text{O}$), CuO, etc., and most importantly, the high ϵ' values of these materials are almost independent over a wide temperature range. This is one of the most intriguing features for their implementations in microelectronics devices, and as a consequence, at present, these materials have been subject of extensive research activities.

The dielectric constant is related to the material polarizability, in particular the dipole polarizability (an atomic property), originating from structures with a permanent electric dipole, can change orientation in an applied electric field. These two quantities are linked through the Clausius-Mossotti relation. In metals, the charge is delocalized and $\epsilon' \ll 0$, while in insulators, the

charge is localized and $\epsilon' > 0$ (materials with a dielectric constant greater than silicon nitride, $\epsilon' > 7$, are classified as high-dielectric constant materials). In general, ϵ values $> 1,000$ are related either to ferroelectrics that exhibits a dipole moment in the absence of an external electric field or to relaxors characterized by a ferroelectric response under high electric fields at lower temperature, but not showing any macroscopic spontaneous polarization. However, both classes of materials show a peak of ϵ' as a function of temperature, which is undesirable for many applications. For instance, capacitors [with capacitance $C = \epsilon'\epsilon_0 A/d$; two plates of area A separated by a distance d] need to have static values to operate properly under a variety of conditions; if the permittivity has a strong temperature dependence, then the device will not be robust and may fail.

A perovskite is any material with the same type of crystal structure as calcium titanium oxide, CaTiO_3 , known as the *perovskite structure*. Perovskites take their name from the compound that was first discovered in the Ural mountains of Russia by Gustav Rose in 1839 and was after named by a Russian mineralogist, L.A. Perovski (1792–1856). The general chemical formula for perovskite compounds is ABX_3 , where “A” and “B” are two cations of very different sizes, and X is an anion that bonds to both. The “A” atoms are larger than the “B” atoms. The ideal cubic-symmetry structure has the B cation in 6-fold coordination, surrounded by an octahedron of anions, and the A cation in 12-fold cuboctahedral coordination. The relative ion size requirements for the cubic structure stability are quite stringent, so slight buckling and distortion can produce several lower-symmetry distorted versions, in which the coordination numbers of A cations, B cations, or both are reduced.

Perovskite materials exhibit many interesting and intriguing properties from both the theoretical and the applications points of view. Colossal magnetoresistance, ferroelectricity, superconductivity, charge ordering, spin-dependent transport, high thermopower, and the interplay of structural, magnetic, and transport properties are commonly observed features in this materials family. These compounds are used as sensors and catalyst electrodes in certain types of fuel cells and are candidates for memory devices and spintronics applications. The understanding of the phenomena involved in the colossal permittivity response is fundamental for the tuning and the engineering of the macroscopic future device applications.

In particular, for the dielectric response of the mentioned materials, the central question is whether the large dielectric response is intrinsic (a property of the perfect crystal) or extrinsic, that is originated by the capacitor structures or by defects, heterogeneities, and impurities in the crystal. In particular, the perovskite-related body-centered cubic material CCTO revealed an extraordinarily high dielectric constant at room temperature of 10^5 [1, 2] and was found to be practically constant between 100 and 600 K. Both properties are very important for device implementation. Among the vast number of papers on the extremely high (“colossal”) dielectric constants found in CCTO, there are at least ten so-called highly cited papers [1–5]. This is a

demonstration of the tremendous interest in advanced high- ϵ' materials, which are prerequisite for further progress in the development of capacitive electronic elements.

In some cases, the macroscopic dielectric response is due to the atomic self-organization (e.g., intrinsic lamellar conducting and insulating structures), or presence of heterogeneities (e.g., extrinsic conducting and insulating phases separation: different grain conductivity, grain–grain boundaries conductivity, surface effects, etc).

The origin of the colossal permittivity in CCTO, if it is intrinsic or not, is still under investigation. Nowadays, usually the results are interpreted within an “internal barrier layer capacitor” (IBLC) model. Polarization effects at insulating grain boundaries between semiconducting grains or other internal barriers generate nonintrinsic colossal values of permittivity, accompanied by a strong Maxwell-Wagner (MW) relaxation mode. As an alternative, a “surface barrier layer capacitor” (SBLC) model was proposed, assuming, for example, the formation of Schottky diodes at the electrode–dielectric interfaces.

This chapter presents an overview on the SPM-based techniques that provided a fundamental contribution for the understanding of physical phenomena involved in the giant permittivity response of electrically heterogeneous materials and it suggests the possibility to employ such techniques on a wider materials families, playing a fundamental role in some of the hottest topics in the dielectrics community.

17.2 Physical Properties of Heterogeneous Materials

17.2.1 Theory of the Dielectric Relaxation: Basic Principles

The complex dielectric function $\epsilon^*(\omega)$, in its dependence on angular frequency $\omega = 2\pi\nu$ (ν – frequency of the outer applied electrical field) and on temperature, originates from different processes: (1) the microscopic fluctuations of molecular dipoles [6, 7] (rotational diffusion), (2) the propagation of mobile charge carriers (translational diffusion of electrons, holes, or ions), and (3) the separation of charges at the interfaces that gives rise to an additional polarization. The latter can take place at inner dielectric boundary layers (Maxwell/Wagner/Sillars polarization) [8, 9] on a mesoscopic scale and/or at the external electrodes contacting the sample (electrode polarization) on a macroscopic scale. Its contribution to the dielectric loss can be orders of magnitude larger than the dielectric response due to molecular fluctuations.

Each of the above processes has specific features in the frequency and temperature dependence of the real and imaginary parts of the complex dielectric function. This paragraph discusses the physical methods to analyze, separate, and quantify the different contributions to the dielectric response, although sometimes they could act in the same frequency range and they can be confused. In this scenario, the SPM can clarify the physical phenomena involved in the dielectric response.

Dielectric relaxation processes are characterized by a peak in the imaginary part ε'' and a step-like decrease of the real part ε' of the complex dielectric function $\varepsilon^*(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega)$ (being $i = \sqrt{-1}$) with increasing frequency. By contrast, conduction phenomena show an increase of the imaginary part of the dielectric function with decreasing frequency. For pure ohmic conduction, the real part of $\varepsilon^*(\omega)$ is independent of frequency, whereas for nonohmic conduction or polarization effects (at inner boundaries or external electrodes), the real part of $\varepsilon^*(\omega)$ increases with decreasing frequency. A schematic representation of the frequency dependence of $\varepsilon^*(\omega)$ is shown in Fig. 17.1.

Mobile Charge Carrier Contribution

According to Maxwell's equations, the current density $\vec{j} = \sigma^* \vec{E}$ [being $\sigma^*(\omega)$ the complex conductivity] and the derivative time of the dielectric displacement $d\vec{D}/dt = i\omega\varepsilon_0\varepsilon^*(\omega)\vec{E}$ (with ε_0 the vacuum permittivity) are equivalent.

Thus, for a sinusoidal electrical field $\vec{E}(\omega) = E_0 e^{i\omega t}$, the $\varepsilon^*(\omega)$ and $\sigma^*(\omega)$ are related to each other by

$$\sigma^*(\omega) = \sigma'(\omega) + i\sigma''(\omega) = i\omega\varepsilon_0\varepsilon^*(\omega), \quad (17.1)$$

the real and the imaginary parts of $\sigma^*(\omega)$ are then given by:

$$\sigma'(\omega) = i\omega\varepsilon_0\varepsilon''(\omega); \quad \sigma''(\omega) = i\omega\varepsilon_0\varepsilon'(\omega). \quad (17.2)$$

In disordered systems, the charge transport takes place due to hopping conduction [10]. Moreover, the motion of charges in these systems is accompanied by a relaxation: an ionic or electronic (polaron) charge is surrounded by negative or positive countercharges. The hopping of charge carriers to a new site can only lead to a successful charge transport if the polarization cloud flows. Otherwise, the charge carrier will jump back with high probability. This mutual movement of the charge carrier and the surrounding polarization cloud requires an electrical relaxation time τ_σ .

A theoretical calculation for a sample with ohmic conductivity but two different conductivity values ($\sigma_0/\varepsilon_0 = 1$ and $\sigma_0/\varepsilon_0 = 10^4$) is reported in Fig. 17.2 as an example. For pure electronic conduction, no contribution arises for ε' (continuous line in both cases) whereas $\varepsilon''(\omega) = \sigma_0/\varepsilon_0\omega$ (dashed and dashed-dotted lines respectively) increases linearly with decreasing the frequency. σ_0 is the (electronic) dc-conductivity and can be estimated by $\sigma_0 = \varepsilon_0\varepsilon_\infty/\langle\tau\rangle$, where $\langle\tau\rangle$ is a mean relaxation time for the conductivity and ε_∞ is the permittivity at the infinite frequency limit.

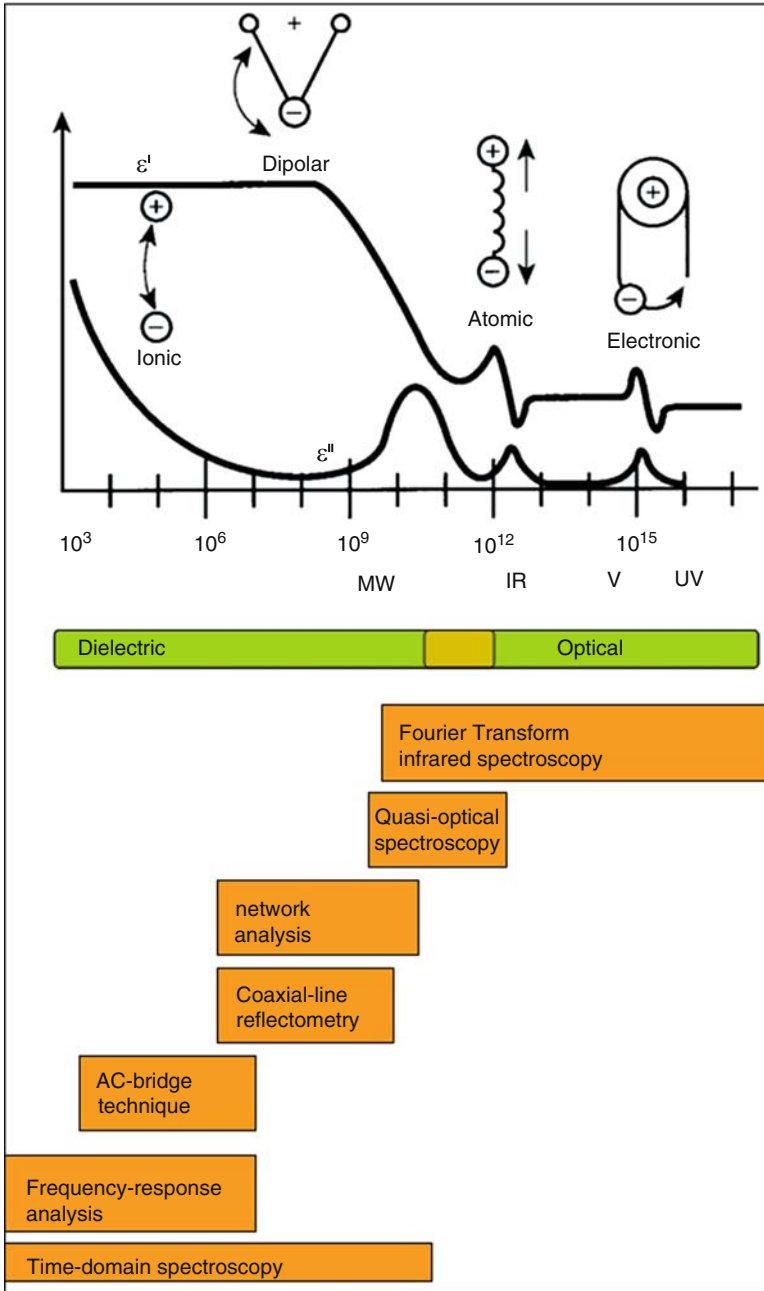


Fig. 17.1. Survey of measurements techniques used in the range frequency from 10^2 to 10^{16} Hz

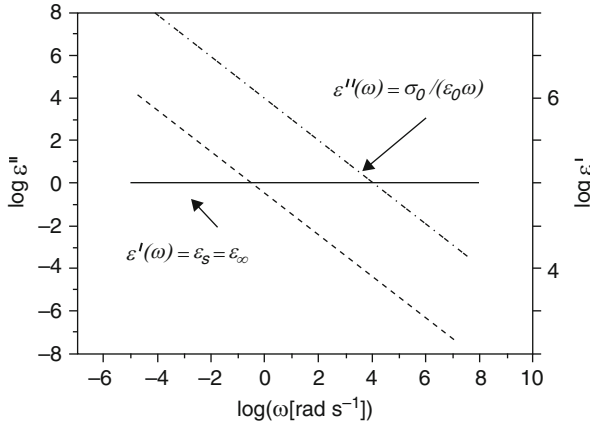


Fig. 17.2. Theoretical example for a complex dielectric function with a pure ohmic contribution $\sigma_0/\epsilon_0 = 1$ (dashed line) and $\sigma_0/\epsilon_0 = 10^4 \epsilon' = 5$

17.2.2 Separation of Charges: Maxwell/Wagner/Sillars Polarization

Charge carriers can be blocked at inner dielectric boundary layers (Maxwell/Wagner/Sillars polarization) [8,9] on a mesoscopic scale or at the external electrodes contacting the sample (electrode polarization) on a macroscopic scale. In both cases, a separation of the charges gives rise to an additional contribution to the polarization. The charges may be separated over a considerable distance.

Mesoscopic Scale: Separation of Charges ⇒ Maxwell/Wagner/Sillars (MW) Polarization

MW polarization processes have to be taken into consideration during the investigation of inhomogeneous materials like suspensions or colloids, biological materials, phase separated materials, and crystalline or liquid crystalline polymers. As an example of the influence of the MW polarization, Fig. 17.3 shows the schematic of an inhomogeneous material and the equivalent circuit. The complex dielectric function can be described as [6, 8, 9]

$$\epsilon^*(\omega) = \epsilon_\infty + \frac{\Delta\epsilon}{1 + i\omega\tau_{MW}}. \tag{17.3}$$

The simplest model to describe an inhomogeneous structure is a double layer arrangement where each layer is characterized by its permittivity ϵ_i and by its relative conductivity σ_i (Fig. 17.3).

Equation (17.3) is similar to the ideal Debye behavior [11] but the parameters have completely different physical meaning. For thicknesses $D_1 = D_2$, it

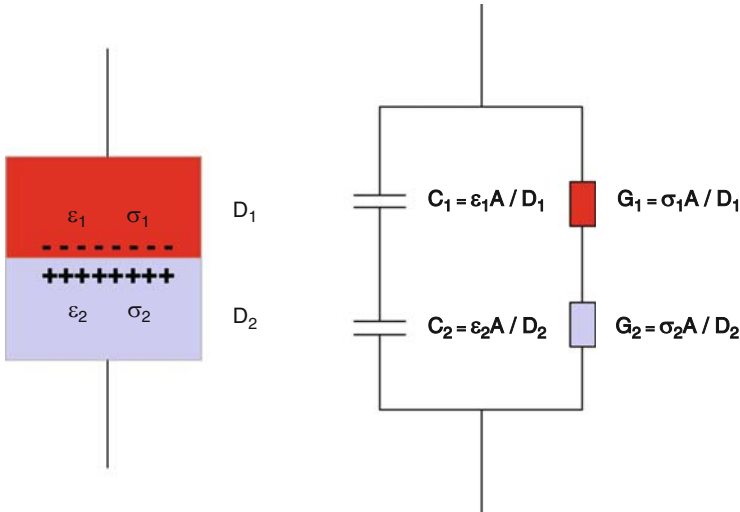


Fig. 17.3. (a) Two different layers in series with ϵ_i and σ_i , the corresponding dielectric permittivities and conductivities. (b) Equivalent circuit where A is the diameter of the electrodes and D_i is the spacing of the layers

holds $\epsilon_\infty = \epsilon_1\epsilon_2/(\epsilon_1 + \epsilon_2)$ and $\Delta\epsilon = \epsilon_2\sigma_1 + \epsilon_1\sigma_2/(\sigma_1 + \sigma_2)^2(\epsilon_1 + \epsilon_2)$. Thus, the relaxation time τ_{MW} of the interfacial polarization is

$$\tau_{MW} = \epsilon_0 \frac{\epsilon_1 + \epsilon_2}{\sigma_1 + \sigma_2}. \tag{17.4}$$

The relaxation time scales inversely with the conductivity of the system. This indicates that the MW effects are more pronounced for conductive materials.

Macroscopic Scale: Electrode Polarization

Electrode polarization is usually an unwanted parasitic effect that can hide the “real” dielectric response of the sample. It occurs mainly for moderately or high conducting samples and influences the dielectric properties at low frequencies. Both the magnitude and the frequency position of the electrode polarization depend on the conductivity of the sample.

At low frequencies, the real part increases strongly with decreasing frequencies up to very high values, which cannot be explained by molecular relaxation processes. Such behavior is a clear sign of electrode polarization. The molecular origin of that effect is the (partial) blocking of charge carriers at the sample–electrode interface. This leads to a separation of positive and negative charges causing an additional polarization.

A theoretical treatment of the electrode polarization in terms of a simple model [12,13], the blocking of charges at the sample–electrode interface, can be described by an electrical double layer with an effective spacing characterized

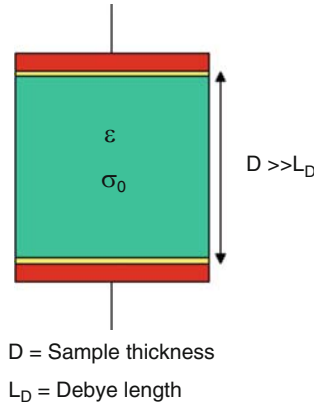


Fig. 17.4. Schematic of a material presenting two depleted layers at the interfaces with the electrodes having thickness D and L_D as Debye length

by its Debye length (L_D). This double layer causes a large capacitance in series to the investigated sample. The time dependence of the electrode polarization is due to charging and discharging of the double layer. The model given in Fig. 17.4 can be used to estimate the effect of the electrode polarization. From (17.4), the characteristic time constant of the present process

$$\tau_{EP} = \frac{\varepsilon_s \varepsilon_0}{\sigma_0} \frac{D}{2L_D} \tag{17.5}$$

can be estimated if the thickness $D \gg L_D$ where σ_0 is the dc-conductivity for the system.

The double-layer model for electrode polarization discussed above leads to a Debye-like frequency dependence [(17.3)]. Because of the fractal nature of the electrode surfaces, only rarely a Debye-like frequency dependence of the electrode polarization is found. Often the fractal power laws for the frequency dependence for the complex dielectric function are observed. Fractal power laws also follow from different theoretical treatment of electrode polarization [14]. Therefore, to correct the measurements for electrode polarization, it is possible to assume

$$\varepsilon'_{EP}(\omega) - \varepsilon_s = A\omega^{-\lambda}; \quad \varepsilon''_{EP}(\omega) \approx \omega^{-\lambda} \quad \omega > \frac{1}{\tau_{EP}},$$

where $\lambda(0 < \lambda \leq 1)$ is a parameter describing the fractal character of the underlying process and ε_s is the permittivity which is due to orientational polarization.

17.2.3 Ultimate Theories on the Dielectric Relaxation

In the last years, several giant permittivity ($\varepsilon' > 10^4$) materials [15] have attracted the scientific attention due to the colossal permittivity response

that could be not easily explained. So far, it was suggested that this behavior was either an intrinsic property associated with some kind of ferroelectric relaxation involving displacement of the internal ions or an extrinsic effect associated with electrical heterogeneities, such as insulating grain boundary layers between semiconducting grains, as found in SrTiO₃-based IBLCs [16] and ZnO-based varistors [17]. In the latter case, the “effective” permittivity from an IBLC-type device can be estimated as $\varepsilon_{\text{eff}} \approx \varepsilon_r(t_b/t_{\text{gb}})$ where ε_r is the relative permittivity of the insulating grain boundary phase, t_b is the average grain size, and t_{gb} is the average thickness of the grain boundaries.

Sinclair et al. [18] used impedance spectroscopy (IS) to demonstrate that some ceramics were electrically heterogeneous, consisting of semiconducting grains (with permittivity < 100) and insulating grain boundaries. The data were analyzed using a simplified equivalent circuit consisting of two parallel RC elements connected in series, one RC element, $R_b C_b$, representing the semiconducting grains and the other, $R_{\text{gb}} C_{\text{gb}}$, representing the insulating grain boundary regions (Fig. 17.5a). A schematic diagram of a typical impedance complex plot, Z^* , obtained at room temperature for electrically heterogeneous ceramics is shown in Fig. 17.5b.

On the basis of this equivalent circuit, and in agreement with the brickwork layer model for electroceramics, the large arc observed at lower frequencies was attributed to the grain boundary response and the nonzero intercept on the Z' axis at high frequencies was attributed to the bulk resistance, R_b .

These studies supported the “extrinsic” IBLC model as the origin of the giant permittivity in electroceramics and also reconciled the “intrin-

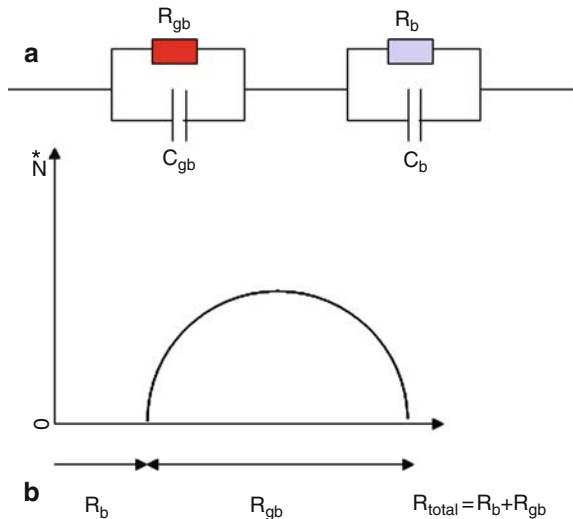


Fig. 17.5. Equivalent circuit (a) and schematic Z^* plot (b) for heterogeneous ceramics

sic” properties of such materials, that is, a permittivity < 100 , which is consistent with a tilted, centrosymmetric perovskite with only electronic and ionic polarizability mechanisms contributing to the permittivity. Adams et al. [3] subsequently used IS to show the influence of average grain size on the electrical properties of CCTO ceramics, reporting extremely high grain boundary capacitance ($\sim 10^{-8}$ F cm $^{-1}$) and effective permittivity values ($\sim 280,000$ at 300 K) for large grain ($\sim 100\text{--}300\ \mu\text{m}$) ceramics, providing further support for the IBLC mechanism in CCTO ceramics. This IBLC model has been supported by other workers [19].

Recently, Chung et al. [20] used a combination of scanning Kelvin probe force microscopy (KPFM) with a lateral bias and I – V measurements on pellets and across individual grain boundaries (using microcontact electrodes on individual grains) to demonstrate the presence of electrostatic barriers at the grain boundaries in CCTO ceramics. The I – V characteristics of the ceramics and across single grain boundaries were nonlinear and consistent with that expected for Schottky-type barriers.

Recently, however, Lunkenheimer et al. [4] claimed that the high permittivity is associated with electrode polarization effects and the development of Schottky barriers at the metal electrode–ceramic interface due to nonohmic contacts and is not due to a grain boundary effect. Such nonohmic electrode contacts are well known to occur with many different types of materials, for example, in positive temperature coefficient of resistance (ptcr)-BaTiO $_3$ ceramics with Au electrodes, but can be eliminated by using an electrode with a lower electron work function such as In–Ga alloy.

The Presence of Inner Schottky Barriers

Schottky barriers are created in many commercial electroceramics such as ptcr-BaTiO $_3$ thermistors and ZnO varistors by deliberate introduction of compositional heterogeneity between the grain and grain boundary regions, so that the grain boundaries act as an extrinsic source of impedance. In many cases, double (back-to-back) Schottky potential barriers are created at interfaces between n -type grains, due to charge trapping at acceptor states, resulting in bending of the conduction band across the grain boundary. This band bending produces an effective potential barrier of height, Φ_b , and depletion width, w , for conduction electrons between the grain and grain boundary regions (Fig. 17.6a). The barrier can typically be characterized as a grain boundary impedance since it is both resistive (in comparison to the bulk) and of high capacitance, due to the thin width of the depletion region with respect to the grain size.

In the absence of a dc bias, Φ_b is given by (17.6)

$$\Phi_b = \frac{qN_s^2}{8\varepsilon_0\varepsilon'_d N_d}, \quad (17.6)$$

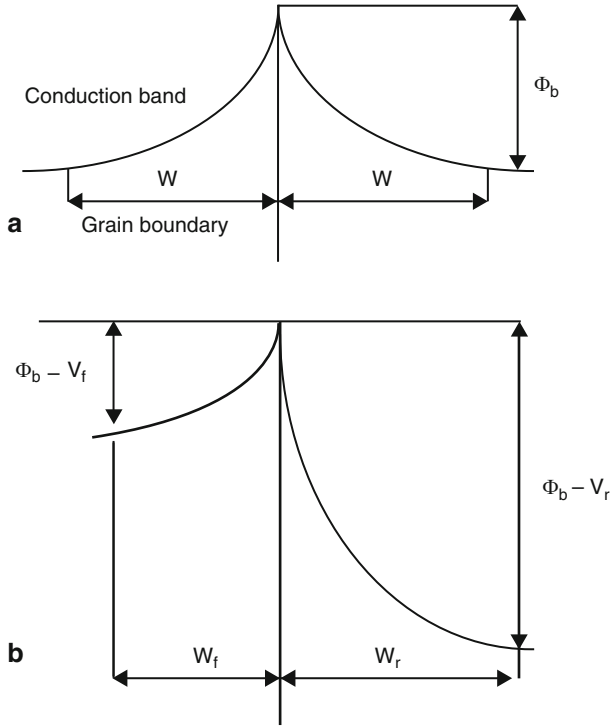


Fig. 17.6. Illustration of a back-to-back double Schottky barrier at a single grain boundary when (a) $V = 0$ and (b) $V > 0$

where q is the electronic charge, N_s is the acceptor (surface charge) concentration, $\epsilon_0 = 8.854 \times 10^{-14} \text{ F cm}^{-1}$, ϵ' is the material permittivity, and N_d is the charge carrier concentration in the grains [21]. The depletion width, on either side of the interface, is identical and proportional to barrier height, Φ_b [(17.7)].

$$w = \left(\frac{2\Phi_b \epsilon'}{qN_d} \right)^{1/2}. \tag{17.7}$$

This situation is shown schematically in Fig. 17.6a. The grain boundary capacitance per unit area, in the absence of a dc bias, is given by

$$C_0 = \left(\frac{\epsilon' q N_d}{8\Phi_b} \right)^{1/2}. \tag{17.8}$$

Under a dc bias, V , the barrier becomes asymmetric such that the depletion width in the reverse direction, w_r , increases, whereas the depletion width in the forward direction, w_f , decreases according to (17.9) and (17.10), respectively (Fig. 17.6b)

$$w_r = \left(\frac{\varepsilon'}{qN_d} \right)^{1/2} \left(\sqrt{4\Phi_b^2 + V^2} - V \right)^{1/2}, \quad (17.9)$$

$$w_f = \left(\frac{\varepsilon'}{qN_d} \right)^{1/2} \left(\sqrt{4\Phi_b^2 + V^2} + V \right)^{1/2}. \quad (17.10)$$

Mukae et al. (13) have shown that the dependence of the grain boundary capacitance associated with the depletion region under a dc bias can be described by

$$\left(\frac{1}{C} - \frac{1}{2C_0} \right)^2 = \frac{2}{\varepsilon'qN_d} (\Phi_b + V), \quad (17.11)$$

where C is the biased grain boundary capacitance per unit area.

The presence of those barriers produces a depleted layer [as it has been discussed, see (17.7)], usually thinner than the semiconducting surrounding grains. This is a particular example, already explained in Sect. 17.3.1, and it is called IBLC effect.

Intrinsic and Extrinsic Mechanisms

The central question is now whether the large dielectric response is intrinsic to a perfect crystal of electroceramics or extrinsic, and if it originates with defects, inhomogeneities, etc., in specific samples. On the basis of first-principles calculations of the materials structures, sometimes it is not possible to explain the colossal permittivity response.

Empirically, for example, the single-crystal perovskite samples are known to be highly twinned [1] and the transport behavior of these domains and their boundaries, as well as those of other such domains and boundaries (in both single-crystal and ceramic samples), could play a significant role in the observed low-frequency giant dielectric response, as was noted in earlier works [5]. The basic idea is that the bulk of the material is either conducting or nearly so, and that the conductivity of the entire sample is only prevented either by a failure of the conducting regions to percolate or by the presence of thin insulating blocking layers at the surfaces or at internal domain boundaries. The various morphologies associated with these possibilities can all be consistent with a huge enhancement of the static dielectric constant and with a Debye-like frequency response, although there are characteristic details of the dielectric response that may help distinguish among them.

Large quasi-static dielectric constants ε' can arise in macroscopic insulators containing conducting regions approaching a percolation threshold. In principle, this may occur for a wide range of morphologies; however, in order to rationalize all the possibilities, here are summarized the six relevant morphology models for single-crystal perovskites possessing the largest values of ε' [2]. Each morphology can be decomposed into two or three broadly defined regions: (1) domains, (2) their associated boundaries, and (3) blocking layers (if any) parallel to the electrodes. Several domain types (and, therefore,

Table 17.1.

Morphology	Acronym
• Nonpercolating, conducting bulk, unblocked	NP/CB/UB
• Nonpercolating, conducting bulk, surface blocked	NP/CB/UB
• Nonpercolating, conducting interfaces, unblocked	NP/CB/UB
• Nonpercolating, conducting interfaces, surface blocked	NP/CB/UB
• Percolating, conducting bulk, surface blocked	NP/CB/UB
• Percolating, conducting interfaces, surface blocked	NP/CB/UB

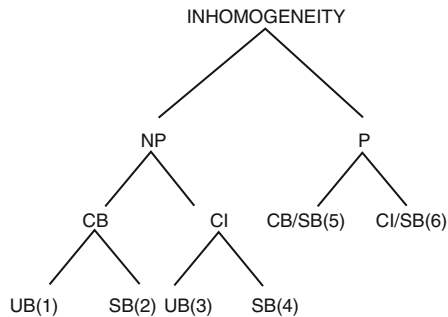


Fig. 17.7. Morphological decision tree showing the possible morphological elements that lead to six distinct morphologies: nonpercolating (*NP*), percolating (*P*), conducting bulk (*CB*), conducting interfaces (*CI*), unblocked (*UB*), or surface blocked (*SB*). The numbers in parentheses correspond to those in Table 17.1

boundaries) are possible. In addition to twin boundaries, a necessary by-product of the observed twinning [22] two other categories may exist, both associated with variant chemical ordering.

These possibilities can be represented by a morphological decision tree, illustrated in Fig. 17.7. Six distinct morphologies result, which are listed in the Table 17.1.

The question now arises whether the large dielectric response observed in electroceramics is due only to an intrinsic or extrinsic contribution (or a contribution of them) or whether there is a contribution from micro- and macroscopic defects, inhomogeneities, etc [23–25] (on a subgranular scale in the case of polycrystalline ceramics), which are also present in single crystals. In this context, it is important to study and compare the dielectric properties of single grains in addition to those of the grain boundaries. However, to date, such data have not been produced because of the

lack of characterization methods with a high lateral resolution, to gain a more complete understanding of electrical phenomena associated with inter- and intragranular microstructure and to analyze their relative contributions toward the dielectric behavior.

Observations with nanoprobe sensitive to the conductivity, structure, and composition at the nanoscale would be of significant experimental interest. A locally increased concentration of electron carrier vacancies leading to activated *n*-type (or *p*-type) conduction in the boundaries can be detected. However, it should be recognized that the conducting boundary morphologies require fine tuning of the boundary volume fraction in order to remain close but not exceed the percolation threshold. It is important to recognize that the mechanisms leading to large dielectric constants can differ in polycrystalline (ceramic) and single-crystal samples.

17.3 Conventional Macroscopic Techniques

17.3.1 Basic Principles

The complex dielectric function $\varepsilon^*(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega)$ can be measured in the extraordinary broad frequency regime [26] from 10^2 up to 10^{12} Hz. To span this dynamic range, different measurement systems based on different measurement principles have to be combined, as summarized in Fig. 17.1.

From 10^2 to 10^7 Hz, lumped circuit methods are used in which the sample is treated as a parallel or a serial circuit of an ideal capacitor and an ohmic resistor. Effects of the spatial extent of the sample on the electric field distribution are neglected. On increasing frequency, the geometrical dimensions of the sample capacitor become more and more important, limiting this approach to about 10 MHz. In addition, parasitic impedances caused by cables, connectors, etc become important at frequencies > 100 kHz.

Using distributed circuit methods, the complex dielectric function is deduced by measuring the complex propagation factor (in reflection or transmission).

In the dynamic range between 10^2 and 10^{10} Hz, the complex dielectric function $\varepsilon^*(\omega)$ can also be deduced from the measurement of the time-dependent dielectric function $\varepsilon(t)$. The latter is related to $\varepsilon^*(\omega)$ via a Fourier transformation

$$\varepsilon^*(\omega) - \varepsilon_\infty = \int_{-\infty}^{+\infty} \varepsilon(t) e^{-i\omega t} dt \quad (17.12)$$

with $\varepsilon_\infty = \varepsilon'(\nu \approx 10^{11} \text{ Hz})$.

The main experimental difficulty in this frequency range arises from the large coherence length of the mm- and sub-mm-wave source resulting in multiple standing wave patterns.

17.3.2 Dielectric Spectroscopy

ac impedance methods are widely used to characterize electrical materials. Data may be analyzed in terms of four possible complex formalisms, the impedance Z^* , the electric modulus M^* , the admittance A^* (or Y^*), and the permittivity ε^* . These are interrelated [27]:

$$\begin{aligned} M^* &= i\omega C_0 Z^*, \\ \varepsilon^* &= (M^*)^{-1}, \\ A^* &= (Z^*)^{-1}, \\ A^* &= i\omega C_0 \varepsilon^*. \end{aligned} \tag{17.13}$$

In order to analyze and interpret experimental data, it is essential to have a model equivalent circuit that provides a realistic representation of the electrical properties. This is chosen on (1) intuition on the kind of impedances that are expected to be present in the sample and whether they are connected in series or in parallel, (2) examination of the experimental data to see whether the response is consistent with the proposed circuit, and (3) inspection of the resistance and capacitance values that are obtained in order to check if they are realistic and if their temperature dependence, if any, is reasonable.

In practice, it is usually possible to find more than one equivalent circuit numerically fitting a given data set, but only one of these is likely to provide a realistic representation of the electrical makeup of the sample. In the present materials, it is clear that both microscopic and submicroscopic features impedance contributions are present in the electroceramic and the electrical properties are determined by several combinations of such submicroscopic features impedances. Each of these components may be represented by a parallel RC element, and the simplest appropriate circuit is a series array of parallel RC elements (Fig. 17.5a).

Considering the series circuit given in Fig. 17.5a, it is desired to separate each of the parallel RC elements and measure their component R and C values. This is best achieved using a combination of the impedance and modulus formalisms since each parallel RC element gives rise to a semicircle in the complex plane (Z'' vs Z' ; M'' vs M') or a Debye peak in the spectroscopic plots of the imaginary components. This is seen from the equation for the impedance of this circuit

$$Z^* = (R_g^{-1} + i\omega C_g)^{-1} + (R_{gb}^{-1} + i\omega C_{gb})^{-1} = \text{Re}(Z^*) - i\text{Im}(Z^*) = Z' - iZ'', \tag{17.14}$$

where

$$Z' = \text{Re}(Z^*) = \frac{R_g}{1 + (\omega R_g C_g)^2} + \frac{R_{gb}}{1 + (\omega R_{gb} C_{gb})^2} \tag{17.15}$$

and

$$Z'' = \text{Im}(Z^*) = R_g \left[\frac{\omega R_g C_g}{1 + (\omega R_g C_g)^2} \right] + R_{gb} \left[\frac{\omega R_{gb} C_{gb}}{1 + (\omega R_{gb} C_{gb})^2} \right]. \quad (17.16)$$

The response of this circuit in the complex plane is shown in Fig. 17.5b.

In this scenario scanning probe microscopy (SPM) allows direct imaging of the ceramic microstructure and can be used to identify the Schottky barrier(s) involved in the IBCL model or other MW-related mechanisms. In particular, it has been demonstrated that the IBLC model can be safely applied to several ceramics having relatively small ($\sim 5 \mu\text{m}$) grains, and the SBLC model can be applied for some single crystal samples [28]. A full understanding might be reached by looking at the nanoscopic electrical properties by SPM techniques.

17.4 Scanning Probe Microscopy

The present paragraph demonstrates and reports on the capabilities of SPM-based techniques to study and compare the electrical and dielectric characteristics within single grains and from the grain boundaries. In particular, an original use of scanning capacitance microscopy, scanning impedance microscopy (SIM) allows the direct imaging of the ceramic microstructure providing information (with high lateral resolution) on the local conduction and on local insulating properties. In fact, applying a bias between the bottom electrode and the tip, the conductive atomic force microscopy (C-AFM) tip acts, on the sample surface, as a sliding metal contact, thus it is possible to collect the impedance signal of the nano-device consisting, in the present case, of tip/electroceramics sample/electrode. Such measurements have been of fundamental importance to get insights on the application of the IBLC model, that is the presence of semiconducting grains (Gs) partitioned from each other by thin insulating grain boundaries (GBs) within the ceramic microstructure [29]. A number of studies have already been performed on CCTO [30, 31] and doped-NiO systems [32] to demonstrate the relationship of the G and GB regions, which finally helps to elucidate the origin of semiconducting behavior of the Gs and insulating nature of the GBs of these giant- κ ceramics (κ or ϵ' are equivalent and interchangeable in the later part of this chapter).

17.4.1 Scanning Tunnelling Microscopy on Giant- κ Materials

In case of CuO [33], IS analysis distinguishably proves a much more conducting nature of the Gs with respect to the GBs, and this is certainly a primary criterion for becoming an IBLC. A small difference between the activation energies of CuO Gs and GBs has also been observed, suggesting that geometry as well as electronic structure of the GBs are slightly different from those of Gs. However, IS technique itself is unable to provide information on the composition or electronic structure of the Gs and GBs of CuO ceramic. So

far, scanning tunnelling microscopy (STM) has been employed to interpret the electronic structure of Gs and GBs of polycrystalline CuO samples. This study provides an understanding of the semiconducting character of the Gs and insulating nature of the GBs, and finally elucidates on the origin of giant permittivity response in CuO ceramics.

The presence of very small percentage of Cu^{3+} in the annealed CuO sample, as reported earlier, [32] was described as the reason for its huge dielectric constant $\kappa(\sim 10^4)$. Electron diffraction pattern obtained from transmission electron microscopy (TEM) studies (not shown here) distinctly proves the existence of superlattice structure, arising from charge ordering in CuO, similar to that observed earlier by Zheng et al. [34] in case of CuO single crystal. The superlattice reflections indeed point out to the existence of minute amount of Cu^{3+} in bulk CuO.

The image of CuO observed from STM study, showed in Fig. 17.8, represents geometrical aspect as well as the electronic structure of the surface. Some depressed surface morphological features in CuO (Fig. 17.8) correspond to the regions of comparatively higher resistivity. These regions are expected to be the grain boundaries, as recognized from the image contrast arising from the variation of local electronic nature of the material surface.

Moreover, to confirm whether a region is electronically homogeneous or not (i.e., whether the region is locally conducting, semiconducting, or insulating), STM studies were performed on the Gs and GBs shown in Fig. 17.8 (marked by the *X* and *Y* points, respectively) and the corresponding results have been displayed in Fig. 17.9a–c. The tunneling *I*–*V* characteristics on the G (the point *X* in Fig. 17.8) represent semiconducting behavior of the material as seen from Fig. 17.9a. The magnitude of the observed tunneling current for maximum applied STM tip bias voltage (~ 1.25 V) was about 100 nA. The

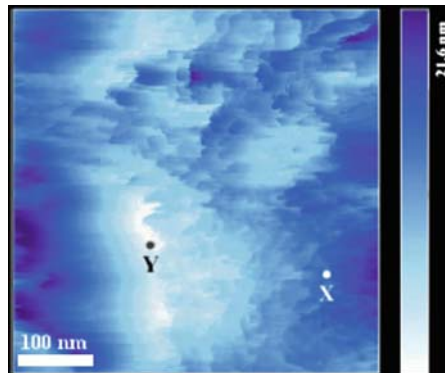


Fig. 17.8. Scanning tunneling microscopy (STM) image of CuO sample showing grain boundary between the CuO Gs. *X* and *Y* are the two points on grain and grain boundary, respectively, where STS studies were performed

tunneling I can be written as [35]

$$I \propto \int_0^{eV} \rho(E)T(E, eV)dE,$$

where $\rho(E)$ is the local density of states (LDOS) of the sample surface and $T(E, eV)$ is the transmission probability of the electron determined by several conditions such as the probe-sample separation and the height of the sample surface potential barrier. To obtain the LDOS profile of CuO grain, differential conductivity (dI/dV) has been calculated, because it is directly related to the LDOS [36]. However, the normalized first derivative of tunneling current $[(dI/dV)/(I/V)]$ is considered to be a relatively direct measure of the LDOS at the Fermi level (E_F) at the position of the STM tip rather than the direct derivative dI/dV , partly because the resultant normalized conductivity provides a measure of the state density, which is approximately independent of the tip-sample separation [37].

However, near the band edges, I/V normally approaches zero faster than dI/dV so that their ratio tends to diverge at a band edge [38]. To overcome this problem, $[(dI/dV)/\overline{(I/V)}]$ instead of $[(dI/dV)/(I/V)]$ has been calculated, considering the approach suggested by Mårtensson and Feenstra [39]. The value of I/V was numerically calculated by adding a small constant to the I/V values. The $[(dI/dV)/\overline{(I/V)}] - V$ plot shown in Fig. 17.9b exhibits the band profile at a point X of a typical CuO grain (Fig. 17.8). The valance band (E_V) and conduction band (E_C) positions are clearly distinguishable from LDOS spectra of the CuO grain as shown in Fig. 17.9b. The LDOS profile on the grain reveals that the band gap of CuO is 1.32 eV. However, the band gap of CuO sample was earlier reported to be ~ 1.4 eV [40]. The zero of voltage, as shown in Fig. 17.9b, corresponds to the E_F position in the CuO grain. The position of E_F on the LDOS profile confirms that the material is p -type [34]. Interestingly, the $I-V$ measurements on the GB (the point Y in Fig. 17.8), shown in Fig. 17.9c, exhibit rectifying nature. The magnitude of the observed current at position Y , for a maximum applied bias voltage (~ 1.25 V) on the STM tip, was about 5 nA. It is worth mentioning that the $I-V$ measurements were also carried out at different spots in the G and GB regions (two such spots X and Y are shown in Fig. 17.8), showing similar results as represented in Fig. 17.9.

The p -type behavior of CuO grains as represented by the $I-V$ characteristics and the corresponding LDOS profile at CuO grains can be attributed to the presence of microquantity of Cu^{3+} as mentioned earlier, whereas the rectifying nature of the $I-V$ characteristics on GBs gives an indication of the Schottky-type barrier formation at GBs. It is also possible to notice the formation of Schottky-type barrier in other giant- κ ceramic such as CCTO [41]. Here, the STM measurements directly exhibit that Cu^{3+} are mostly accumulated within the Gs, which are semiconducting, whereas the concentration of Cu^{3+} is negligible in GBs, which become insulating. This type of electrically heterogeneous nature of the CuO ceramics is, in turn, responsible for the

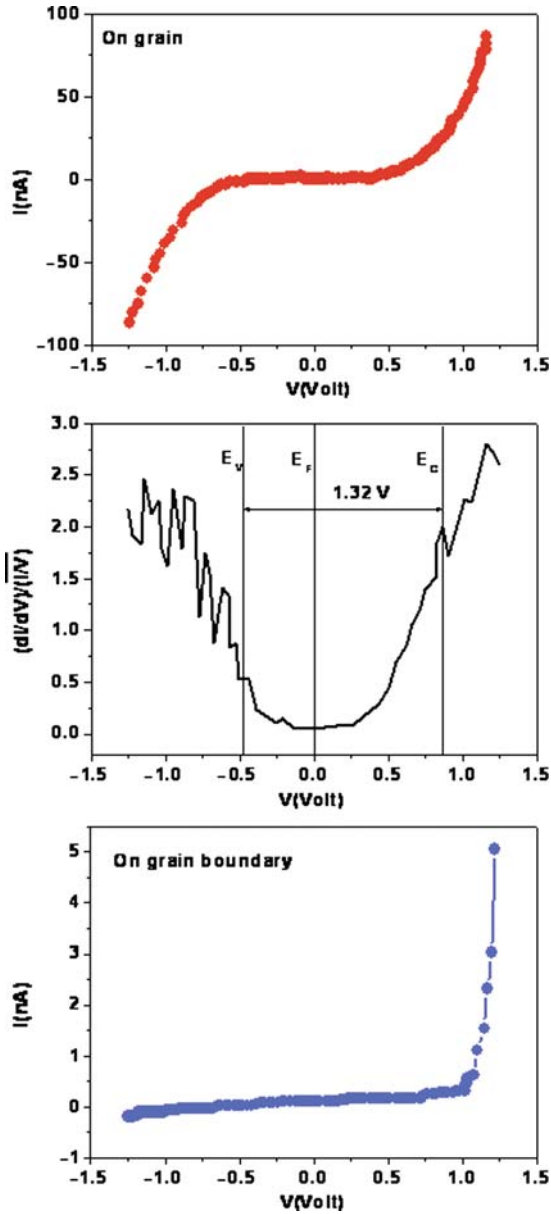


Fig. 17.9. (a) Tunneling spectroscopic I - V characteristics on the CuO grain at point X (Fig. 17.8) and (b) the $[(dI/dV)/(I/V)] - V$ curve at point X , obtained from the I - V data at that point. E_V , E_F , and E_C represent the positions of valance band, Fermi level, and conduction band, respectively. (c) Tunneling spectroscopic I - V behavior on the grain boundary of CuO sample at point Y of Fig. 17.8

formation of barrier layer capacitance elements (BLCEs) in polycrystalline CuO. Assembly of series–parallel array of BLCEs is exclusively formed and the system as a whole acts as an IBLC and, hence, giant dielectric response is exhibited by CuO. At the same time, it is evident that the conduction mechanism in bulk CuO is mainly governed by the hole hopping. It should, however, be mentioned that further extensive work is needed to elucidate the still unknown defect chemistry of CuO enlightening the chemical compositions of the Gs and GBs of CuO ceramics.

17.4.2 Kelvin Probe Force Microscopy on Giant- κ Materials

Many $ATiO_3$ -type perovskite titanates (where $A = \text{Ca, Sr, Ba, Pb}$, or their solid solutions) are key materials as critical components in electronic devices because of their outstanding electrical properties [42]. On the contrary, the related CCTO (or, equivalently, $\text{Ca}_{1/4}\text{Cu}_{3/4}\text{TiO}_3$) showing a giant dielectric response is a fairly recent discovery [2]. The crystal structure of this titanate was identified in 1979 as a perovskite-based body-centered cubic with slightly tilted $[\text{TiO}_6]$ octahedral facing each other [43]. Compared with the dielectric constants of most ferroelectrics and relaxor materials, which usually range from 1,000 to 50,000 at room temperature, the constant of more than 10^5 observed in CCTO, which does not show a ferroelectric transition or relaxor behavior, is extraordinarily high. A number of theoretical studies [5, 22, 25, 44] and experimental observations [45] have attempted to elucidate this remarkable dielectric property. Several possible explanations [25], including an IBLC [3], have been proposed to describe the giant dielectric phenomenon.

By I - V measurements and KPFM, it has been shown that a large potential barrier intrinsically exists at the grain boundaries in addition to the intriguing dielectric response. This barrier acts as an obstacle to the current flow through the conductive bulk of the grains in a polycrystalline specimen, and thus results in enormous nonlinearity between the current and the applied voltage.

To get access to the presence of a potential barrier at the grain boundaries, patterned microelectrodes on a polished surface have been fabricated and the I - V characteristics across individual grain boundaries have been directly measured, as shown in Fig. 17.10a. Figure 17.10b shows the measured I - V curves across grain boundaries and those within a grain for comparison. In contrast to the nearly ohmic behavior between electrodes 1 and 2 within a grain (blue curve), the I - V curves measured across grain boundaries are highly nonlinear, for any direction of the bias voltage. For single grain boundaries between electrodes 1 and 3 (black curve), and also 1 and 4 (red curve), the flowing current is suppressed up to a threshold voltage of about 2.2 V, showing a strong nonlinearity.

The threshold voltage almost doubles when measured across two individual grain boundaries, between electrodes 1 and 5 (green curve). From these observations, it is possible to conclude that back-to-back electrostatic barriers form at the grain boundaries in polycrystalline CCTO and they are consequently

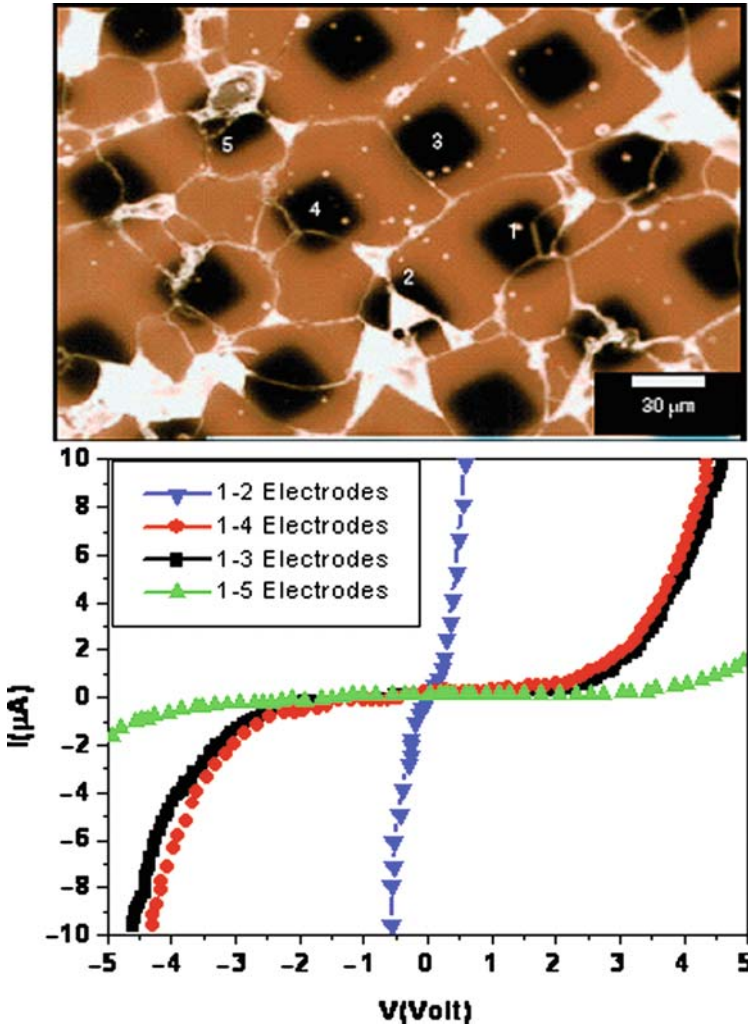


Fig. 17.10. Surface morphology with patterned gold electrodes, and I - V characteristics within single grains and across individual grain boundaries. (a) Electrodes ($30\ \mu\text{m} \times 30\ \mu\text{m}$) deposited on the polished surface. They are numbered for reference. (b) Variations in I - V between pairs of electrodes

responsible for the nonlinear characteristics. The presence of electrostatic barriers at grain boundaries in CCTO has been confirmed by scanning KPFM [46] under a lateral bias, which can visualize the surface potential difference between grains caused by the electrostatic barrier. This type of microscopy is a variant of atomic force microscopy (AFM) using a conductive cantilever tip. The regions containing grain boundaries were easily identifiable from the configuration of pores in the sintered microstructure, even though the grain boundaries could not be precisely located during the scanning of the polished

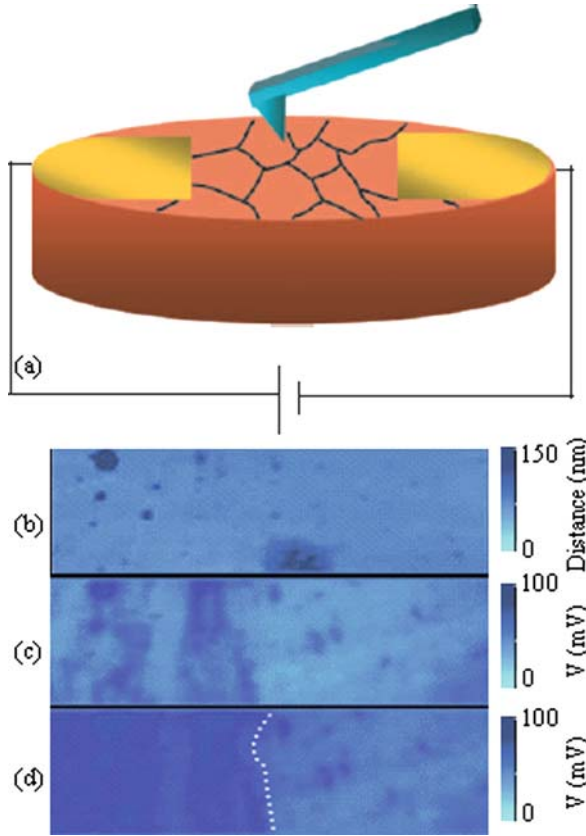


Fig. 17.11. Surface topography and potential variations under a lateral bias (two- and three-dimensional). (a) Schematic diagram of a specimen configuration for the applied lateral bias. The interval between gold electrodes on the surface is 4 mm. (b) atomic force microscopy (AFM) surface topographic image. The scanned region is $30\ \mu\text{m} \times 10\ \mu\text{m}$. Impurity particles adsorbed on the surface are visible. (c) Surface potential image of the scanned region for the topography in (b). (d) Surface potential image with an applied lateral bias of 40 V. The potential drop is clearly observed across the boundary indicated by a broken line in the two-dimensional image. Note also that the overall potential is higher than that in (c) because of the applied bias

surface. Figure 17.11a shows a schematic of the application of a lateral bias on the surface. A surface potential image without lateral bias (Fig. 17.11c) shows no notable variation in surface potential, although some potential fluctuations occur because of the foreign particles adsorbed on the surface, as shown in the topographic image of Fig. 17.11b. By contrast, a potential drop of about 50 mV is clearly observed across the broken line indicated in Fig. 17.11d, in the same region under a lateral voltage of 40 V. This abrupt drop in potential across the boundary confirms the existence of an electrostatic barrier at the grain boundary [47].

The resistivity of single grains can be estimated from the surface current that flows between two microelectrodes under various voltages, using the equation $\rho = 2\pi s(\Delta V/\Delta I)$, where s is the spacing between electrodes and $\Delta V/\Delta I$ is the resistance equivalent to the reciprocal of the slope in an I - V plot [48]. When calculated from the linear relationship between the current and voltage as in the inset of Fig. 17.10b, the average resistivity of five individual grains was $310 \Omega \text{ cm}$, or an electrical conductivity of $3 \times 10^{-3} \text{ cm}^{-1}$ at room temperature. This conductivity is consistent with previous data obtained through IS [18]. Compared with other typical dielectrics, CCTO grains have high conductivity, indicating that there are sufficient mobile charge carriers in the lattice.

17.4.3 SIM on Giant- κ Materials

CCTO Polycrystalline Ceramics

It is not clear if the giant permittivity is an intrinsic property of the materials or if it can be related to extrinsic effects, for example, point and extended defects, contaminants, electrical domain boundaries within grains, grain boundaries, surface layers, and nonohmic electrode contacts. All of these extrinsic effects have been proposed to lead to higher than expected permittivity values in different oxide-based ceramics [49–51]. IS has revealed CCTO ceramics to be electrically heterogeneous and to consist of semiconducting grains and insulating grain boundaries [3]. In this case, the giant permittivity effect in CCTO ceramics has been explained using the already described IBLC model. However, the IBLC results do not scale exactly with the grain dimension as should be according to the model, so other effects should also be included when explaining the phenomenon of the colossal permittivity [15].

The question now arises whether the large dielectric response observed in ceramics is due only to an IBLC effect associated with the grain boundaries or whether there is a contribution from defects, inhomogeneities, etc [18,23,25] on a subgranular scale, which are also present in single crystals. In this context, it is important to study and compare the local permittivities of single grains in addition to those of the grain boundaries network. Such data have been recently produced by innovative characterization methods with a high lateral resolution, based on SPM but considering the IS.

In particular, it has been shown as an original use of scanning capacitance microscopy and SIM [52–54], which can be implemented, and it allows the direct imaging of the ceramic microstructure providing information (with high lateral resolution) on the local conduction and insulating properties. In fact, applying a bias between the bottom electrode and the tip, the C-AFM tip acts, on the sample surface, as a sliding metal contact, thus it is possible to collect the impedance signal of the nano-device consisting of the tip/CCTO/electrode.

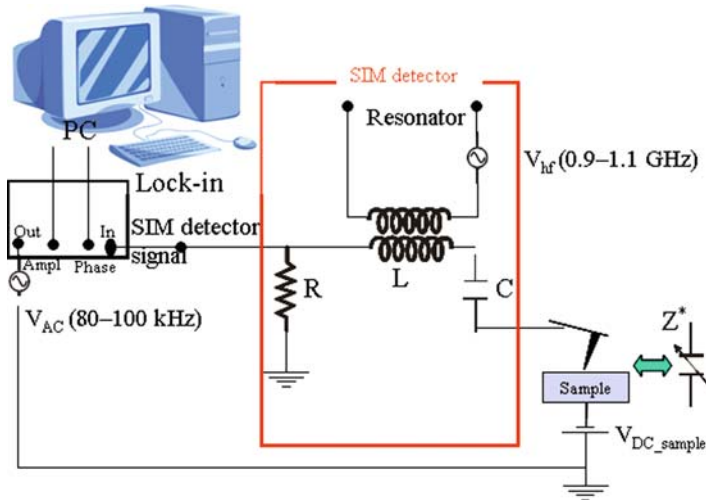


Fig. 17.12. Schematic of the scanning impedance microscopy (SIM) experimental setup. A resonator is connected to the AFM conductive tip and to the lock-in amplifier. The applied oscillating signal can be varied in the range of a few kilo Hertz. The resonator frequency is fixed at ~ 1 GHz

Measurements were performed using a back side contact, obtained by silver paint, opposite to the polished surface. SIM measurements were carried out applying an ac bias between the tip and sample was varied in the 1–10 V peak-to-peak range at 90 kHz; the resonator frequency was in the range of 1.0 ± 0.1 GHz.

The SIM described in refs. [52, 53] is an ordinary AFM equipped with a conductive tip, which has been connected to a resonator (Fig. 17.12), and the collected SIM signal amplitude is related to the sample impedance. The impedance due to the current flowing through the sample is collected at the nanometre scale, and can be described by the following equation:

$$A_{SIM} = \frac{RV_{hf}}{\sqrt{[R + Re(Z^*)]^2 + [(\omega L - (1/\omega C)) + Im(Z^*)]^2}}, \quad (17.17)$$

where R , L , C , ω , and V_{hf} are the characteristics of the resonator circuit, whereas Z^* is the complex impedance and is modelled by an ideal equivalent circuit consisting of resistors and capacitors.

A SIM map provides simultaneous imaging of both the real and imaginary parts of the complex impedance, as shown by (17.17). In particular, the higher resistance value at the grain boundary regions produces an increase of the real part of the sample impedance and [by consideration of (17.17)] a decrease in the SIM signal amplitude.

Other workers [41, 55] have demonstrated the presence of a double Schottky barrier at the grain boundary regions; however, the depletion region width has

not yet been reported. It is possible to obtain an average depletion width using SPM techniques. In particular, according to the simple IBLC model [29], the ceramics permittivity can be expressed as:

$$\varepsilon^* \approx \varepsilon_b \frac{d}{t}, \quad (17.18)$$

where ε^* is the measured permittivity, ε_b is the “real” or intrinsic bulk permittivity, d is the mean grain dimension, and t is the width of the depleted layer at the grain boundaries. The experimental value of the permittivity of the CCTO sample is about 7,000 at 100 kHz. The mean grain size is $\sim 5 \mu\text{m}$, whereas the depletion width at the grain boundaries has been evaluated to be $130 \pm 20 \text{ nm}$ from the SIM image (Fig. 17.13b), thus the effective bulk permittivity has been calculated to be ~ 160 . Despite the width value of the depletion region in this study is quite different from that of a few nanometers as suggested in the literature [56], our bulk permittivity value is in good agreement with both theoretical predictions and conventional IS measurements [6].

The dependence of the depleted region width at the grain boundaries on applied AC and DC voltage has also been evaluated. No evident variation on changing the longitudinal applied voltage has been observed. Moreover, several tips with different contact areas have been used to exclude any possible electrode-related artifacts (Pt and polycrystalline high conductivity diamond).

Finally, the influence of the grain size on the grain boundary depletion regions has been studied. Figure 17.13a shows a SIM map obtained on large sample region having a grain size distribution from $\sim 1 \mu\text{m}$ to $\sim 10 \mu\text{m}$ in diameter. The SIM image shows that the width of the grain boundary depletion regions is almost independent of the grain dimensions. SIM high-resolution image of a single CCTO grain $\sim 6 \mu\text{m}$ wide is reported in Fig. 17.13b. At the center of the grain, there is a region having two different electrical behaviors: a darker ring surrounds a spot having lighter contrast that is very similar to the main part of the grain. In Fig. 17.13c is shown the I - V characteristics performed stopping the C-AFM tip on both the dark ring (blue curve) and the light circle (red curve). The current flowing through the light regions is two orders of magnitude higher than that found in the dark ring. This indicates that the dark ring possesses a resistance two orders of magnitude higher and this signal is comparable with the resistance shown by the depletion regions at the grain boundaries. Thus, a conductive domain surrounded by insulating blocking areas has been found as a subgrain feature that influences the electrical microstructure of the CCTO ceramics.

Usually, polycrystalline materials can be described as a sequence of parallel RC elements connected in series. In the classic version of the brick-work layer model, one RC element can be related to the grains and another one to the grain boundary contribution as already discussed in Sect. 17.2.

In this context, the standard equivalent circuit of two parallel RC elements connected in series ($R_b C_b - R_{gb} C_{gb}$) needs to be modified adding a third parallel RC element ($R_d C_d$) because of the insulating subgrain feature observed

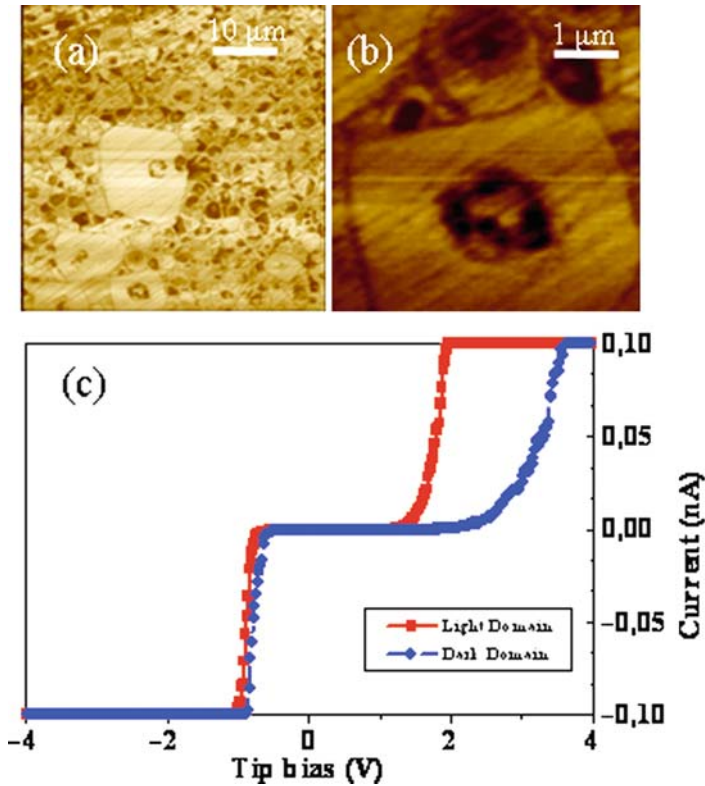


Fig. 17.13. (a) scanning impedance microscopy (SIM) image of a polycrystalline $\text{CaCu}_3\text{Ti}_4\text{O}_{12}$ (CCTO) sample with a wide grain size distribution. The width of the depleted regions at the grain boundaries is 130 nm and is independent of the grain diameter and applied voltage. (b) High-resolution image of the insulating secondary-phase inclusion within a single CCTO grain. A semiconductive grain part (light circle at the center) is surrounded by insulating secondary phase (dark ring). (c) I - V measurements collected by conductive atomic force microscopy (C-AFM) on the light circle (red) and on the dark ring (blue), respectively. The secondary phase resistance is about two orders of magnitude higher than the grain resistance and it is comparable to the grain boundaries resistance

in this study. The complex impedance of this circuit is given by rewriting (17.14) as follows including the third contribution:

$$\begin{aligned}
 Z^* &= (R_g^{-1} + i\omega C_g)^{-1} + (R_{gb}^{-1} + i\omega C_{gb})^{-1} + (R_d^{-1} + i\omega C_d)^{-1} \\
 &= \text{Re}(Z^*) - i\text{Im}(Z^*).
 \end{aligned}
 \tag{17.19}$$

The results of conventional IS studies on CCTO ceramics show that $C_g \ll C_{gb}$ and $R_g \ll R_{gb}$. Comparing our SIM and C-AFM data, no difference has been found, thus indicating that the impedance variation in the domains

is basically due to resistance variations. In fact, as previously explained, SIM maps demonstrate that the impedance variation observed in correspondence of such electrical domains is basically due to resistance variation (Fig. 17.13). Moreover, the I - V measurements have shown that $R_d \approx R_{gb} \gg R_g$ and it is reasonable to suppose that $C_d \approx C_{gb} \gg C_g$. It is therefore possible to suggest that the presence of $R_d C_d$ may give rise to a dielectric anomaly in single crystals and this may, at least in part, explain the giant permittivity reported for CCTO crystals.

The chemical nature of the insulating domains within the CCTO grains has been evaluated by energy filtered transmission electron microscopy (EFTEM). Figure 17.14a shows a TEM micrograph within a CCTO single grain. It

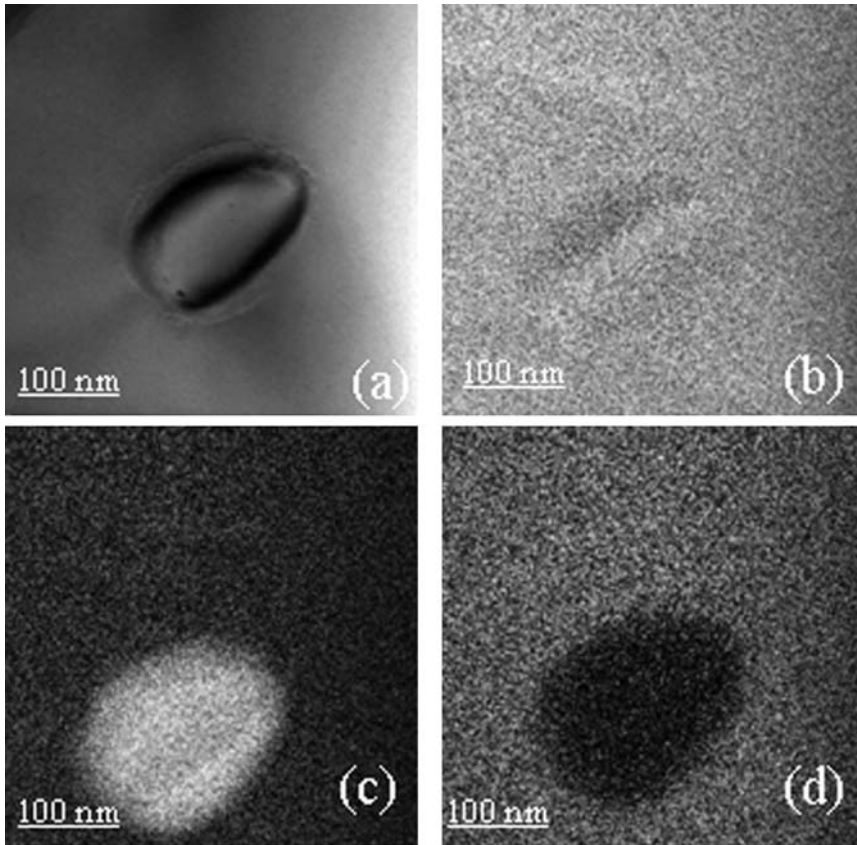


Fig. 17.14. (a) Transmission electron microscopy (TEM) image of an inclusion within $\text{CaCu}_3\text{Ti}_4\text{O}_{12}$ (CCTO) grain. It is few hundreds of nanometres in diameter. The chemical composition of such precipitates has been obtained by energy filtered TEM (EFTEM). The titanium (b), calcium (c), and copper (d) maps have been recorded and they indicate that it consists of a copper-free calcium titanate phase

is clearly visible a crystalline domain about 200 nm wide. The titanium (Fig. 17.14b), calcium (Fig. 17.14c), and copper (Fig. 17.14d) bidimensional maps indicate that it is a copper-free domain consisting of calcium titanate phase [53].

Thus, the presence of insulating precipitates may be a contribution for the raising of the colossal permittivity of CCTO ceramics or one of the possible explanations of the CCTO single crystal response.

CCTO Single Crystal

The IBLIC works reasonably well for CCTO ceramics, the observation of a giant permittivity effect in CCTO single-crystals remains perplexing as grain boundary features should not be present in single crystals. The origin of the effect in single crystals may therefore be related to some other features not observed in the ceramics, for example, nonohmic contact, modification of crystal surface composition, and/or internal boundary layers associated with defects, twins, dislocations, etc.

In CCTO single crystals, the presence of insulating domains within the material has been also found, as shown in Fig. 17.15. Both SIM and C-AFM show their insulating nature.

In fact, stopping the SPM probe on the darker region depicted in Fig. 17.15b, it is possible to collect both the conducting and insulating localized properties of the material. Figure 17.16a shows the $I-V$ curves obtained both on the domain region and on the surrounding parts, respectively. The lower current flowing through the dark domain region (Fig. 17.15b) indicates the blocking effect played by such structure attributing an insulating behavior.

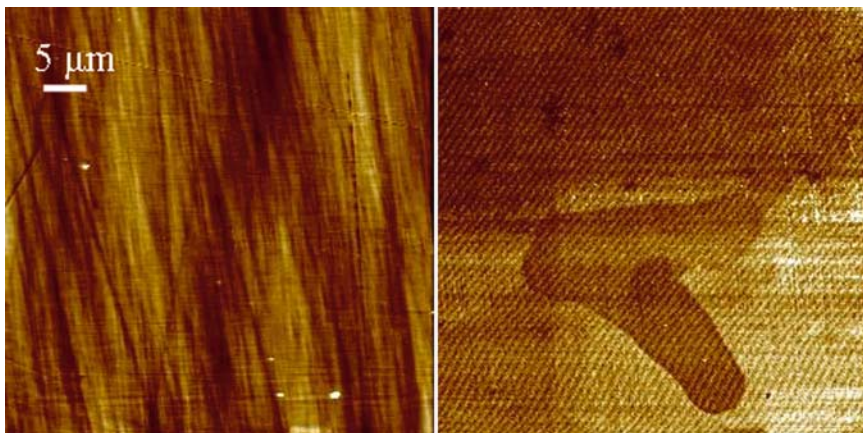


Fig. 17.15. Experimental evidence of the insulating domain presence within $\text{CaCu}_3\text{Ti}_4\text{O}_{12}$ (CCTO) single crystals by C-AFM. (a) AFM morphology image and (b) current map, respectively

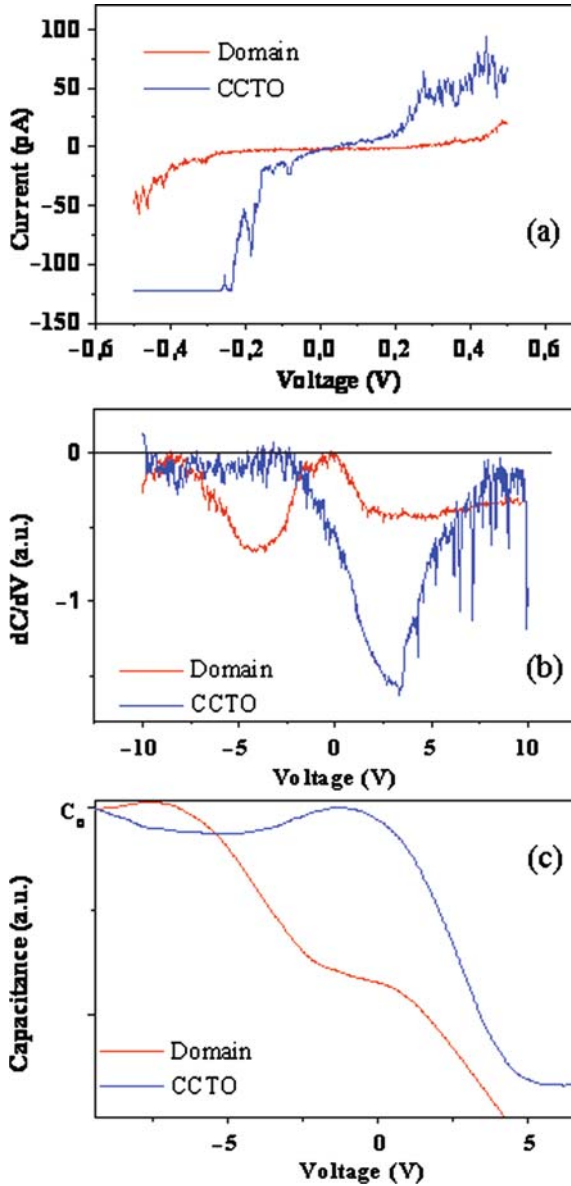


Fig. 17.16. (a) Localized $I-V$ measurements performed by conductive atomic force microscopy ($C-AFM$) inside and outside the insulating domain. $I-V$ curves demonstrated the blocking effect played by the domains. (b) Localized capacitance measurements demonstrated different insulating behavior explained by the integrated curves shown in (c) where the presence of electrical heterostructures is demonstrated

Local dC/dV versus voltage measurements on both the domain region and the surrounding part, respectively, have been collected and shown in Fig. 17.16b. The curve obtained on the pure CCTO single crystal part shows a single peak centered at about 3 V, but on the contrary, the curve obtained on the darker insulating domain shows two peaks; the first centered again at 3 V and the second at -5 V. Going at the equivalent integrated $C - V$ curves shown in Fig. 17.16c, it is evident how the pure CCTO possesses an ideal single step-like behavior of a “nano-MIS” (metalinsulatorsemiconductor) structure; on the contrary, the $C - V$ obtained on the insulating domain shows a double-step-like behavior.

TEM performed in the CCTO single crystals provided the presence inclusions, and as in the case of the polycrystalline ceramics, the inclusions are CaTiO_3 precipitates. Such composition can create the rising of heterostructures in the band diagram with the related creation of barriers for the charge carrier motion that could explain the macroscopic behavior, in particular in the case of the observed second relaxation or of the low-temperature conductivity measurements.

17.5 Summary and Conclusions

The present chapter represents an overview on the SPM-based techniques that provided a fundamental contribution on the understanding of physical phenomena involved in the giant permittivity response of electrically heterogeneous materials. Here it has been described as the potentialities of such techniques on a wider materials family, thus demonstrating a fundamental role in some of the hottest topics in the dielectrics community. In particular, after a brief introduction, in the first paragraph, the basic concepts on the dielectric relaxation in heterogeneous materials have been presented. Moreover, the relevance to employ SPM techniques to deeply understand the physical phenomena involved in the macroscopic relaxation process has been highlighted focussing on the lacks of the standard techniques and on the fundamental help provided by such SPM techniques.

Second paragraph gives a brief overview on the standard macroscopic characterization techniques as a background for the interpretation of the local SPM measurements. In particular, it has been emphasized how a deep understanding at submicrometer scale is needed to distinguish different contributions and to get reasonable numerical fitting of the equivalent circuits used for the macroscopic theoretical treatments of the dielectric relaxation.

In the third and last paragraph is presented an overview on the SPM contributions for the understanding of the physics of different “hot” materials presented also on highly relevant international journals.

Several materials and techniques have been presented to demonstrate the capabilities of reliable SPM techniques for the full understanding of the physical phenomena. In particular, STM, KPFM, SIM, and C-AFM clarified the

only hypothesized macroscopic phenomena like the IBLC effect in several polycrystalline and also in single crystal materials. In particular, SIM recognized in polycrystalline and also single crystal CCTO samples the presence of unknown subgranular features that have been subsequently identified as CaTiO_3 precipitates, and it also demonstrated their relevance on the macroscopic dielectric response.

The aim of this chapter has been to emphasize the relevance of the contribution given by SPM for the understanding of the dielectric phenomena in heterogeneous materials and the possibility to spread the employment of such techniques on a wider family of materials [57–59].

References

1. M.A. Subramanian, D. Li, N. Duan, B.A. Reisner, A.W. Sleight, J. Solid State Chem. **151**, 323 (2000).
2. C.C. Homes, T. Vogt, S.M. Shapiro, S. Wakimoto, A.P. Ramirez, Science **293**, 673 (2001).
3. T.B. Adams, D.C. Sinclair, A.R. West, Adv. Mater. **14**, 1321 (2002).
4. P. Lunkenheimer, R. Fichtl, S.G. Ebbinghaus, A. Loidl, Phys. Rev. B **70**, 172102 (2004).
5. L. He, J.B. Neaton, M.H. Cohen, D. Vanderbilt, C.C. Homes, Phys. Rev. B **65**, 214112 (2002).
6. Friedrich Kremer, Andreas Schönhals, *Broadband Dielectric Spectroscopy* (Springer, Berlin, Heidelberg, New York, 2003); ISBN 3-540-43407-0.
7. C.J.F. Böttcher, P. Bordewijk, *Theory of Electric Polarization, Vol II. Dielectrics in Time-Dependent Fields* (Elsevier, Amsterdam, Oxford, New York, 1978).
8. R.W. Wagner, Arch. Elektrotech. **2**, 371 (1914).
9. R.W. Sillars, J. Inst. Elect. Eng. **80**, 378 (1937).
10. H. Boettger, V.V. Bryksin, *Hopping conducting in solids* (Akademie, Berlin, 1985).
11. K.S. Cole, R.H. Cole, J. Chem. Phys. **9**, 341 (1941).
12. E. Warburg, Ann. Phys. **6**, 125 (1901).
13. H. Fricke, Phys. Rev. **26**, 678 (1925).
14. S.H. Liu, Phys. Rev. Lett. **55**, 529 (1985).
15. T.B. Adams, D.C. Sinclair, A.R. West, Phys. Rev. B **73**, 094124 (2006).
16. D.C. Sinclair, A.R. West, J. Appl. Phys. **66**, 3850 (1989).
17. D.R. Clarke, J. Amer. Ceram. Soc. **82**, 485 (1999).
18. D.C. Sinclair, T.B. Adams, F.D. Morrison, A.R. West, Appl. Phys. Lett. **80**, 2153 (2002).
19. R.K. Grubbs, E.L. Venturini, P.G. Clem, J.J. Richardson, B.A. Tuttle, G.A. Samara, Phys. Rev. B **72**, 104111 (2005).
20. S.Y. Chung, I.L.D. Kim, S.J.L. Kang, Nat. Mater. **3**, 774 (2004).
21. F. Greuter G. Blatter, Semicond. Sci. Technol. **5**, 111 (1990).
22. C.C. Homes, T. Vogt, S.M. Shapiro, W. Wakimoto, M.A. Subramanian, A. Ramirez, Phys. Rev. B **67**, 092106 (2003).

23. Shao S.F., Zhang J.L., Zheng P., Zhang W.L., Wang C.L. *J. Appl. Phys.* **99**, 084106 (2006).
24. Sinclair D.C., Adams T.B., Morrison F.D., West A.R. *Appl. Phys. Lett.* **80**, 2153 (2002).
25. Cohen M.H., Neaton J.B., He L., Vanderbilt D. *J. Appl. Phys.* **94**, 3299 (2003).
26. R.H. Boyd, in *Methods of Experimental Physics*, vol. 16 (Academic, Orlando FL, 1980).
27. I.M. Hodge, M.D. Ingram, A.R. West, *J. Electroanal. Chem.* **74**, 125 (1976).
28. S. Krohns, P. Lunkenheimer, S.G. Ebbinghaus, A. Loidl, *Appl. Phys. Lett.* **91**, 022910 (2007).
29. J. Wu, C.-W. Nan, Y.H. Lin, Y. Deng, *Phys. Rev. Lett.* **89**, 217601 (2002).
30. C. Wang, H.J. Zhang, P.M. He, G.H. Cao, *Appl. Phys. Lett.* **91**, 052910 (2007).
31. V. Raineri, P. Fiorenza, R. Lo Nigro, D.C. Sinclair, *Solid State Phenom.* **131–133**, 443 (2008).
32. K. Jana, S. Sarkar, B.K. Chaudhuri, *Appl. Phys. Lett.* **88**, 182901 (2006).
33. S. Sarkar, P.K. Jana, B.K. Chaudhuri, H. Sakata, *Appl. Phys. Lett.* **89**, 212905 (2006).
34. X.G. Zheng, C.N. Xu, Y. Tomokiyo, E. Tanaka, H. Yamada, Y. Soejima, *Phys. Rev. Lett.* **85**, 5170 (2000).
35. J. Tersoff, D.R. Hamann, *Phys. Rev. B* **31**, 805 (1985).
36. N.D. Lang, *Phys. Rev. B* **34**, R5947 (1986).
37. R.M. Feenstra, J.A. Stroscio, A.P. Fein, *Surf. Sci.* **181**, 295 (1987).
38. R.M. Feenstra, J.Y. Lee, M.H. Kang, G. Mayer, K.H. Rieder, *Phys. Rev. B* **73**, 035310 (2006).
39. P. Mårtensson, R.M. Feenstra, *Phys. Rev. B* **39**, 7744 (1989).
40. J. Glijsen, L.H. Tjeng, J. van Elp, H. Eskes, J. Westerink, C.A. Sawatzky, M.T. Czyzyk, *Phys. Rev. B* **38**, 11322 (1988).
41. I.-D. Kim, A. Rothschild, H.L. Tuller, *Appl. Phys. Lett.* **88**, 072902 (2006).
42. A.J. Moulson, J.M. Herbert, *Electroceramics: Materials, Properties and Applications*, 2nd edn. (Wiley, New York, 2003).
43. B. Bochu, et al. *J. Solid State Chem.* **29**, 291 (in French) (1979).
44. L. He, J.B. Neaton, D. Vanderbilt, M.H. Cohen, *Phys. Rev. B* **67**, 012103 (2003).
45. Chen, L. et al. *Appl. Phys. Lett.* **82**, 2317 (2003).
46. M. Nonnenmacher, M.P. O'Boyle, H.K. Wickramasinghe, *Appl. Phys. Lett.* **58**, 2921 (1991).
47. B.D. Huey, D. Lisjak, D.A. Bonnell, *J. Am. Ceram. Soc.* **82**, 1941 (1999).
48. S.-Y. Chung, Y.-M. Chiang. *Electrochem. Solid State Lett.* **6**, A278–A281 (2003).
49. J. Wu, C.-W. Nan, Y.H. Lin, Y. Deng, *Phys. Rev. Lett.* **89**, 217601 (2002).
50. B. Benner, P. Lunkenheimer, M. Schetter, A. Loidl, A. Reller, S.G. Ebbinghaus, *J. Appl. Phys.* **96**, 4400 (2004).
51. P. Lunkenheimer, V. Bobnar, A.V. Pronin, A.I. Ritus, A.A. Volkov, A. Loidl, *Phys. Rev. B* **66**, 052105 (2002).
52. P. Fiorenza, R. Lo Nigro, V. Raineri, R.G. Toro, M.R. Catalano, *J. Appl. Phys.* **102**, 116103 (2007).
53. P. Fiorenza, R. Lo Nigro, C. Bongiorno, V. Raineri, M.C. Ferarrelli, D.C. Sinclair, A.R. West, *Appl. Phys. Lett.* **92**, 182907 (2008).
54. P. Fiorenza, R. Lo Nigro, S. Sciuto, P. Delugas, V. Raineri, R.G. Toro, M.R. Catalano, G. Malandrino, *J. Appl. Phys.* **105**, 061634 (2009).

55. G. Zang, J. Zhang¹, P. Zheng, J. Wang, C. Wang, *J. Phys. D* **38**, 1824 (2005).
56. T.-T. Fang, C.P. Liu, *Chem. Mater.* **17**, 5167 (2005).
57. Y. Lin, L. Jiang, R. Zhao, C.-W. Nan, *Phys. Rev. B* **72**, 014103 (2005).
58. Sudipta Sarkar, Pradip Kumar Jana, B.K. Chaudhuri, *App. Phys. Lett.* **92**, 142901 (2008).
59. S. Sarkar, P.K. Jana, B.K. Chaudhuri, H. Sakata, *Appl. Phys. Lett.* **89**, 212905 (2006).

Controlling Wear on Nanoscale

Mario D'Acunto

Summary. One definition of wear in a tribological sense is that it is the progressive loss of mass from the operating surface of a body as a result of relative motion with another surface. Generally, tribology may be defined as the study of friction wear and lubrication, but in comparison with friction forces or lubrication methods, very little work has been done on the extremely important area of wear and surface damage. For systems consisting of common materials (metals, polymers, ceramics), there are at least four main mechanisms by which wear and surface damage can occur between solids in relative motion: abrasive-, adhesive-, chemical-, corrosive-, or fatigue-wear mechanisms. The occurrence of such different wear mechanisms involves the fundamental mechanics of molecular and supra-molecular scale. As a consequence, the understanding of surface properties on nanoscale level should be generated to also have a satisfactory knowledge of materials response on micrometer scale. Investigation of the fundamental characteristics of wear at the nanoscale is complicated by some factors and forces that acting on such scale level have not yet been addressed in the tribology of macro-systems. Since these forces are sensitive to the environment and surface condition of the specimens, it is quite difficult to determine them accurately. Further, quantification of wear is not straightforward since the amount of wear is often too small to be detected by surface-sensitive instruments. Nanowear measurements have been object of rapid development of precise measuring tools following to the introduction of scanning-probe microscopy (SPM) family. The advent of SPM family opened a new unique opportunity to study wear mechanisms on nanoscale, making possible to simulate one single asperity surface sliding on the other one and controlling many physical parameters during the contact dynamics. Further, SPM made possible to have a direct connection between experimental data and theoretical models.

This chapter is devoted to give a panoramic lecture of some experimental results and models on the scientific and technological advantage when the basic wear mechanisms on nanoscale are known and controlled.

Key words: Abrasive- and adhesive-wear mechanisms, Archard formula, Atomistic wear single asperity sliding contact, Double-well model, Scanning probe microscopy.

18.1 Introduction

The atomistic mechanisms and dynamics of the interactions of two materials during relative motion need to be understood in order to develop fundamental understanding of adhesion, friction, wear, indentation, and lubrication processes. Consequently the importance of investigating single asperity contacts in studies of the fundamental mechanical and tribological properties of surfaces and interfaces on nanoscale has long been recognised. The recent emergence and proliferation of scanning-probe microscopies (SPMs) and computational techniques for simulating tip-surface interactions and interfacial properties has allowed systematic investigations of interfacial problems with high resolution as well as ways and means for modifying and manipulating nanoscale structures. These advances have led to the appearance of the new field of nanotribology, which pertains to experimental and theoretical investigations of interfacial processes occurring during adhesion, friction, wear, scratching, nanoindentation, and thin-film lubrication at sliding surfaces on scales ranging from the atomic- and molecular- to the microscale [1, 2]. The nanotribological studies are needed to develop fundamental understanding of interfacial phenomena on atomic, molecular, and supra-molecular scale and to study interfacial phenomena in nanostructures used in the area commonly defined as nanotechnology.

A definition of wear mechanisms can be as follows: the succession of events whereby atoms, products of chemical conversion, fragments, debris, are induced to leave the system and identified in a manner that embody or suggest solutions [3]. These solutions may include choice of materials, lubricants, contact condition, and manner of operation of the mechanical system. This definition also includes (a) the conditions imposed upon surfaces, such as contact stress, shear stress, and number of cycles of these stress applications, (b) the changing composition of substances in the sliding interfaces due to both sliding and these will likely influence the friction and shearing stresses imposed in sliding contact, (c) the processes whereby particles (atoms, molecules) are separated from a surface, and (d) the circulation and path of those particles before they are expelled as debris. There are more than 24 common terms to describe wear mechanisms [4]. Generally, such wear mechanisms are listed in six categories of terms. Among the most important categories, we limit to discuss the most remarkable, such as adhesion-based wear mechanisms, fatigue-derived wear mechanisms, and abrasion.

Adhesive wear is the most difficult term to define. It may denote a particular type of material loss due to high local friction (which is often attributed to adhesion) and is a tempting term to use because high local friction produces tearing and fragmentation, whereas lubricants diminish tearing. Some time, lubricated wear is taken to be the opposite of adhesive wear. An often overlooked aspect of adhesive wear is that if adhesion is effective, the particles torn loose by adhesion will re-adhere and not be expelled as debris.

Wear mechanisms deriving from cyclic stressing imply fatigue of materials such as fretting. Fretting is the oscillatory sliding motion of small slip amplitude between two nominally stationary contacting surfaces. It occurs when two quasi-static surfaces in contact are actually subjected to deformation and/or vibrations. Damage caused by fretting is one of the severe problems in engineering since it contributes substantially to shorten life of mechanical components. Analogous to fretting is the delamination mechanisms, which describes a type of wear debris developed by low cycle fatigue when surfaces are rubbed repeatedly by a small shaped slider.

Abrasive-wear mechanisms denote a wide class of wearing processes occurring when a hard solid, particle, or asperity rubs against a softer surface. Examples of abrasive wear include the smoothing of wood by sandpaper, the action of emery paper in polishing metals, grinding processes, and harmful effects of hard particles (e.g., sand, dirt, carbide or silica inclusions, oxides, and work-hardened wear debris) in lubricated systems. Abrasive wear is a form of microcutting, involving a hard particle or shape that indents, grooves, and then cuts material out of the surface.

On macroscale, wear has been the subject of much study and many general mechanisms are reasonably well understood [3,5]. On nanoscale, this is yet not so. The mechanisms are largely unknown and not readily determined. Nevertheless, the occurrence of wear processes at nanoscale has been demonstrated [6–11].

Let us consider the possible mechanisms of wear to be essentially two types: abrasive- and adhesive-wear mechanism. On macroscale, if the contribution to wear is essentially of adhesive nature, it is possible to quantify the volume of wear for plastic contacts as

$$V = \frac{k_{\text{adh}}Lx}{H} \quad (18.1)$$

where k_{adh} is a non-dimensional wear coefficient dependent on the materials in contact, L is the load applied, H is the hardness coefficient, and x is the sliding distance. Equation (18.1) is commonly referred as Archard's equation [12]. It suggests that the wear rate is independent of the apparent area, directly proportional to the applied load, and the wear rate is constant with sliding distance (or time). These properties seem to be analogous to Amontons' laws of friction. Furthermore, (18.1) suggests that the probability of decohesion of a certain volume of material and formation of a wear particle increases with real area of contact, $A_r (= L/H)$ and the sliding distance. An equation of the form similar to Archard's equation for adhesive wear is found to cover a wide range of abrasive situations

$$V = \frac{k_{\text{abr}}Lx}{H} \quad (18.2)$$

From the Archard formulation, the wear rate, that is, the volume of material transferred from one asperity to another per unit of sliding length, is simply proportional to the contact area multiplied for a constant (wear coefficient) to be defined by the contact conditions when the contact is fully established.

In this sense, there is no discontinuity between wear occurring at nanoscale and at larger ones: if the real contact area decreases to nanometric scales, the wear rate should decrease proportionally. From a predictive point of view, the Archard formulation for wear rate is of limited help, because there is an unyielding uncertainty related to the wear coefficient involved [13,14]. As the contact scale decreases to near-atomic length, the wear mechanisms involve plastic deformation, crack propagation, such as grooving, delamination, or fatigue, and they cannot be predominant if only a few atomic layers have to be taken in account in the contact. Molecular dynamics simulations operating with contact scales limited to few atomic layers showed that at this small length scale, wear is essentially controlled by the dragging of atoms from their initial positions [15,16]. Moreover, a recent experiment by Gotsmann and Lantz confirms that Archard formula seems to be limited to macroscale, while atomistic wear in a single asperity sliding contact shows that wear occurs through an atom-by-atom loss process, which in turn implies the breaking of individual bonds [17]. Such bond breaking can be described by a thermally activated process governed by Arrhenius kinetics [18–21].

Several SPM-based techniques are used for fundamental studies of tribological features on micro/nanoscale. SPM family including atomic force microscopy (AFM), friction force microscopy (FFM), scanning force microscopy (SFM), in addition to surface forces apparatus (SFA), operates in ultrahigh vacuum (UHV) or in controlled environments and they constitute the basic imaging and force measuring tools. These techniques have made possible the study of single-asperity contacts in a controlled fashion, eliminating complications encountered in more traditional methods, such as pin-on-disk, indenters, profilometers, in which the contact is made between rough surfaces involving a large and uncontrolled number of asperities (Fig. 18.1). The

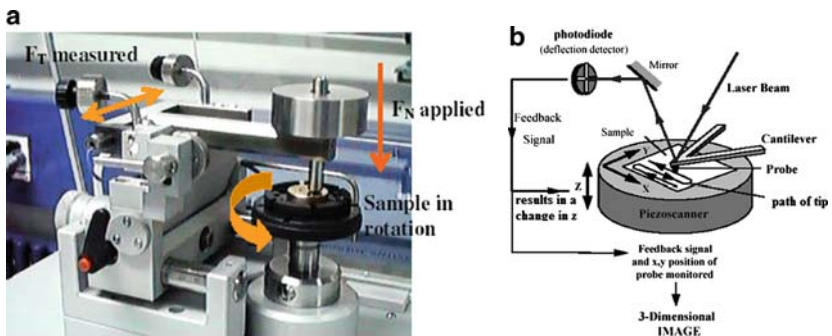


Fig. 18.1. On *left*, typical wear machine where a flat or a sphere shaped indenter is loaded on to the test sample with a precisely known weight. A pin is mounted on a stiff lever, designed as a frictionless force transducer. As the sample is either rotating or sliding, resulting frictional forces acting between the pin and the sample are measured by very small deflections of the lever. Wear coefficients for both the sample and material are calculated from the volume of the material lost during a specific friction run. (Reprinted from CSM instruments, Switzerland, with permission). On *right*, typical experimental set up for an AFM configuration

possibility of single asperity sliding contact regime where the tribological conditions are well known is addressing new knowledge giving the possibility for the first time for controlling wear mechanisms on nanoscale. The scientific and technological advantages deriving from this type of knowledge could be impressive in the next decade.

This chapter is structured as follows: in Sect. 18.2, prominent wear mechanisms on nanoscale are discussed; in Sect. 18.3, activation process describing onset wear on the nanoscale will be presented.

18.2 Molecular and Supra-Molecular Features for Basic Wear Mechanism

Any SPM-based method for the observation of the nanowear processes is given by the production of defects during sliding contact making use of a probe tip. Generally, the probe tip is used to scan repeatedly the same area of the sample. The first example on which we focus is the wear response of mica to AFM tests. The AFM has been used to study such processes on mica both in controlled atmosphere [19] and in aqueous environment [22]. Muscovite mica, a layered aluminosilicate in the form of $\text{KAl}_2(\text{Si}_3\text{AlO}_{10})(\text{OH})_2$, has been frequently employed as a model surface for the investigation of the atomic scale relationship between friction [23, 24], adhesion [25–27], and wear [8, 28] by the AFM and SFA [29] measurements due to the ability to readily generate large domains of atomically smooth surface. The layered sheet structure of mica has approximately 1 nm periodicity normal to the (001) crystal plane, which is defined by the repeating boundary layers of K^+ (Fig. 18.2). The K^+ ions electrostatically bind the layers together and act to neutralise the net negative charge associated with the partial substitution of Si with Al (on average one out of every four) in the basal plane. The K^+ interlayer demarcates

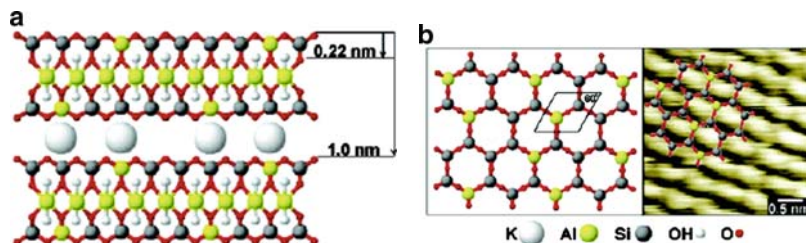


Fig. 18.2. (a) Cross-section depicting micas layered structure along the c axis with interlayer distances referenced to the $\text{SiO}_3/\text{AlO}_3$ tetrahedral basal plane. (b) A graphic representation and an unfiltered AFM topographic image ($3.4 \times 3.4 \text{ nm}^2$) of the mica (001) crystal plane showing the 0.52 nm lattice periodicity. (Reprinted from [27], with permission)

the preferred cleavage planes and defines the atomic boundaries along which shearing of the layers most readily occurs. Under aqueous conditions, the outmost K^+ layer is exchanged into solution and moreover, with the K^+ , the outermost mica surface exposes a hexagonally arrayed pattern of oxygen atoms with a 0.52-nm periodicity (Fig. 18.2).

AFM studies of mica surfaces have shown that under sufficiently high loads, the surface becomes damaged and deep holes of approximately 0.2 or 1 nm are formed, depending on the imaging conditions. Wear scars have been produced by repeatedly scanning over the same region at a fixed load. Point defects are created as a result of the lowering of the activation energy for bond breaking by the applied stresses [22]. In the tip sample contact area, the number of defects created is given by [19]

$$N_{\text{def}}(L) = t_{\text{res}} n_0 A(L) \nu e^{\{-[\varepsilon_0 - \delta\varepsilon(L)]/k_B T\}}, \quad (18.3)$$

where t_{res} is the residence time of the tip, n_0 ($\sim 10^{20} \text{ m}^{-2}$) the surface density of atoms, and $A(L)$ the contact area as a function of total load L , ν is the attempt frequency ($\sim 10^{13} \text{ s}^{-1}$) to overcome the energy barrier, ε_0 , to break an Si–O bond, $\delta\varepsilon(L)$ is the decrease in activation energy due to the applied stress, which can be estimated by multiplying the force acting on an atom by the resulting bond stretch. The average force per atom is given by the pressure $P(L)$ in the contact area divided by n_0 . The relative deformation (bond stretch), $\Delta l/l$, should be on the same order as the Hertzian indentation depth h , divided by the characteristic length of the affected area, which is on the order of the contact radius $r(L)$: $\Delta l/l \approx h/r(L)$, with $h \approx r^2(L)/2R$ for a tip radius $R \gg h$. Thus the decrease in activation energy due to the applied stress $\delta\varepsilon(L)$ changes as $L^{2/3}$. Wear damage will occur when the defect density, $N_{\text{def}}/n_0 A(L)$, reaches a critical value η_{crit} , and gives rise to nucleation of a hole. Thus the number of consecutive scans necessary to cause a visible damage (0.2 nm holes) on the muscovite mica surface is given by the equality of critical value with the product of number scans with the defect density (Figs. 18.3 and 18.4). Such threshold value for number scans can be expressed by the following cumulative curve [19]:

$$N_{\text{scans}} = AL^{-2/3} e^{-BL^{2/3}}, \quad (18.4)$$

where A and B are two constants.

Helt and Batteas reported the wear of mica surfaces under aqueous environments of varying pH [22]. Using the AFM, they have evaluated the simultaneous changes in surface topography, friction, and adhesion at the mica surface under controlled pH conditions. By varying pH and applied loads, the incidence and mechanisms of the observed wear were found to be influenced dramatically. In fact, with the appropriate control of applied load, number of scans, and pH, they have been able, for the first time, to directly observe

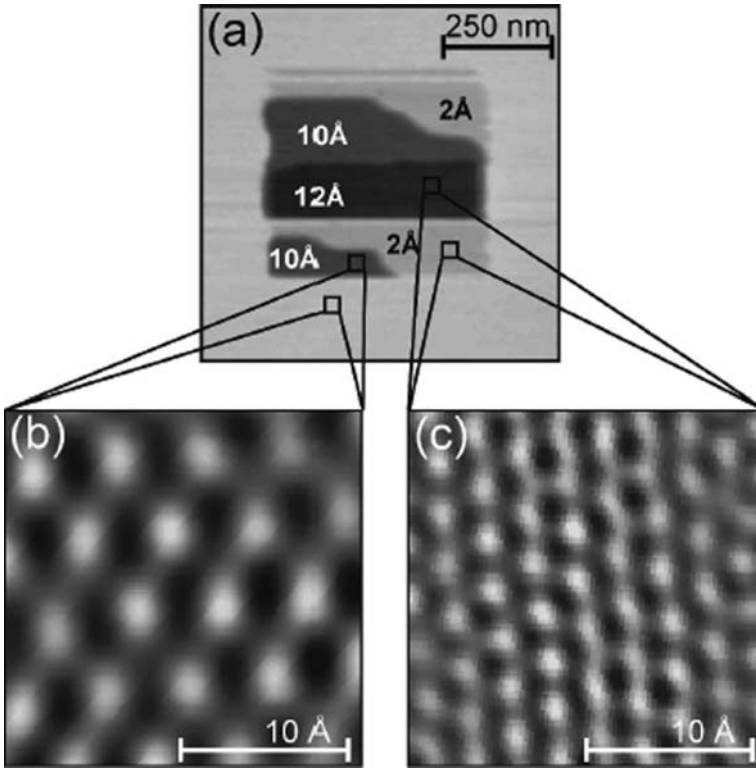


Fig. 18.3. AFM images of a hole scratched into muscovite mica with an AFM tip. The images were recorded making use of a silicon cantilever with a nominal force constant of 2.1 N/m and a nominal tip radius of 50 nm was used. The hole was created after eight consecutive scans with a load of 230 nN, over an area of $500 \times 500 \text{ nm}^2$ at a scan speed of 60 ms/line. (a) Image size is $0.75 \times 0.75 \mu\text{m}^2$. The hole obtained with these parameters shows 2, 10, and 12 Å deep regions (denoted by grey levels). Marked with a square are regions where the atomic lattice resolution images, shown Fourier filtered in (b) and (c), were obtained. The periodicities in these images are 5.0 Å (b) and 3.0 Å (c). The latter is rotated by $\sim 30^\circ$. (Reprinted from [19], with permission)

defect production, consisting in the breaking of terminal surface bonds, as detected in the form of a surface restructuring prior to the formation of 2 Å pits or scars, generally, considered gross wear. A further result shown by Helt and Batteas is that the degree of wear can be completely controlled by modifications of imaging conditions and pH. Under low pH conditions ($\text{pH} \leq 3$), the wear threshold pressure relating to the formation of 2 Å pits is equivalent to that of a simply humid environment and complete inhibition of wear is observed for tests conducted at the threshold pressure for higher pH studies.

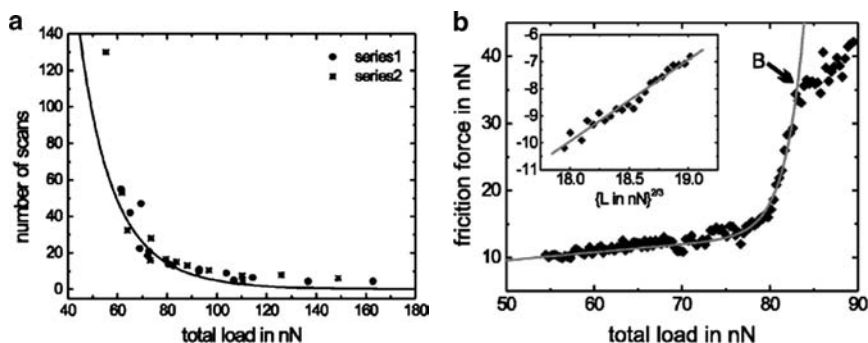


Fig. 18.4. On *left*, it is shown the number of consecutive scans necessary to cause visible damage to the mica surface (0.2 nm holes). Data from two series of experiments measured with the same tip (force constant 0.4 N/m) on two consecutive days are shown. On *right*, friction versus load curve obtained during the creation of 0.2 and 1 nm deep holes. Rapid increase in friction is observed above 75 nN load, corresponding to a large increase in the production of defects. When the density of defects reaches a critical value, the friction levels off and a wear scar detectable by AFM is produced. The inset shows the region, where the friction increases rapidly, in a semilog representation: $\ln\{[F(L)-cL^{2/3}]/L^{2/3}\}$ is plotted versus $L^{2/3}$ to allow for a simple linear fit of the coefficients B (4). (Adapted from [19], with permission)

At the highest pH conditions ($\text{pH} \approx 8$), the mica surface undergoes rapid wear. Even under low loads, wear scars of approximately 2 Å in depth form readily. The threshold pressure for onset wear at $\text{pH} = 8$ was found to be nearly 1.2 GPa. With additional scans (>10), wear scars of 10 Å or more are rapidly formed (Fig. 18.5). Interestingly, the depth of the wear scars formed under these pH conditions is found to scale roughly linearly with the number of scans for wear scars of up to 60–70 Å in depth [22].

A remarkable approach for controlling abrasive wear has been shown by Dickinson et al. who performed a series of fundamental experiments on nanowear on single-crystal calcite in saturated CaCO_3 solutions [30–32] to quantify the shear stress required to detach single crystal salt particles from a soda lime glass substrate under controlled atmospheres of known humidity [33], and corrosive wear on a single-crystal brushite ($\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$) in aqueous solution [34]. Generally, in such experiments the tribological stimulus was provided by the scanning-probe tip itself. In the case of single-crystal calcite, the dissolution and growth of calcite crystals in aqueous solution have been extensively studied due to their importance in mineral formation, global CO_2 exchange, and interactions with environmental contaminants. Localised corrosive wear can readily be induced by the SFM tip on calcite surfaces imaged in aqueous solutions. Localised dissolution along pre-existing steps can also be produced by simply indenting the surface near the edge of an etch pit. Wear produced by drawing the probe tip back and forth in a repetitive, linear fashion across a monolayer step can be unambiguously interpreted in terms

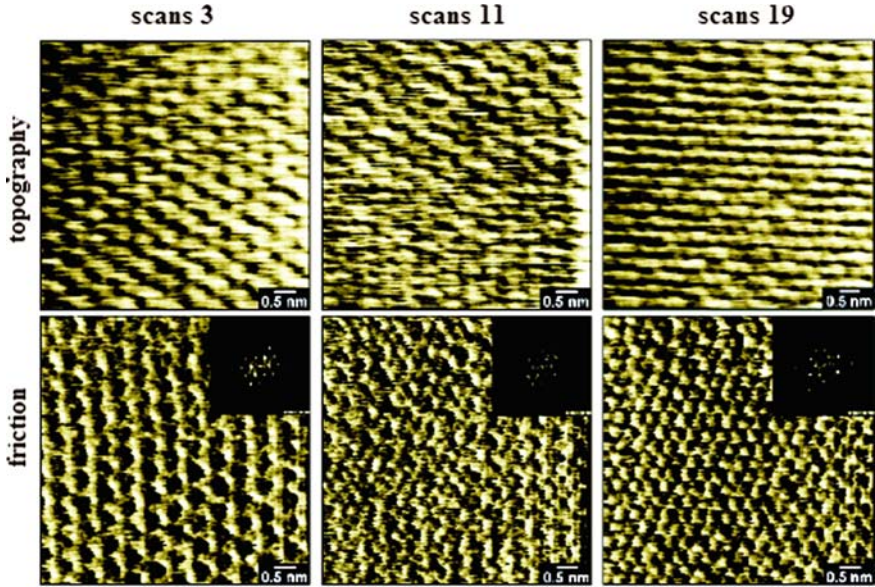


Fig. 18.5. Topography and frictional images monitoring the tribochemical surface reconstruction of the native mica lattices from a 5.2 \AA (scans 3) to an approximately 3 \AA (scans 19) periodicity under the local stress of an AFM tip. Images represent the 3rd, 11th, and 19th scans during continuous scanning at a rate of 33 Hz over a $7 \times 7 \text{ nm}^2$ region at $\text{pH} = 5$ and an applied load of 9 nN (mean pressure of 1.25 GPa). The 2D FFT insert of the friction micrographs illustrates the lattices transition where the intermediate state (11th scan) possesses a mixed 3 and 5.2 \AA periodicity. (Reprinted from [22], with permission)

of volume activated, double-kink nucleation in the strain field of the SFM tip (Fig. 18.6). Experimental details can be found in [30]; here, we like to discuss the activation process.

Wear tracks can be easily imaged in a geometry as presented in Fig. 18.6. The wear track growth rate ($\Delta x/\Delta t$) is a strong function of contact force. In Fig. 18.7, it is presented in a plot of the wear track growth rate for wear tracks force.

The magnitude of the stresses surrounding the SFM tip can be estimated assuming an isotropic, elastic behaviour. Under such assumptions, the maximum tensile stress involves the radial components, σ_r , along the circle where the SFM tip contacts the substrate. Double-kink nucleation (the removal of ions from initially complete or intact steps) should be enhanced over a modest range of distances from the tip. The radial stress, σ_r , can be calculated by the Hertzian relation [35]:

$$\sigma_r = \frac{(1 - 2\nu)}{\pi} \left(\frac{2PE^2}{9R^2(1 - \nu^2)^2} \right)^{1/3} \quad (18.5)$$

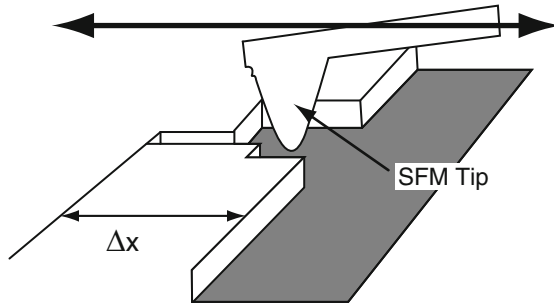


Fig. 18.6. Schematic sketch of wear geometry of the experiment design from Dickinson and co-workers. The SFM tip is drawn back and forth in a linear fashion across the edge of pre-existing etch pit. (Reprinted from [30], with permission)

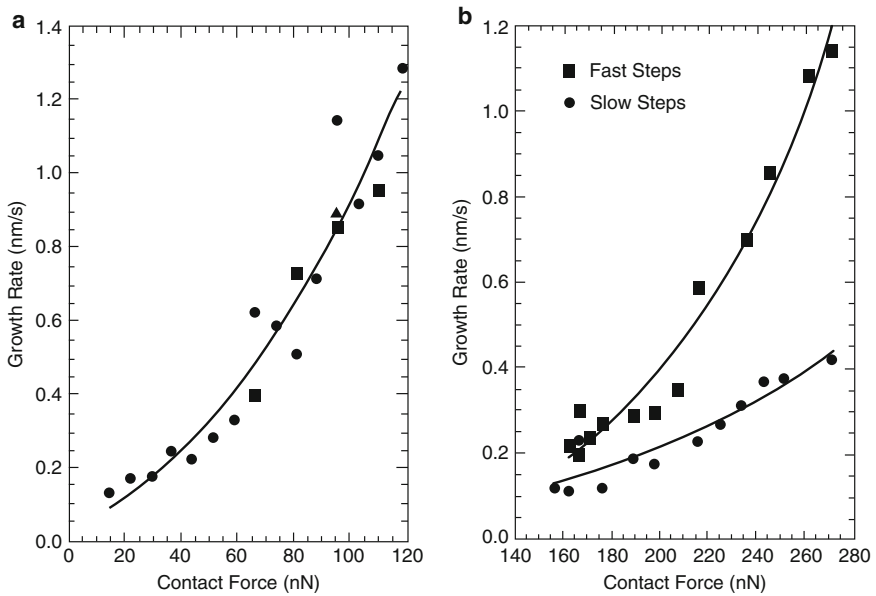


Fig. 18.7. (a) Wear track growth rate versus contact force during flow of solution and repeated linear scanning. The data points represented by different symbols were made on different calcite samples with different cantilevers. The *continue line* is a least-squares fit of (5) to the data. The measurements employed $500 \times 500 \text{ nm}^2$ scans acquired at a tip velocity of $12 \mu\text{m/s}$. (b) Wear track growth rate for both fast and slow steps versus contact force under high degree of saturation. As in the case (a), the *continue lines* represent least-squares fits of (5) to the data. (Reprinted from [30], with permission)

where ν is the Poisson's ratio and E is the reduced Young's modulus (approximately 58 GPa), R is the tip radius and $P = F_N/A_r$, where A_r is the real area of contact, in the normal pressure. An applied normal force of 100 nN yields a maximum radial stress of about 500 MPa. The dependence of the growth rate of the wear track V on contact force is readily modelled with an Arrhenius-type expression [30]

$$V = V_0 \exp \left[- \left(\frac{E_{\text{act}} - \nu^* \sigma_r}{k_B T} \right) \right], \quad (18.6)$$

where V_0 is the appropriate pre-exponential, E_{act} is the zero stress activation energy for double-kink nucleation and ν^* is an activation volume. The best-fit curve of (18.6) to the data of Fig. 18.7a (unsaturated solution) corresponds to $V_0 \exp(-E_{\text{act}}/k_B T) \approx 5.8$ pm/s and $\nu^* \approx 4.4 \times 10^{-29}$ m³. This activation volume is slightly larger than the average volume per ion in the CaCO₃ lattice (3.3×10^{-29} m³), making it reasonable to suppose that this activation volume corresponds to the displacement of one or perhaps two ions from a step site. A similar analysis of the wear track growth rates at slow steps under saturated conditions (Fig. 18.7b) yields $\nu^* \approx 4.4 \times 10^{-29}$ m³. This agreement is consistent with the expectation that the degree of saturation controls the lifetime of nucleated kinks and not the nucleation rate.

The description of the wear track in terms of the rate of mechanically enhanced double-kink nucleation was found particularly efficient to explain the corrosive wear of single-crystal brushite in aqueous solution. In the case of atomic-scale corrosion, an understanding of fundamental effects is particularly necessary, requiring accurate details of interactions of the water molecules with stressed surfaces. The measurements of E_{act} , ν^* , and the factors for the key dissolution processes (such as pit nucleation, double-kink nucleation) provide an experimental foundation for a complete predictive model for corrosive wear of crystalline materials.

Chemical factors have been shown to be relevant for the particles removal, as in the case of single-crystal NaCl salt [33]. The nominal shear stress required to remove such particles from a soda lime glass substrate with an SFM-probe tip is a strong function of particle size and relative humidity (RH). At 3% RH, particles larger than about 500 nm on an edge could not be removed by the probe tip at any an accessible contact force. Increasing the RH to 30% dramatically reduces the stress required for particle removal and promotes the removal of much larger particles. Further increases in RH yield smaller decreases in the failure stress. The humidity dependence of the shear stress can be described in terms of the expected change in interfacial energy due to water absorption. The stress required for particle removal scales roughly with the inverse square root of the particle area. Fracture mechanics can help to suggest that this is associated with an initial flaw size that is roughly proportional to their area. An improved understanding for particle removal could be an important factor in a number of technologies including integrated circuit manufacture and optical component manufacture. The experiments

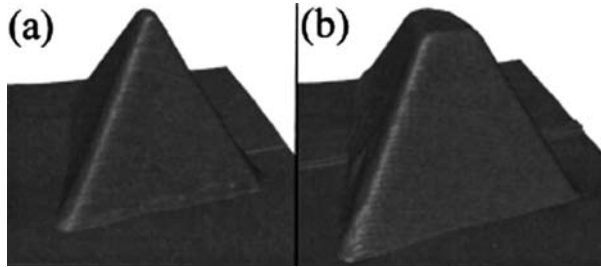


Fig. 18.8. AFM images of (a) a fresh AFM tip and (b) the same tip after wear. The AFM tip was scanned 45 scans times across a sodium trisilicate glass substrate in ammonium hydroxide solution ($pH \sim 11$) at an applied force of 125 nN. Each image has been slightly cropped from the original $1,500 \times 1,500 \text{ nm}^2$ image. The vertical (z) axis has the same scale as x and y axes. (Reprinted from [36], with permission)

of Dickinson and co-workers showed in special way how the combination of chemical and mechanical effects plays a key role for tribological features on molecular scale.

The connection between chemical environments and onset wear mechanisms was further investigated by Dickinson and co-workers directly on silicon nitride tips examined in a variety of substrates in aqueous solutions [36]. They showed that the chemical nature of the substrate plays an important role: significant wear was observed only when the substrate surface is populated with appropriate metal-hydroxide bonds. Mica and calcite substrates, whose water-exposed cleavage surfaces lack these bonds, produced little if any tip wear. As a function of contact force F_N and scan duration t , the length of the tips decreases approximately as $(F_N t)^{0.5}$ (Fig. 18.8). It is a reasonable assumption that pressure-induced intermediate states involving hydroxyl groups form on both the tip and the substrate; chemical reactions subsequently form transient bridging chemical bonds that are responsible for tip wear (Fig. 18.9).

In a similar way, Stevens et al. [37] showed that changes in contact stress due to the wear of the AFM tip dramatically affect the observed wear rates in single-asperity wear on sodium trisilicate glass in basic solutions using AFM, where the silicon nitride tip was used to both tribologically load and image the surface [37]. The changes in surface elevation for both the tip and the substrate display the same dependence on true stress and time for the case of square-raster scanning over micron dimensions. In the case of linear reciprocal scanning, conformal wear substantially modifies the rates and morphology of the wear of both the tip and the substrate.

Recently, a series of nanowear tests where the SPM in UHV has been used to realise and detect atomic-scale abrasion on KBr (001) have been performed [38, 39]. Such wear tests showed how the debris extracted from the KBr (001) surface by a silicon tip is reorganised on the atomic scale (Fig. 18.10). The dependence of wear versus time was investigated scratching across a single line for a long time at fixed load and velocity (5,120 scans at $F_N = 11.2 \text{ nN}$

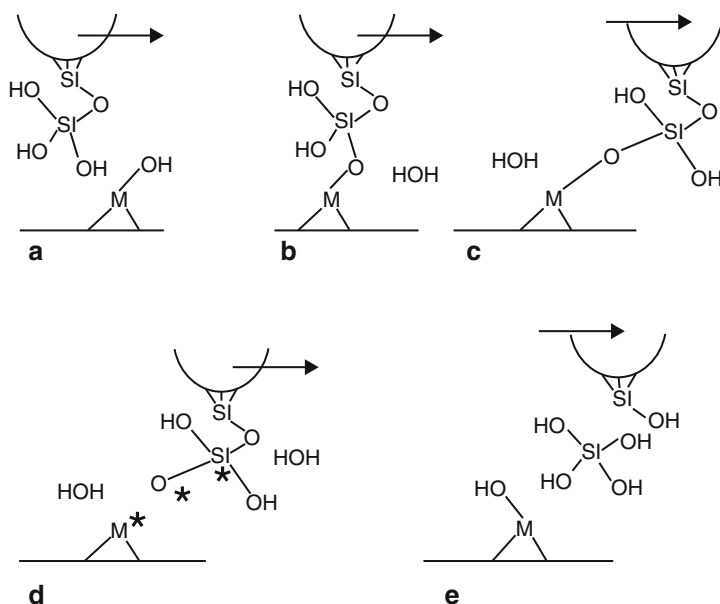


Fig. 18.9. Chemical reaction model for tribochemical tip wear. In water, the silicon nitride AFM tip is coated with a layer of silicon oxide. (a) During scanning, a hydroxyl group on the tip (Si-OH) encounters a hydroxyl group on the surface (M-OH). (b) The two groups react to form a bond, releasing water. (c) As the tip continues its motion across the surface, the bridging bond is stretched. (d) The bond breaks. Energy that was formerly stored in the stretched bridging bond is deposited in the atoms on either side (asterisks). (e) These “activated” atoms react with water, which may also break additional bonds. In this case, a bond between the silicon atom and the AFM tip breaks, releasing the silicon atom from the tip and producing aqueous $\text{Si}(\text{OH})_4$. Further tip motion yields additional reactions and more broken Si-O bonds on the tip surface, as well as M-O bonds on the substrate. (Reprinted from [36], with permission)

and scanning velocity of $2.5 \mu\text{m/s}$). The mean lateral force $\langle F_L \rangle$ increased continuously from an initial value as a function of the number of scratches N . The behaviour of the lateral force is a clear demonstration of the abrasive character of the wear process. In fact, the ions seem to be removed and released by pairs or by small cluster of ion pairs, and no dramatic crack events were observed in thousands of scans, and the regular arrangement of the deposited material resembles an epitaxial growth of thin ionic films, where the ions are piled up pair by pair. The ion removal can also be recognised in the irregularity of the friction loops, where the number of displaced ion pairs is comparable to the number of slip events, as the ions were detached when the tip jumps from one stick position to the next one. Generally, the stick-slip appears more irregular compared to atomic stick-slip observed without wear.

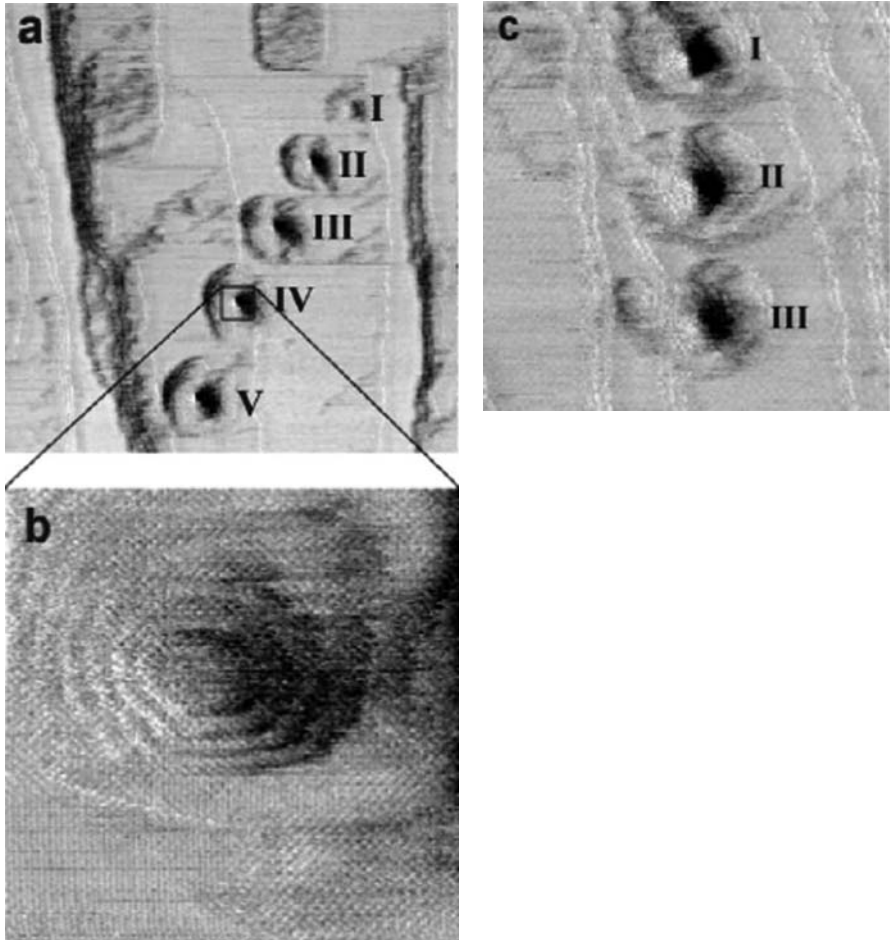


Fig. 18.10. (a) Lateral force images of pits and mounds produced by 256 scratches on $5 \times 5 \text{ nm}^2$ areas with $F_N = 5.7, 10.0, 14.3, 18.6,$ and 22.8 nN (up to down) and $v = 25 \text{ nm/s}$. Frame size: 150 nm . (b) Lateral force image of the fourth pit. Frame size: 20 nm . (c) Lateral force images of pits and mounds produced by 256 scratches on $5 \times 5 \text{ nm}^2$ areas with $v = 25, 50,$ and 100 nm/s (upper to lower) and $F_N = 14.1 \text{ nN}$. (Reprinted from [38], with permission). Frame size: 74 nm . (Reprinted from [38], with permission)

The AFM scratch process was also used to produce nanoscale regular patterns on KBr (001) in UHV. In such tests, wear of single atomic layers has been observed when the AFM tip is repeatedly scanned across a line. The initially flat surface is rearranged in a quasi-periodic pattern of mounds and pits (Figs. 18.11 and 18.12). The distance between the pits was observed to be about 40 nm when normal forces of a few nN are applied, and it slowly increases with the load. If a square area is scanned, a pattern of regular

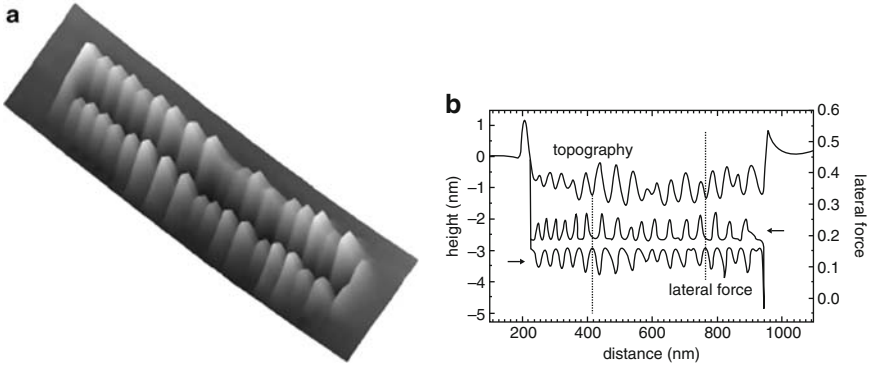


Fig. 18.11. (a) Topography image of a groove formed on KBr(001) after 512 scratches along the (100) direction with a normal load of $F_N = 26.6$ nN and a scanning velocity of $2.67 \mu\text{m/s}$. The groove is surrounded by a pattern of 34 mounds aligned along the two sides, and by two mounds piled up at its ends. The height difference between the top of the mounds and the bottom of the pits in the groove is about 2.5 nm. The distance between two adjacent pits or mounds is about 40 nm, as well as the distance between the two rods or mounds. (b) Cross-section of the topography along the ground of the groove and lateral force acquired while scratching the groove. The *lower* and *upper* parts of the lateral force loop were obtained while scanning forwards and backwards, as indicated by the arrows. The *dotted lines* indicate positions in the pits where the tip is decelerated as derived from an increase in the lateral force in either direction. (Reprinted from [39], with permission)

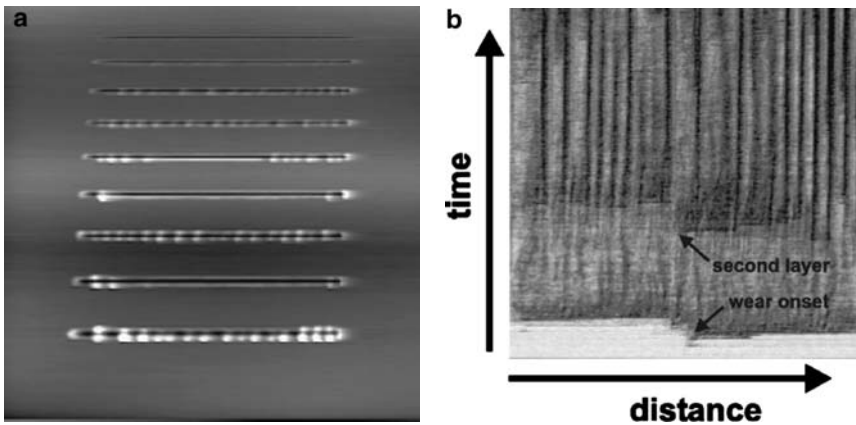


Fig. 18.12. On *left*, topography map of a region scratched with the following loads (up to down): 1.7, 5.3, 8.8, 12.4, 15.9, 19.5, 23, 26.6, and 30.1 nN; frame size: $1.2 \mu\text{m}$. On *right*, development of lateral force with time while scratching a single line of 800 nm length forwards and backwards. In horizontal direction, the lateral force during forward scan is plotted, where darker colour means stronger lateral force; the vertical direction is the time axis. (Reprinted from [39], with permission)

ripples was formed. Ripple formation due to perturbations acting on a surface is commonly observed on macroscopic scale. Well-known examples are given by wind-blown sand dunes observed in the desert and in the submarine sands [40]. In the case of the continuous scratching of AFM-probe tip, such features can be interpreted within an erosion–diffusion process induced by a periodic increase in the strain produced by the scanning tip.

Recently, Gotsmann and Lantz reported a study into the nature of single-asperity wear of sharp silicon tips sliding on flat polymer surface [17]. The silicon tip on polymer surface serves as an excellent model systems of wear, and is also technologically relevant due to the emergence of scanning-probe-based data storage [41–43] and nanolithography [44]. Gotsmann and Lantz quantify tip wear in situ as a function of sliding distance for applied loads ranging from 5 to 100 nN. It is interesting that, in contrast to similar previous nanotribology works, they study extremely large sliding distances of up to 750 m that is a requirement for the practical implementation of AFM-based data storage and lithography. The abrasive wear mechanism is described as a thermally activated atom-by-atom loss process resulting from a lowering of the barrier to remove an atom due to the stretching of atomic bonds by frictional shear stress. The experimental results confirm the theory of double-well potential applied to activated production of debris, previously developed by D'Acunto and Salmeron [18–21]. Gotsmann and Lantz modify the previous theories in order to fit their experimental results, because they are interested to quantify tip wear rather than the sample surface induced wear. In details, the wear experiments were performed using a homebuilt AFM and compliant (0.15 N/m) silicon cantilevers with tips initially covered with a 2–3 nm native oxide. The AFM is housed in a vacuum chamber to facilitate control of the experimental environment. As a counter surface, a 100-nm film (roughness rms \approx 0.4 nm) of cross-linked polyaryletherketone spun cast on silicon. Such material was chosen for its technical relevance to AFM data-storage applications. Tip-wear experiments were performed by raster scanning a tip over a $30 \times 6 \mu\text{m}^2$ area of the sample in constant-height mode. The load variation was less than 3 nN. The fast-scan axis (mean velocity of 1.5 mm/s) was driven with a sinusoidal waveform to prevent ringing of the scanner. Normal force versus displacement curves were taken every 62 cm of sliding to provide slow feedback control of the load force and to monitor the state of the tip. After every 62 cm of sliding, a reduced scan velocity was used to acquire an image of the area in which the fast sliding had been performed, to monitor the state of the polymer surface. In such experiments, the sample polymer does not show significant signs of wear. Figure 18.13 shows an example of a typical wear test performed with an applied load of 5 nN in a dry environment and a sliding distance of 750 m. After the wear test, a wear volume of $1.5 \times 10^4 \text{ nm}^3$ for the tip wear was determined from pre-experiment and post-experiment scanning electron microscopy (SEM) images (Figs. 18.13a and 18.13b, respectively). Figure 18.13c shows the adhesion data acquired during the experiment that has been converted into radius versus sliding distance using the relation

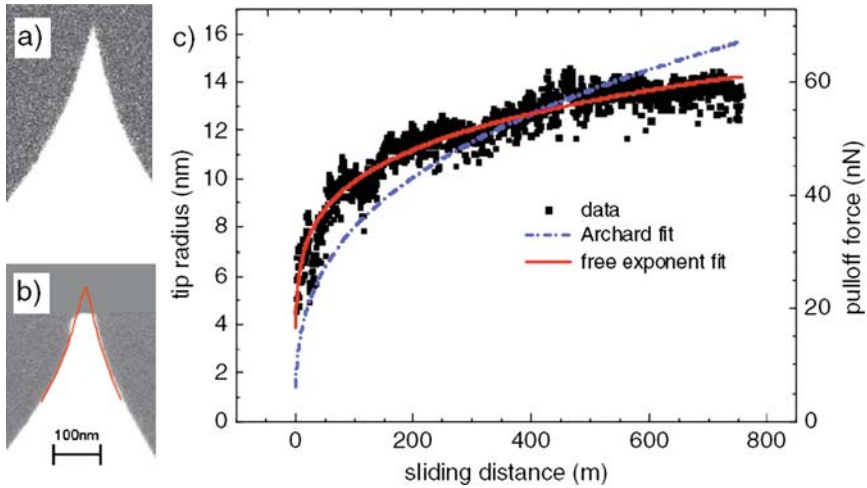


Fig. 18.13. Wear data for a tip sliding on a polymer surface with a load force of 5 nN: (a) SEM image of the tip before and (b) after testing. A contour of the primitive tip is overlaid to visualise the volume loss during the wear test ($1.5 \times 10^4 \text{ nm}^3$). (c) Plot of the adhesion force and contact radius versus sliding distance. Initially, the tips are extremely sharp and well modelled by a cone with a spherical cap having a radius of 3–5 nm. After a few metres of sliding, the spherical cap is worn away and the tip can be modelled as a truncated cone. The pull-off force (de-cohesion force) F_{adh} is proportional to the radius of the flat end a , $F_{\text{adh}} = k_{\text{adh}}a$. For the 11 tips analysed by Gotsmann and Lantz with the final radii between 10 and 50 nm obtained $k_{\text{adh}} = 4.6 \pm 0.9 \text{ N/m}$. The data are fitted using (7) with $m = n = 1/3$, corresponding to the Archard’s wear law, and a fit free m yielding $m = 0.18$. (Reprinted from [17] with permission)

$F_{\text{adh}} = k_{\text{adh}}a$ in a combination with the radius and adhesion force (pull off force) measured at the end of the experiment. Initially, the adhesion increases rapidly, roughly doubling within the first few metres of sliding, followed by a continuous decrease of the rate of change in adhesion. Gotsmann and Lantz use Archard’s wear model to fit the experimental data, obtaining the wear volume V by the relation $V = kF_{\text{N}}d$, where d is the sliding distance, F_{N} is the load force, and k is a constant. Applying such model to a conical tip geometry and solving for the final flattened radius a is [17]

$$a \propto d^m F_{\text{N}}^n, \quad (18.7)$$

with $m = n = 1/3$. The dash-dotted line in Fig. 18.13 is a least-squares fit of the relation to the data illustrating that the Archard’s law is not the best fit for the experimental wear data.

The experimental results indicate that Archard’s law is not applicable to nanoscale wear. The situation is analogous to friction, where macroscopic friction laws fail on the nanoscale [1, 14, 45]. On contrary, the results show

that the wear rates are in general low. For example, for the data in Fig. 18.13, the average wear rate roughly corresponds to the loss of one atom per micron of sliding. Moreover, the wear process proceeds as a smooth process, without indication of evident fractures. Finally, the wear rate increases with increasing applied load and decreases at large sliding distance. First, we can deduce that the wear occurs through an atom-by-atom loss process, which implies the breaking of individual bonds. Such bond breaking can be described by a thermally activated process governed by Arrhenius kinetics, and consequently the height loss rate $\partial h/\partial t$ of the tip is described as [17, 18, 20]

$$\frac{\partial h}{\partial t} = b\omega \exp(-E/k_{\text{B}}T), \quad (18.8)$$

where b and ω are the lattice parameters and attempt frequency, respectively, E is an activation energy, k_{B} is Boltzmann's constant, and T the absolute temperature. In general, the previous Arrhenius-based models describe the wear rate for worn sample surfaces, where the abrasive wear is induced by an AFM tip that remains unchanged by the mechanical contact with the wearing surface [18–21]. On contrary, the experiments made by Gotsmann and Lantz regard the tip wear, as a consequence, some modifications of the previous model is necessary. The first modification to be considered is that during the wear process the tip radius changes and as a consequence also the shear stress, τ , changes. The activation barrier E is reduced though bond stretching by the shear stress τ acting on the bonds, as already predicted by Kopta and Salmeron [19] and D'Acunto [46]. The functional form used by Gotsmann and Lantz for the shear stress is given by the analysis of Briscoe and Evans [47] so that (18.8) can be written now, replacing the height h by the tip radius a , as [17]

$$\frac{\partial a(d)}{\partial d} = \tan(\theta) \frac{\omega b}{v_0} \exp \left[-\frac{E}{k_{\text{B}}T} + \frac{V}{k_{\text{B}}T} \left(\tau_0 + \xi \frac{F_{\text{appl}} + k_{\text{adh}}a}{\pi a^2} \right) \right], \quad (18.9)$$

where the partial derivative is expressed in term of the sliding distance using the relation $v = \partial d/\partial t$, θ is the opening angle of the conical tip, the normal force F_{N} is given by the applied and adhesive forces as $F_{\text{N}} = F_{\text{appl}} + F_{\text{adh}}$, V is an activation volume (an empirical material related parameter without direct correspondence to real space), and τ_0 , ξ , and v_0 are constants. Figure 18.14 shows a numerical solution of (18.9).

In Fig. 18.14, three distinct regimes can be identified in the figure. Initially (i), a small slope is observed in the $\log(a) - \log(d)$ relation. The slope is approximately constant and corresponds to a small exponent m in (18.7). In this regime, the tip is rapidly blunted, and the wear rate depends strongly on load force but only weakly on the constant part of the activation energy $E_{\text{eff}} = E - V\tau_0$. This is because the change in activation energy is largely driven by the pressure under the tip. Region (ii) is particularly important for interpreting the nanoscale wear data. In this region, the exponent m changes from

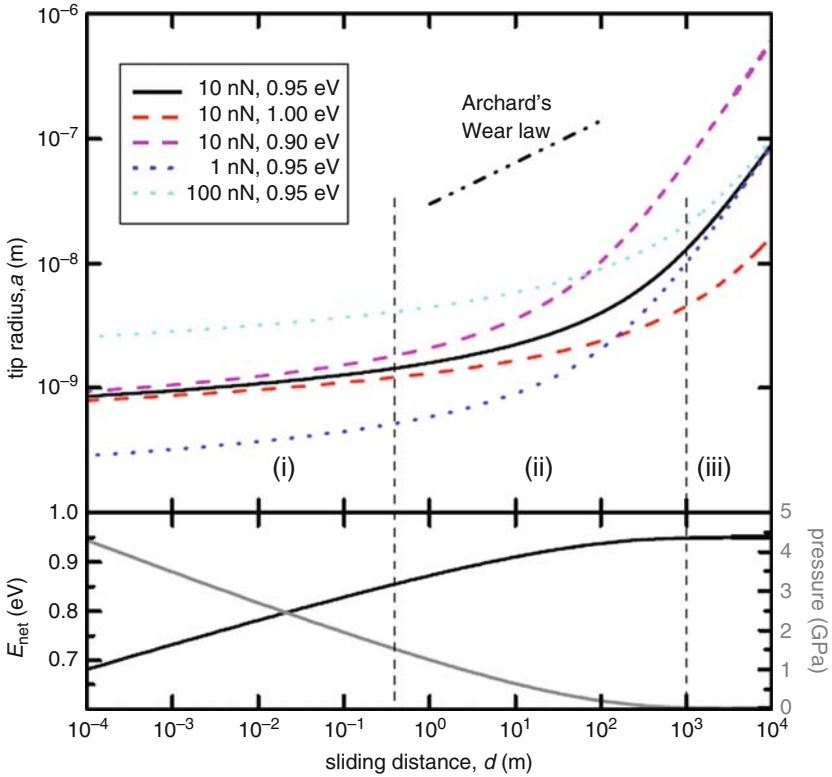


Fig. 18.14. Wear modelled using (9) using $F_N = 10\text{ nN}$, $E_{\text{eff}} = E - V\tau_0 = 0.95\text{ eV}$, and $\xi V = 10^{-29}\text{ m}^3$ (solid lines). The dotted lines correspond to 0.95 eV and load forces 1 and 100 nN . The dashed lines correspond to 10 nN and $E_{\text{eff}} = 0.9$ and 1.0 eV . The slope of Archard’s wear law is plotted in the dash-dotted line. (Reprinted from [17] with permission)

approximately 0 to 1 over several orders of magnitude of d . The change of slope is controlled by a continuous change of pressure on the tip and consequently a change of the net activation barrier $E_{\text{eff}} = E - V\tau$. Thus, for small sliding distance wear tests, a fit using a single exponent m may be possible, and under certain conditions Archard’s wear law may be a good approximation. In the third regime (iii), the wear behaviour becomes independent of pressure and is strongly affected by the effective energy barrier E_{eff} . In this region, the normal pressure is too small to contribute to the wear rate. The tip radius a follows the relationship $a = h \tan(\theta)$ for a perfect cone [17]. The wear data and the appropriate model describing them are a useful tool of the emerging probe technologies, where surely new studies should be addressed in order to clarify all the intriguing and complex aspects connected to wear mechanisms.

One common recognised explanation of the results showed in the various experiments mentioned in the section converges on the attribution that the onset wear is an activated process, where the stripping off atomic or molecular debris is stimulated essentially to the scanning-probe tip induced stresses. The next section 18.3 is devoted to examine thoroughly the activated process when applied to the onset wear problem. In the next subsection 18.2.1, the occurrence of wear for an important class of materials such as viscoelastic materials is described.

18.2.1 Abrasive Wear Mechanisms for Viscoelastic Materials

The use of viscoelastic materials such as polymeric materials constantly increases in the field of nanotechnology. These materials are softer than metallic and inorganic ones, and, because of that, they are easier to wear and deform. Therefore, the wear mechanisms occurring for viscoelastic materials are rather complex, and, generally, present more complications with respect to metals or ceramics materials [1, 2, 48, 49].

The AFM tip-induced wear for the analysis on the onset mechanisms of wear on polymer surfaces consists in the repeated scanning of a prefixed area of the film surface at constant force. Generally, the tip can be moved along the horizontal axis forward and backward along the same line, then moves perpendicular to it and starts the next line. The motivation behind such abrasive tests can be the recognition of the potential that AFM tip-induced wear experiments offer for the characterisation of the molecular organisation of polymer surfaces. For example, it is generally recognised that the glass transition temperatures of polymer surfaces differ from those of the bulk, suggesting that viscous behaviour may be observed at the surface at temperatures where the bulk of the material is glassy. Studies of tip-induced wear behaviour may be valuable in assessing the mechanical properties of polymers surfaces. Moreover, for some classes of thin-film materials, the application of bulk mechanical testing methods is not feasible and other techniques associated with surface mechanical investigation, such as nanoindentation, are inappropriate.

One technique for wear testing consists of using the AFM probing tip to abrade the surface of interest while simultaneously imaging the area where the polymer is being progressively damaged by the scanning tip. Such technique permits the following wear properties to be observed both qualitatively and, when possible, quantitatively: (1) qualitative evolution of the surface during the test, (2) evaluation of the wear volume, for a comparison with macroscale wear tests, and (3) observation of the AFM-probe tip degradation as a consequence of possible adhesion–wear mechanisms for polymer debris. It is commonly found that AFM tip-induced wear tests comport plastic deformations involving the formation of bumps and ripple structures [50–54] (Fig. 18.15).

The wear of the polymer surface caused by the tip is a complex process and it depends on the contact conditions, namely applied force, tip size, and

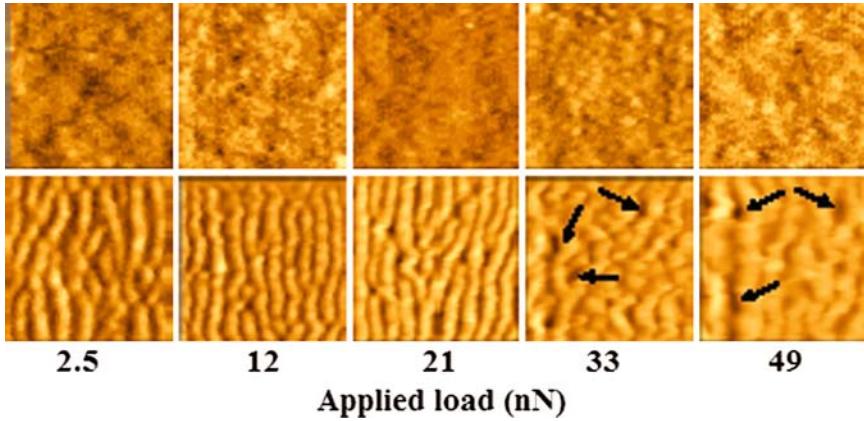


Fig. 18.15. Topographic images for an AFM-tip-induced wear test performed in air (1–40 scanning cycles) with the indicated applied loads on annealed polystyrene [54]. All experiments were conducted on annealed films with a tip nominal radius of 15 nm and scanning directions is horizontal. Each image is $1 \times 1 \mu\text{m}^2$. The topographic variation is 0–2 nm for the *upper row* images (first scanning cycles) and 0–55 nm for the last *bottom row* image (40th scanning cycle). The arrows indicate areas where the film has broken. (Reprinted from [54] with permission)

molecular surface polymer structure. Since the influence of all these parameters is in close connection with the sample properties, one can also expect a dependence of the wearing process on the sample mechanical properties that can vary significantly from the bulk properties, if cross-linking is made or, on contrary, residual solvents is added at the polymer. Recently, Gotsmann et al. focused the attention on controlling the wear characteristic of synthetic polymers for contact mechanical operations in NEMS [55]. They have observed the dependence of wear on segmental relaxation dynamics by tailoring the polymer synthesis with predetermined cross-linking levels. As a consequence, a reduction of nanoscale wear in a polymeric NEMS is possible. Moreover, two wear regimes have been identified: (1) below a critical cross-link density, a ductile wear mode exhibits weak dependence on the spacing between cross-links and (2) above a threshold cross-link density, a brittle wear mode becomes operative and the wear rate decreases rapidly with additional cross-linking (Fig. 18.16). The threshold cross-link density that separates the two wear modes occurs at a critical segment weight of approximately 2,150 Da. This weight corresponds to a critical cross-link spacing of approximately 3 nm, suggesting a competition between the cross-link spacing and the cooperation length for segmental backbone relaxation. When the spacing between cross-links falls below the cooperation length, constraints are imposed on the backbone mobility, essentially stiffening the material. Under these conditions, increased hardness and reduced wear are strongly correlated with cross-link spacing. If the concentration of cross-links is insufficient to interfere with natural backbone relaxation,

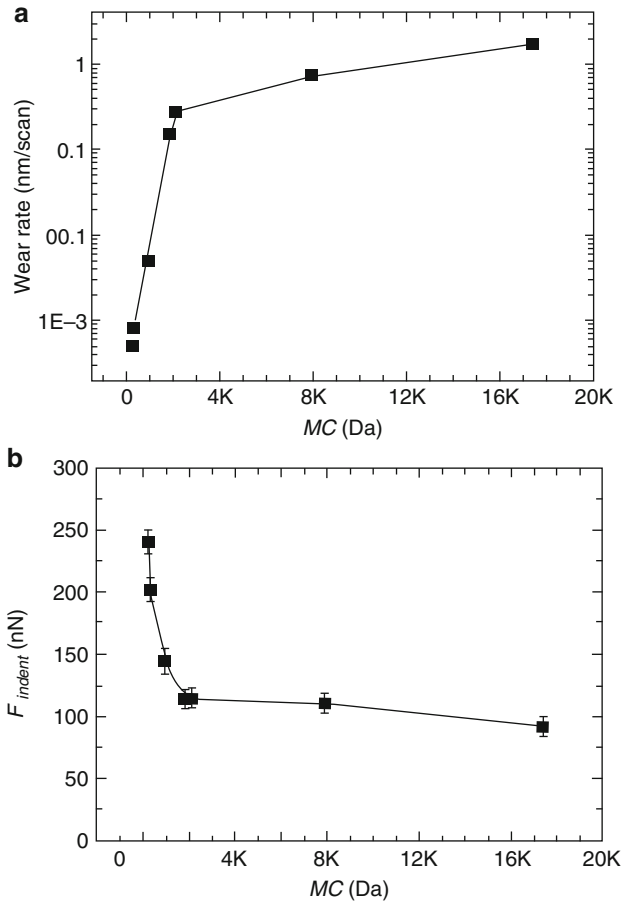


Fig. 18.16. On *right*, the impact of cross-link spacing on wear rate is shown. Two wear regimes are separated by a critical molecular weight between cross-links, $M_{c,crit}$. Below $M_{c,crit} = 2,150 \pm 150$ Da, wear is highly sensitive to the cross-link density, while above 2,150 Da, there is little wear dependence on M_c . The uncertainty level represents the intersecting domain defined by the root mean square variance of regression fits above and below $M_{c,crit}$. On *right*, the impact of cross-link spacing on the hardness of the PSBCB thin films is shown. The minimum force needed to make a permanent 0.5 nm deep indentation, F_{indent} , is plotted versus the molecular weight between cross-links, M_c . (Note: for a given indentation geometry and penetration depth (i.e., equal contact area), the hardness is directly proportional to F_{indent}). (Reprinted from [55] with permission)

the presence of cross-links has little impact on the polymer response and the material behaves similarly to the un-cross-linked native polymer [55].

Section 18.3 is devoted to discuss some models that describe the abrasive- and adhesive-wear mechanisms as an activated process, in a similar way to

chemical reactions, and under some conditions, the wear basic mechanisms can be considered as a chemical reaction between solid surfaces.

18.3 Modelling Wear as an Activated Process

In the preview section 18.2, there have been shown some experimental results linking the abrasive wear mechanisms to the activation rate for the breaking of bonds and consequent stripping off of atomic debris. The onset wear so described seems to be very close to the activation threshold for chemical reactions. Ever since Kramers seminal paper [56], the fluctuational escape over a potential barrier has been a paradigm for a thermal activation process. Kramers understood chemical reactions as the climbing of an energy barrier between two energy wells in a large configuration space, where the first well represents the state of the reactant molecules and the second well that of the product molecules. Activated escape underlies diffusion in crystals, protein folding, nucleation rate in phase transitions and provides a paradigm for activated chemical reactions. In a wear experiment, we can be mainly interested to quantify the amount of mass stripped off from a surface in a relative motion with another generally harder surface. As a consequence of the formed wearing debris, one possibility is to describe the change of spatial location for the defects produced during the damaging as described in Fig. 18.17.

In an activated process, roughly we have a system of interest at thermal equilibrium connected to a driving force responsible for the activation of a process from the system of interest. The mechanism responsible can be readily understood for adiabatically slow driving, where the driving frequency is small compared to the relaxation rate in the absence of fluctuations and the system of interest remains in quasi-equilibrium. For systems in thermal equilibrium, the fluctuation probabilities are given by the activation law, $W \propto \exp(-E/k_B T)$, where E is the activation energy of escape that is tuned by driven force, and even where the small fluctuation of amplitude $|\delta E|$ is small compared to E , it may still substantially exceed $k_B T$, in which case W will be changed very strongly. In our picture, W represents the probability that a portion of mass (debris) can be detached by one surface when it is in relative motion with respect to other surface (Fig. 18.18). The formation of wear debris is a random process and stochastic picture should be the most adequate tool to approach quantitative results. A stochastic process is the correspondence between a random experiment and a group of functions; as a consequence, the system can be predicted in terms of probability as for example the velocity of a Brownian particle [57, 58]. Effectively, the connection between Brownian motion and debris can be made only in terms of products of the wear process, when a debris is formed being detached from the native surface and its mechanical features (diffusion, adsorption, etc.) are subject to environmental fluctuations.

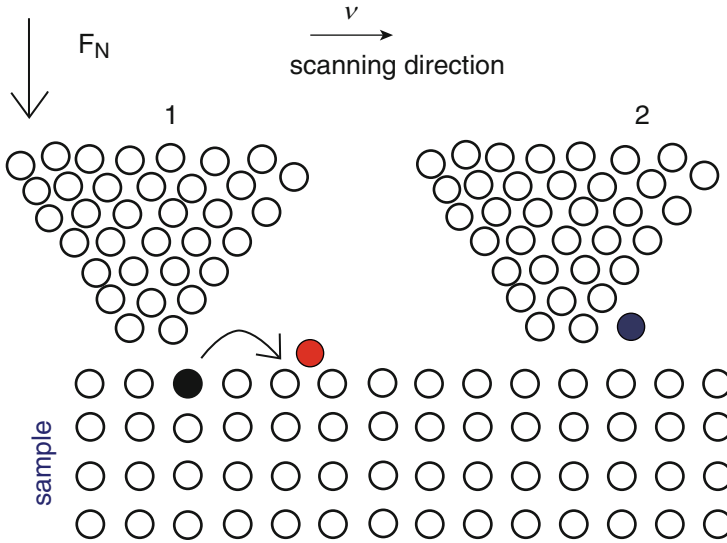


Fig. 18.17. Schematic sketch on the production of debris (defects) due to abrasion wear-mechanism or adhesion wear-mechanism. The scanning tip in contact with the sample produces a damage on the regular surface moving the black atom, relocating in a new position (red atom), or adsorbed by the scanning-probe tip (blue atom). The AFM-induced activated possible detachment of the black atom should be described by a conditional probability $P_1(x, t_1/x_0, t_0)$, where at time t the atomic debris is in the x position after the tip stress contact is made starting from x_0 position before tip contact, and the possibility of relocation on the native surface or adhesion to the stressing tip after the contact should be given by another conditional probability, $P_2(y, t_2/x, t_1)$. The whole conditional probability is the convolution between these two probability functions

Two surfaces in contact, such as the AFM tip and the scanned sample, share a common surface for a certain time. Each exerts an equal and opposite force on the other determined by the pressure it exerts on every elements of the surface of separation as required by the physics of an open systems [59]. Like general approach to open systems, we can consider a sample of interest in contact with external or environmental forces. In the tip-sample contact, the external forces are represented by the tip-induced stresses. One practical assumption can be made in order to simplify the tip-sample systems looking at the different time scale for the motion of the AFM tip and the atomic or molecular components of the sample investigated. The AFM has an effective cantilever mass around 10^{-11} kg, many orders of magnitude larger than the atomic mass unit ($\sim 10^{-27}$ kg). The resonant frequency of the cantilever is around 10^5 Hz, while the atomic vibration frequency is of the order of 10^{13} Hz. Practically, in the contact area we have two kinds of dynamics – the slow dynamics of the cantilever and the fast vibrations of atoms. At the time scale

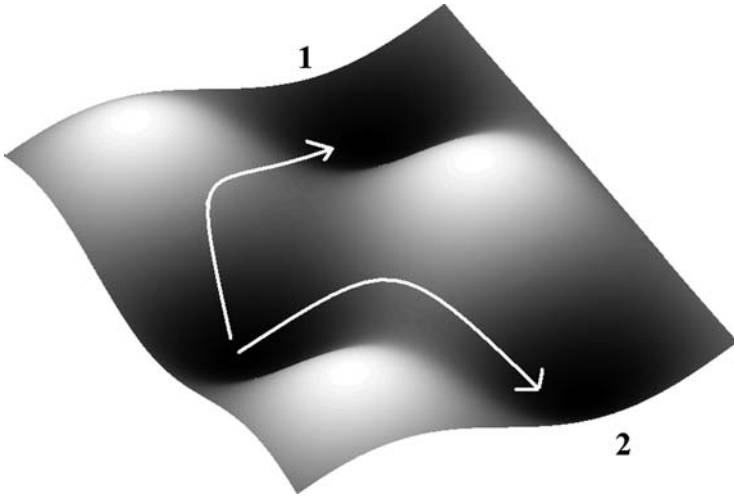


Fig. 18.18. Schematic sketch to describe the well potential transition applied to describe the possible location of a debris stripped off from the native surface during sliding motion between two interacting surfaces, as in Fig. 18.17. For molecular scale contact, the debris has molecular dimensions, and the different final locations of debris are connected to the specific wear-mechanisms. As a consequence, the final wells 1 and 2 simulate the two possible different surfaces [in analogy to Fig. 18.17, (18.1) and (18.2) wells represent the re-positioning on sample surface or tip surface] and the different wear mechanisms

of picoseconds, we encounter thermal vibrations of atoms, while the cantilever is stationary. On contrary, at the time scale of microsecond, we observe the scanning vibrations of the cantilever in contact with a cluster of atoms. If the relaxation time of this atomic cluster is much smaller than the cantilever time scale, then we will always observe an equilibrium cluster of atoms that follows the slow motion of the cantilever. Therefore, the energy modes between the cantilever and the atomic substrate can be decoupled. In molecular dynamics simulations, integration of dynamic equations of motion for the cantilever and the atomic substrate can be performed separately. As a consequence, we can limit to study the possible escape process for the sample atoms considering the effect of the AFM tip in terms of induced stresses lowering the energy activation for the surface atoms and considering the tip rigid.

For example, let us consider a general situation where the fundamental wear-mechanisms are abrasive- and adhesive-wear mechanisms. The quantification of wear due to adhesive or abrasive forces has been calculated in terms of rate of transition between two subsequent well potential. The problem can be reduced to a bi-dimensional question. Along the z -axis, the vertical forces are defined by a van der Waals term leading to the adhesion mechanism. Shear forces along the x -axis, that is, the direction of motion of the tip leading to the abrasive mechanism for the atomic debris. The occurrence of one or both

the mechanisms depends on the intensity of the forces along the two dimensions. Such situation can be resumed by the sketch in Fig. 18.18, where the simultaneous double well for the two wear processes have been shown. In both the cases, the wear volume rate could be calculated within the microscopic diffusive processes that can be expressed by the following relation [18, 20, 21]

$$V = n_{\text{tot}} w \Theta, \quad (18.10)$$

where n_{tot} is the total number of possible debris lying on the primitive sample surface, w is the rate of transition associated to the activated process, and Θ is the mean debris volume. In experiments performed on ionic crystal, the mean debris volume can be identified in ionic pairs, so the wear volume rate can be reduced in the calculation of wear rate for activated de-bonding under applied tip stresses. Equation (18.10) should be considered to be equivalent of macroscopic wear volume (18.1) and (18.2) on nanoscale. In a simplified vision, wear contributions due to adhesive- and abrasive-wear mechanisms can be considered separated and summarised as $V = V_{\text{abr}} + V_{\text{adh}} = (n_{\text{abr}} + n_{\text{adh}})\Theta$, where V_{abr} and V_{adh} are the abrasive- and adhesive-wear volumes, respectively, and n_{abr} , n_{adh} are the number of atoms involved in the two distinct wear mechanisms. In reality, the overlap of the two wear processes can be easily supposed. The experimental results performed on ionic compound in UHV by Gnecco et al. [38] clearly show that $n_{\text{abr}} \gg n_{\text{adh}}$ to be valid, that is, the abrasive-wear mechanism is stronger compared to the activation of adhesive-wear mechanism. The advantage of modelling the interface between two surfaces by a double-well potential is that the wear volume can be quantified essentially in terms of rate of transition between the well minima. Moreover, the rate and (as a consequence) the wear volume depend on the potential barrier height, and the potential barrier height would be connected to experimental and material quantities such as hardness, elasticity, load applied, and time (or sliding velocity). Obviously, the wear volume will be a function of the energy barrier. Essentially, vertical forces lead to the adhesion-wear mechanism (transition from surface sample to tip), and shear forces lead to abrasion-wear mechanism (transition from surface sample to surface sample). Independently by the wear mechanisms activated, the rate of transition in (18.10) is given anyway by the Kramers picture [see also (18.8) and (18.9)]

$$w_{\text{abr,adh}} \approx e^{-\Delta E_{x,z}/k_{\text{B}}T}, \quad (18.11)$$

where the index is referred to the x -dimension for abrasion and z -dimension for adhesion. To find the force of interaction between an AFM tip and a flat sample, the interatomic potential has to be summed over the whole number of atoms involved in the contact between sample and tip. The accurate derivation of such rate of transition for a double-well potential model was derived in [21].

The probabilistic approach to the production of wear debris can be performed within the classical double-well potential with time-fluctuating barrier. The barrier is fluctuating because during the contact between the tip and the

sample, we have two different time scales. Here, we limit to comment the wear rate obtained as a function of the frictional state in the initial well potential. Each crossing particle is defined by a frictional term, proportional to $\gamma\dot{x}$, where \dot{x} is the effective velocity of the particle living in the initial well and γ is the damping factor. The physical nature of such frictional term can be explained in terms of irreversible transferring of energy to the environment, other particles or phonon states. Nevertheless, there is another irreversible process, which costs energy leading to a non-conservative lateral force during the contact between the probe tip and the sample. This supplement of friction is manifested during the transition between well potentials. The total amount of lost energy for unity of the tip residence time for the wearing sample can be quantified as follows:

$$E_{\text{loss}} = n_A w \gamma \dot{x} L, \quad (18.12)$$

where L is the spatial distance between the minima of the well potential, n_A described the atomic density on the native surface (initial well potential), and $n_A w$ is the debris formed during the contact with the tip (it has the dimension of a mass for unity of residence time of the probe tip during the contact). The frictional state of particles in the initial well modifies the wear rate. In this sense, we can define a correlation between molecular friction and adhesive wear. Considering high internal viscosity, that is, when the cluster velocity relaxes much more rapidly than the position, constant wear rate is

$$w(t) = \frac{\Omega^2}{2\pi\gamma} \exp\left(-\frac{\Delta U(t)}{k_B T}\right), \quad (18.13)$$

where Ω is an instantaneously effective oscillation frequency and $\Delta U(t)$ is the energy barrier between the two wells as a function of the time of the local contact of the tip and the sample surface. As can be expected, the wear rate depends on an Arrhenius factor and it is interesting to note that the rate does not depend on the shape of the potential barrier, but only on the difference of the top barrier. The total number of particles passing the barrier in one second is equal to $n_A w$. This is the case of high viscosity. It is also possible to treat in a similar fashion the limiting case of low viscosity with result obtaining the expression [58]

$$w(t) = \gamma \frac{\Delta U(t)}{k_B T} \exp\left(-\frac{\Delta U(t)}{k_B T}\right), \quad (18.14)$$

where now, the wear rate is directly proportional to the friction term γ . Equations (18.13) and (18.14) can be slightly modified if we consider additive fluctuating terms to potential barriers due to noise of other nature with respect to the thermal noise due to the complexity of dynamic interaction between tip and sample in a sliding process, where anomalous diffusion processes take place [20]. The presence of such fluctuations in the potential barrier is important since it may be hypothesised that in certain catalytic reactions, the catalyst–environment interaction could give rise to effects that can be

represented with an effective potential that contains a supply noise coupled with position variable [60, 61].

The double-well potential as described earlier can be used in special way for describing some features of the surface wear. In surface wear, the rearranging and re-crystallisation of removed atoms may be observed [38]. Here, we like to focus the attention on two aspects. The first one is that the induced AFM tip wear mechanisms can be due to the combined effects of the sliding tip and interatomic forces and can be summarised in terms of shearing forces and vertical forces. The second relevant aspect is that the debris can diffuse on the surface growing in regular structures as shown in Fig. 18.11.

Now, we focus the attention on the problem of the quantification of the effects due to shear forces and normal forces. The first ones could be considered responsible for the detachment of the ionic atoms from its crystal site and subsequent dragging until a new equilibrium crystal site is reached. The new position is achieved when the equilibrium is approached; this new position is a step on the flat atomic surface (Fig. 18.17). The ionic crystal shows that the process involves pair-by-pair ionic atoms. Vertical forces are due to the interaction of the tip probe approaching the surface sample. This interaction is led by a van der Waals force and stimulates the ionic atoms to follow the tip adhering to it (Fig. 18.17). In terms of effective wear rendering, the adhesion causes a vacancy along the crystal surface. Following (18.10), the abrasive- and adhesive-wear mechanisms can be quantified considering the rate of transition as given by Kramers picture for the ionic atoms in a two-dimensional well potential (Fig. 18.18). The corresponding height of energy barriers determines the incidence of wear both for abrasion and for adhesive case. In the case of shear forces, the energy barrier can be deduced by the surface potential U_0 , and a reasonable energy barrier along the x -direction $\Delta E_x \approx U_0$. At the minimum value of the barrier, the vertical force is given essentially by the energy binding of the attractive term of the van der Waals force. The apex tip is considered a cylinder of radius $R \cong 10$ nm, the potential barrier can be expressed like $\Delta E_z \approx AR^{1/2}r_0/z^{3/2}$, where z is the distance of separation between the surface and the tip and A is the Hamaker constant [29]. The distance of separation, z , is directly proportional to the applied load. After the minimum value of the barrier is reached, corresponding to the minimum position z_{\min} , the first term of a Lennard-Jones potential is not negligible. Moreover, the temperature should increase the rate lowering the well potential. Although the effect of temperature could engrave in a different way on the shear forces compared to the vertical forces. Nevertheless, the role of thermal fluctuations should be investigated more appropriately. A numerical simulation has been performed on a volume of $120 \times 120 \times 1.8$ nm³; the density of atoms involved in the wear processes is given by the summation on the first three layers of the quantity $n_L \exp[-(i-1)]$, where n_L is the number of atoms for layer, and $i = 1, 2, 3$ denotes the layers [21]. The exponential correction is introduced because the atoms lying on the inner layer are less involved in the possible wear mechanism compared to the first

layer one. The lattice constant is considered to be $a = 0.6$ nm, any layer contains 200 atoms, and the total number of atoms could be involved in the wear process for the volume considered nearly 6×10^4 . The calculation for an applied load of $F_N = 0.65$ nN is performed at room temperature $T = 298^\circ\text{K}$, the thermal energy is $k_B T = 0.041 \times 10^{-19}$ J, the lattice potential barrier is $U_0 = 0.432 \times 10^{-19}$ J, and the Hamaker constant is $A = 10^{-19}$ J. Inserting these values in the expression of the barriers ΔE_x and ΔE_z , we obtain that $\Delta E_x/k_B T \cong 19$, while $\Delta E_z/k_B T$ is nearly 15 times $\Delta E_x/k_B T$. Because the functional dependence of the rates by the ratio between the energy barrier and the thermal energy is $x \exp(-x)$, the second term corresponding to vertical transitions toward the tip is practically negligible. The rate of atoms passing a potential barrier and undergoing a wear mechanism is given dividing the wear volume by the atomic volume [21]:

$$V/\Theta = nw_{\text{abr}} \approx n \left(\frac{2\pi/T_0 T_1}{\beta} \exp\left(-\frac{U_0}{k_B T}\right) \right), \quad (18.15)$$

where T_0 and T_1 are the instantaneous effective oscillation potential of well periods. For a load of $F_N = 0.65$ nN and a cycle of 1,000 scans, (18.15) gives nearly 2,400 atoms involved in the abrasion-wear mechanism, that is, the 3.9% of the total number of atoms, while the number of atoms undergoing an adhesive-type process is practically irrelevant. It is remarkable that this model describes in a direct manner that the incidence of abrasive-type mechanism for wear is the most prominent mechanism for wear on nanoscale for low applied loads.

At low loads, it can be interesting to know the effects of the tip on the atoms, which are not able to pass the energy barrier. These atoms do not form the material loss but can be subject to a lattice modification, like dislocations, and consequent formation of gradient stress.

The second relevant aspect, that is, the rearrangement of structures due to the surface wear debris as the ripples in Fig. 18.11, is still an open problem, but recently, a solution is going to be proposed [62]. Here we limit ourselves to show some results obtained for a theoretical model introduced to describe the experimental results [39] (Fig. 18.19).

18.3.1 Self-assembled Monolayers as a Frame for Modelling Wear in Viscoelastic Materials

The increasing use of viscoelastic materials requires that the basic wear mechanisms should be deeply analysed and, because of the complex structure of such materials, simplified models are needed. Self-assembled monolayers (SAMs) are very useful for the systematic modification of the physical, chemical, and structural properties of a surface by varying chain length, tail group, and composition. Many of these properties can be studied making use of the AFM, and the interaction between AFM-probe tip and SAMs can be considered also

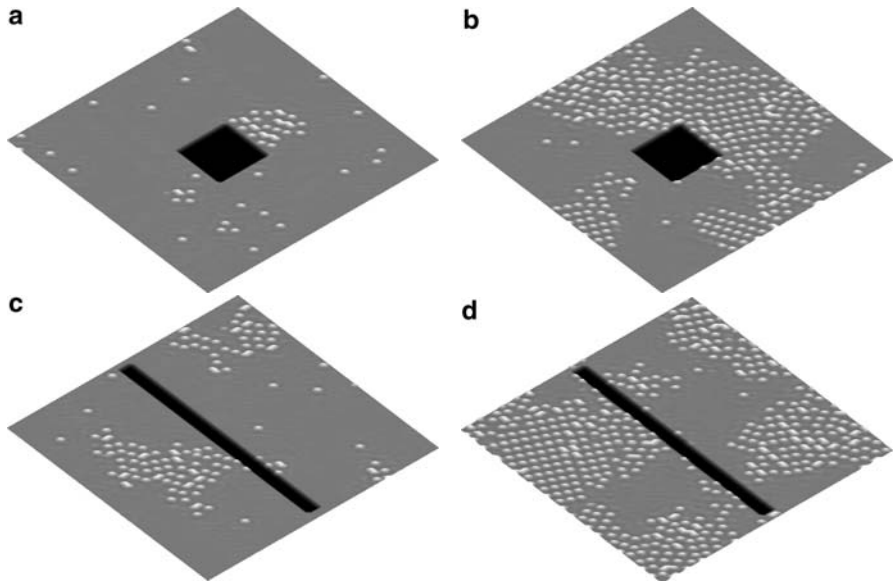


Fig. 18.19. Simulated surfaces where a scratching wear test has been performed. The black areas are the scratched areas (squared or linear), and the white dots are the cluster nucleated formed by the adatoms AFM induced during the scratching [62]. The growing clusters are a product of the re-crystallisation of the atomic AFM wear induced debris, different patterned surface structures are possible

to be an excellent reference to study fundamental properties of dissipation phenomena and onset wear for viscoelastic materials on the nanoscale.

In this section, we deal with a numerical study showing that the fundamental mechanism for the onset wear in an SAM is a process of nucleation of domains starting by initial defects [46]. An SAM surface repeatedly sheared by an AFM-probe tip with enough applied loads shows the formation of progressive damages enucleating in domains. The AFM-induced surface damages involve primarily the formation of radicals from the carbon chain backbones, but also the deformations of the chains resulting in changes of period lattice have to be taken into consideration. The nucleation of the wear domains starts generally by initial surface defects where the energy cohesion between chains is lower. Moreover, the presence of surface defects is consistent with changes in lateral force increasing the probability of the activation for the removal of carbon debris from the chain backbone. The quantification of progressive worn area can be performed making use of the Kolmogorov-Johnson-Mehl-Avrami (KJMA) theory for phase transition kinetic processes [63, 64].

When the AFM-probe tip slides on a deformable viscoelastic surface, relaxation of inner polymer structure leads to inevitable contact damping, which makes surface forces non-conservative. Since the pioneering paper of Nuzzo and Allara [65, 66], the fabrication of nanoscale structures using SAMs has

attracted attention due to their scientific importance and potential applications [67–69]. SAM and Langmuir-Blodgett (LB) techniques are the most common methods for forming monomolecular films on solid substrates for use as model lubricants [70–72]. In the case of LB films, the molecular interaction with the substrate is through weak van der Waals forces, and repeated shearing results in rapid wear of the molecules. On contrary, SAMs are ordered molecular assemblies formed by the chemical adsorption of an active surfactant on a solid surface and are, therefore, better candidates as lubricants. SAMs can be formed by immersion of an appropriate substrate into a solution of an active surfactant in an organic solvent. SAMs consist of three building groups: a head group that binds strongly to a substrate, a tail group that constitutes the outer surface of the film and a spacer chain (backbone chain) that connects the head and tail groups (Fig. 18.20). As for an example, we focus the attention on alkanethiol monolayers on Au(111). It is now well established that complete alkanethiol monolayers formed on Au(111) adopt a commensurate $(\sqrt{3} \times \sqrt{3})R30^\circ$ structure, with a nearest neighbour spacing of 0.497 nm, which corresponds to a surface area of 0.2 nm^2 per chain [74]. The self-assembly mechanism of alkanethiol C_n , $[-\text{SH}(\text{CH}_2)_{n-1}\text{CH}_3]$, molecules on Au(111) has also been studied using several techniques including SFM techniques.

It is commonly recognised that during the sliding of a tip onto a SAMs-coated substrate, there are two main factors that create wear: (a) the rupture of SAMs/substrate bonds and (b) the rupture of SAM chain carbon–carbon

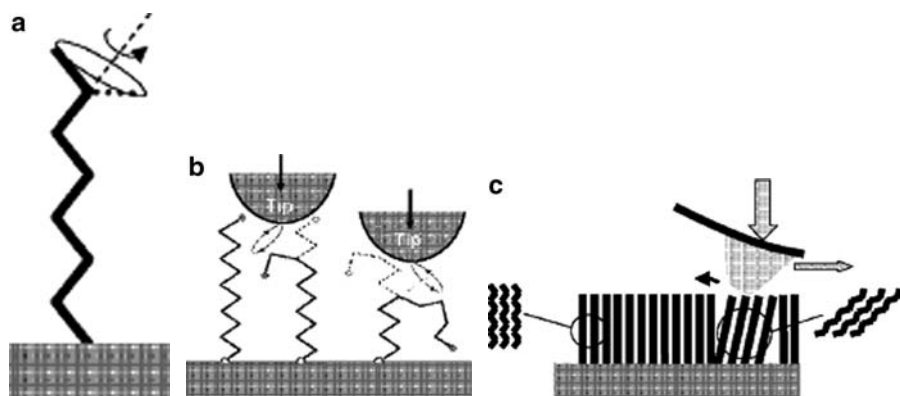


Fig. 18.20. Schematic representation of three types of viscoelastic deformations of alkane molecules proposed by Salmeron [73]. (a) Terminal gauche distortions, (b) internal gauche distortions, which require displacement of the neighbouring molecules, and (c) rigid molecular tilts. The requirement of displacing neighbouring molecules is less stringent in this case due to the concerted motion. (Reprinted from [73], with permission)

bonds. These two mechanisms can be catalysed by the disorder induced by the amplification of primitive defects [75].

AFM sliding friction experiments on C_n -alkanethiol SAMs showed a friction behaviour dependent on chain length [76]. The coefficient of friction decreased as the chain length was increased and the low friction of methylic group terminated SAMs is due largely to low surface energy coming from their hydrophobic tail groups and high load-bearing capacity due to their compact structure minimising van der Waals interaction energy.

Onset wear tests can be simulated on an SAM surface taking a typical AFM test consisting in many scanning cycles on the same area and acquiring data of such progressive damaged area scan after scan. In the simulated approach, the percentage of the worn area can be studied as a function of several parameters, such as applied loads, lateral friction force and chain lengths, and curves are plotted indicating evolution of progressive damages (Figs. 18.21 and 18.22). Consequently, the progressive damaging of the surface due to the continuous passage of the AFM tip can be monitored scan by scan. Because of the SAM lattice structure, this implies the possibility of a two-dimensional quantification for the wear damage rather than classical macroscopic volume test quantification. As a consequence, we can quantify the progressive wear formation on the SAM surface observing how the primitive regular lattice area is progressively changed and reduced during the wear test. When the applied load is not low enough, deformations of the SAM structure or defects or formation of free radicals due to applied stress lead to changes in the two-dimensional regular structure. The connection between dissipation mechanisms and wear is essentially given by the fact that initial surface defects correspond to the points where the lateral force and exchange of energy between tip and sample are higher. It is well known that alkanethiol monolayers on Au(111) exhibit a distribution of pit-like defects [77]. These pit-like defects should be a consequence of already existing defects of underlying Au surface. Other defects can be formed during alkanethiol monolayer assembly. They diminish locally the physical and chemical properties of alkanethiol chains, other than changing the height profiles. During the AFM-induced compression, the molecular spring assembly can experience orientation and compression under normal load [73], so the SAM chains experiment further deformations that can be considered in the class of defects. Generally, three types of AFM-induced viscoelastic defects can be considered (Fig. 18.20). The first type involves the creation of gauge distortions at the free end groups. This results in a 120° rotation of the methyl group around the C–C backbone. The height shortening of the chain is expected by about 0.1 nm, and the energy activation for such rotation is around 0.1×10^{-19} J. The second type of AFM-induced defect is internal gauge deformations, which require displacement of the neighbouring molecules. The third type is given by the rigid chain tilts. These two last types of defects are more difficult to produce, because their energy activation is considerably higher. Nevertheless, in the surrounding of structural initial defects such as pit-like defects, the alteration

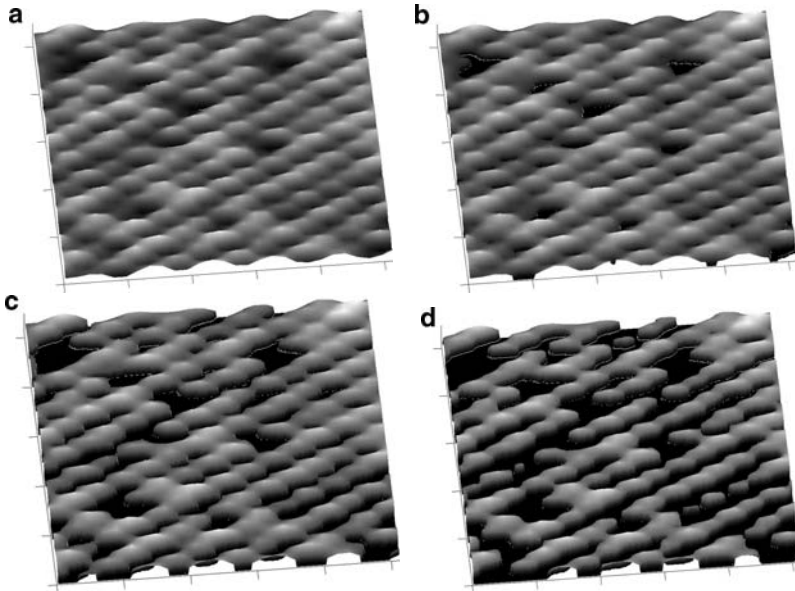


Fig. 18.21. C_{18} -chain progressive damage of simulate methyl group end surface. The initial image (*left upper*) is based on an AFM experimental acquired data on a C_{18} sample, the typical commensurate $c(4 \times 2)$ superlattice of $(\sqrt{3} \times \sqrt{3})R30^\circ$ arrangement is slightly altered by the applied load. The final image (*right down*) shows that nearly 20% of methyl groups have been stripped off by the AFM tip after 100 scans with an applied load of 5 nN. The black portions represent the wear domains. (Reprinted from [46] with permission)

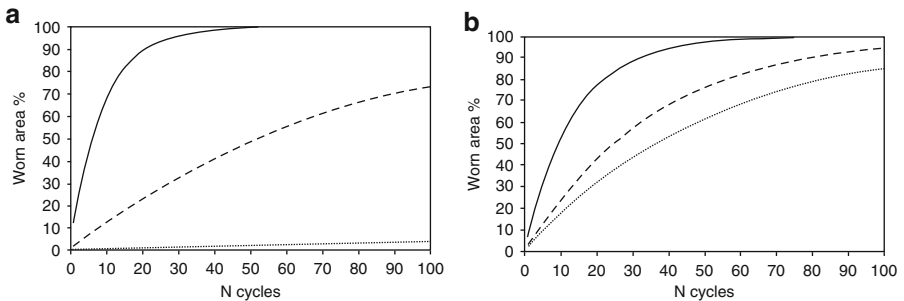


Fig. 18.22. On *left*, worn area as a function of the number of scans for the C_8 -alkanethiol chain at different applied loads (10 nN, *continuous line*; 5 nN, *dashed line*; 2 nN, *dotted line*). It is interesting that just after only one scan, the worn area is about 12% with an applied load of 10 nN. On *right*, worn area as a function of number of scans for a different number of carbon atoms on the backbone chains (C_8 , *continuous line*; C_{15} , *dashed line*; C_{18} , *dotted line*) is displayed, for an applied load of 10 nN. (Reprinted from [46], with permission)

of the molecular packing density can lower their energy activation. Barrena et al. observed that the heights of self-assembled alkylsilanes decrease in discrete amounts with normal load [71]. This step-like behaviour is due to the discrete molecular tilts, which are dictated by the geometrical requirements of the close packing of molecules. Only certain angles are allowed due to the zigzag arrangement of the carbon atoms.

It is reasonable to suppose that the energy exchanged between the probe tip and a SAM sample involves the initial defects points for nucleation activation for larger defects when the scanning-probe tip slides more and more. The idea is that at the edge of a defect, the energy cohesion due to van der Waals interaction and responsible of chain packaging should be lowered by the extra-space available due to the presence of defects. The van der Waals energy cohesion ε_{vdW} is nearly $0.6\text{--}1.3 \times 10^{-20}$ J, and the van der Waals radius of the methyl groups is approximately 0.2 nm in a SAM packed structure [73]. The bond energy of C–C is nearly $1\text{--}2 \times 10^{-19}$ J, and S–C and S–Au binding energy fall in the range of $2\text{--}4 \times 10^{-19}$ J [78]. Under opportune applied loads, the shearing stress induced by the passages of the probe tip should provoke the enlargement of the defects area provoking modifications of period lattice with subsequent changes of orientation of the chains and, over threshold loads, the breaking of covalent bonds with the formation of holes. These holes are due to the stripping off of radicals, such as $\text{CH}_3\text{--}$, macroradicals such as $(\text{CH}_2)_x\text{--CH}_3\text{--}$, or simply H-; here subscript x denotes the fraction stripped off from the original chain backbone. But, it is reasonable to suppose that the AFM-induced de-bonding regards essentially C–C connections. Any carbon end group can be considered in a well potential and the passage of probe tip should make possible to overcome this potential barrier with the probability of transition enhanced by the stress applied (Fig. 18.18). Generally, the formation of active debris should involve the adhesive attachment of such SAM debris to the sliding-probe tip changing the shape of the probe tip (adhesion tip wear) and as a consequence the real area of contact. Adhesion tip wear due to the attachment of progressive formed debris to the probe tip could be described satisfactorily by a double-well potential, where any well potential corresponds to one surface in dynamic contact with other one, as described in the preview section 18.2 on activated process [18, 20, 21]. Although the adhesion tip wear is a common occurring wear mechanism, we will not include its description, limiting ourselves to study general condition for progressive damaging of SAM surface without considering the dynamics of SAM debris when detached by the carbon backbone.

It is reasonable to suppose that in an AFM abrasive test consisting of a certain number of scans, the density of defects should increase before gross wear will take place. This process is essentially a process of nucleation of defects forming holes increasing in time, where the time is the product of the residence time of the probe tip for the number of scanning cycles. Nucleation is an activated process, such as its rate has an exponential dependence on the free-energy barrier to nucleation defects. If at the i th site of the SAM surface

the free energy barrier is ΔU_i , and the attempts frequency is v_i ($\sim 10^{13}$ Hz), then the rate of nucleation at the i th site is [46]

$$\Gamma_i \approx v_i \exp\{-[\varepsilon_{0,i} - \delta\varepsilon_i(F_N, F_1)]/k_B T\}, \quad (18.16)$$

where $\Delta U_i = \varepsilon_{0,i} - \delta\varepsilon_i(F_N, F_1)$ is expressed by the mean value ε_0 of the C–C bond and the stress-induced deviation represented by the term $\delta\varepsilon(F_N, F_1)$ that is a function of the applied load F_N and lateral force F_1 , k_B is the Boltzmann's constant, and T is the temperature. The stress-induced deviation is a combined effect of compression and shear stresses and it lowers the energy barrier making possible the carbon stripping off from the chain backbone. The lateral stress is generally higher corresponding to defects with respect to regular lattices; as a consequence, the rate of nucleation must be calculated locally and cannot be considered homogeneous. The number of defects produced in the absence of stresses is rather small; under the tip stress the maximum possible rate of debris production, $\Delta N_i/\Delta t$ at the i th point contact should be given related to the thermal activated detachment [79]:

$$\frac{\Delta N_i}{\Delta t} \approx \frac{R^2}{a^2} \Gamma_i, \quad (18.17)$$

where $\Delta t = (A_r/A_{\text{image}})t_{\text{image}}$ is the residence time of the probe tip on the sample, A_r is the real area of contact, A_{image} is the area to be scanned, and t_{image} is the image required for an image [19,46], R is the radius of the real area of contact between the tip and the SAM, a is the van der Waals average radius of single methyl groups, and Γ_i is the rate of activation for the creation of AFM-induced defects. Here, we assume that the continuum mechanics could give a good approximate value for the real area of contact and the depth of penetration even if the contact is discrete [80]. For a paraboloid shape of the probe tip in non-adhesive contact with a flat surface, one finds that the real area of contact and penetration depth are, respectively [35]

$$A_r = \pi \left(\frac{3F_N \rho}{4E^*} \right)^{2/3} \quad \text{and} \quad z_d = \frac{1}{2} \left(\frac{3F_N}{4\rho^{1/2}E^*} \right)^{2/3}, \quad (18.18)$$

where $E^* = [(1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2]^{-1}$ is the reduced modulus, $E_{1,2}$ are the Young moduli and $\nu_{1,2}$ the Poisson's ratios, F_N is the applied load, and $\rho (= 1 \text{ nm})$ is the tip radius. For typical moduli of Young and an applied load of 2 nN, we obtain a value for the reduced modulus of nearly 29.5 GPa [80]; as a consequence the real area is approximately 0.4 nm^2 . A single C_n -alkanethiol head group formed by $-\text{CH}_3$ has an area of 0.2 nm^2 ; this means that approximately two final head groups are simultaneously in contact with the probe tip and support the applied load. It should be noted that differently to friction behaviour, SAM films show chain-length independence of elastic modulus [80].

We suppose that the decrease of activation energy $\delta\varepsilon(F_N, F_1)$ in (18.16) has two major contributions depending on the normal stress or shearing stress.

The first contribution, due essentially to the normal stress, $\delta\varepsilon_1$ (divided by the C–C bond energy, ε_{C-C}) can be estimated by multiplying the force acting on the methyl group carbon atoms involved simultaneously in contact with the tip by the resulting bond stretch

$$\frac{\delta\varepsilon_1}{\varepsilon_{C-C}} \approx \frac{a^2}{R^2} \frac{z_d}{(A_r/\pi)^{1/2}}, \quad (18.19)$$

where z_d is the penetration depth, A_r is the real area of contact, R is the radius of the real area of contact between the tip and the SAM, and a is the van der Waals average radius of single methyl groups. The relation (18.19) can be justified by the assumption that the applied load-induced bond stretch should be on the same order as the Hertzian penetration depth z_d , divided by the characteristic length of the contact area, which is on the order of the contact radius. The ratio between the areas of contact and the van der Waals region of single methyl group takes in account the number of methyl groups supporting the tip–SAM contact. The relation (18.19) shows that the contribution of applied load stress to decrease in activation energy changes as $F_N^{1/3}$. In the second term of the decrease of activation energy, the shearing stress plays an important role and should contribute for a lowering factor as [46]

$$\frac{\delta\varepsilon_2}{\varepsilon_{C-C}} \approx \frac{d_{C-C}}{(A_l/\pi)^{1/2}} \frac{z_d \times \text{sticking force}}{\varepsilon_{vdW}}, \quad (18.20)$$

where $d_{C-C} = 0.153$ nm is the C–C bond length along chain axis, A_l is the lateral real area calculated at the edge of penetration depth, ε_{vdW} is the van der Waals contribution to the energy of cohesion between chains, and sticking force is the maximum value of the lateral force in sticking phase during the typical stick-slip motion of the probe tip sliding on the SAM surface. The existence of a penetration depth implies that the carbon atoms set at the edge of the penetration depth should support entirely the shearing stress provoked by the coming forward probe tip along the sliding direction [81]. The cohesion energy ε_{vdW} between chains involves a resistance to the shearing inducing debris production on the edge of the penetration depth as the tip slides. An advantage of a probe-scanning based wear test is that the state of a progressively worn surface can be monitored scan after scan. The time evolution of such worn areas progressively damaged can be visualised as in a real AFM experiment. For example, Fig. 18.21 proves the progressive AFM-induced damaging for a C_{18} surface for an applied load of 5 nN.

The growing number of AFM-induced surface defects and the total worn area at the end of the wear tests can be deduced by the classical nucleation theory [46]. If we define $\rho(x, t)dt$ the probability of finding an untransformed surface scratched area during the wear test in the linear range of the sliding direction between the site x and $x + dx$, at the time t , the rate equation for the time evolution of the probability of finding the untransformed surface scratched area is [46]

$$\frac{\partial \rho(x, t)}{\partial t} = \lambda \frac{\partial \rho(x, t)}{\partial x} - \Gamma(t)x\rho(x, t) + 2\Gamma(t) \int \rho(y, t)dy, \quad (18.21)$$

where λ is the growth velocity of nucleation of defects and $\Gamma(t)$ is the nucleation rate, eventually deducible by (18.17). After that, a wear test consisting of N scanning cycles wear has been performed; the total damaged area $X(t)$ corresponds to a summation of the wear domains generated via nucleation by the initial defects given by the following relation [46]:

$$X(t) = 1 - \int x\rho(x, t)dt, \quad (18.22)$$

where $\rho(x, t)$ is obtained by (18.21) and the integral expresses the unworn SAM area. Equation (18.22) can be simplified considering the residence time, τ , multiplied for the N scanning cycles [46]. The progressive worn area can be plotted as a function of the scanning cycles (Fig. 18.22).

The interfacial wear mechanisms as shown in Figs. 18.3.1 and 18.22 are mainly governed by a kinetic energy transfer from the AFM tip to the molecular film. The additional energy is dissipated through the SAM film by creating modifications in the molecular conformations and disorder, leading in some cases to plastic deformations up to breaking the covalent involving real damage in the SAM chains. The theoretical results confirm the common recognised experience that the higher the density of the defects, the faster the wear [72, 75, 82].

18.4 Conclusions

The chapter has been devoted to give a review about experimental data and correspondent models on the onset wear mechanisms occurring on nanoscale. Wear is the progressive loss of mass from the operating surface of a body as a result of relative motion at the surface. The occurrence of different wear mechanisms involves the fundamental mechanics of molecular and supra-molecular scale. As a consequence, a fundamental understanding of surface properties on nanoscale level should be generated to have a satisfactory knowledge of responses of materials and machine components also at macrometer scale. Investigation of the fundamental characteristics of wear at the microscale is complicated by some factors and forces that act on a nanoscale level, which have not yet been addressed in the tribology of macrosystems. Since these forces are sensitive to the environment and surface condition of the specimens, it is quite difficult to determine them accurately. Further, quantification of wear is not straightforward since the amount of wear is often too small to be detected by surface-sensitive instruments. Nanowear measurements have been the object of rapid development of precise measuring tools following the introduction of SPM family as, for example, AFM. The advent of AFM opened a powerful opportunity to study wear mechanisms on nanoscale, making possible to simulate one single asperity surface sliding on other one and controlling

many physical parameters during the contact dynamics. Some recent different experiments and few relevant theories show that the atomic abrasive wear is based on a smooth atom-by-atom wearing process without fracture. The wear rates are generally lower with respect to macroscale, and such rates increase with increasing load. The surface abrasive wear offers new opportunity for the possible role of the surface re-crystallisation that can lead to new patterned structures. In the following years, with the possibility to address new dedicated experiments, it is reasonable to suppose that our knowledge on the basic wear mechanisms and their control can open new technological challenges and opportunity on nanoscale.

References

1. B. Bhushan, *Nanotribology and Nanomechanics*, (Springer, Berlin, 2005).
2. B. Bhushan, *Springer Handbook of Nanotechnology*, (Springer, Berlin, 2003).
3. K. Ludema in *Fundamentals of Tribology and Bridging the Gap between the Macro- and Micro/Nanoscale*, ed. by B. Bhushan (Kluwer, 2001).
4. K. Ludema, *Friction, Wear and Lubrication, a Textbook in Tribology*, (CRC, USA, 1996).
5. I.M. Hutchings, *Tribology: Friction and Wear of Engineering Materials*, (CRC, 1992).
6. Y. Kim, W.N. Huang, C.M. Lieber, *Applied Phys. Lett.* **59**, 3404–08 (1991).
7. B. Bhushan, V.N. Koinkar, *J. Appl. Phys.* **75**, 5741–45 (1994).
8. J. Hu, X.D. Xiao, D.F. Ogletree, M. Salmeron, *Surf. Sci.* **327**, 358–61 (1995).
9. R.W. Carpick, M. Salmeron, *Chem. Rev.* **97**, 1163–94 (1997).
10. J. Schöfer, E. Santner, *Wear*, **222**, 74–83 (1998).
11. R. Gählin, S. Jacobson, *Wear*, **222**, 93–102 (1998).
12. J.F. Archard, *J. Appl. Phys.* **24**, 981–8 (1953).
13. K.H. Chung, D.E. Kim, *Tribol. Lett.* **15**(2), 135–44 (2003).
14. R. Colaço, in *Fundamentals of Friction and Wear on the Nanoscale*, ed. by E. Gnecco, E. Meyer (Springer, Berlin Heidelberg, 2006) p. 453–80.
15. U. Landman, W.D. Luedtke, E.M. Ringer, *Wear* **153**, 3–30 (1992).
16. J.A. Harrison, R.J. Colton, C.T. White, D.W. Brenner, *Wear* **168**, 127–133 (1993).
17. B. Gotsmann, M.A. Lantz, *Phys. Rev. Lett.* **101**, 125501(1–4) (2008).
18. R. Bassani, M. D'Acunto, *Tribol. Int.* **33**, 443–52 (2000).
19. S. Kopta, M. Salmeron, *J. Chem. Phys.* **113**, 8249–52 (2000).
20. M. D'Acunto, *Tribol. Int.* **36**, 553–61 (2003).
21. M. D'Acunto, *Nanotechnology* **15**, 795–801 (2004).
22. J.M. Helt, J.D. Batteas, *Langmuir* **21**, 633–9 (2005).
23. R. Erlandsson, G. Hadzioannou, C.M. Mate, G.M. McClelland, S. Chiang, *J. Chem. Phys.* **89**, 5190–93 (1988).
24. R.W. Carpick, N. Agrait, D.F. Ogletree, M. Salmeron, *J. V. Sci. Technol. B* **14**, 1289–95 (1996).
25. R.W. Carpick, N. Agrait, D.F. Ogletree, M. Salmeron, *Langmuir* **12**, 3334–40 (1996).
26. L. Xu, M. Salmeron, *Langmuir* **14**, 2187–90 (1998).

27. E. Barthel, S. Roux, *Langmuir* **16**, 8134–38 (2000).
28. S. Miyake, *Appl. Phys. Lett.* **65**, 980–2 (1994).
29. J.N. Israelashvili, *Intermolecular and Surfaces Forces*, 2nd edn. (Academic, New York, 1992).
30. J.T. Dickinson, N.S. Park, M.W. Kim, S.C. Langford, *Tribol. Lett.* **3**, 69–80 (1997).
31. N.S. Park, M.W. Kim, S.C. Langford, J.T. Dickinson, *Langmuir* **12**, 4599–604 (1996).
32. N.S. Park, M.W. Kim, S.C. Langford, J.T. Dickinson, *J. Appl. Phys.* **80**, 2680–6 (1996).
33. J.T. Dickinson, R.F. Hariadi, L. Scudiero, S.C. Langford, *Tribol. Lett.* **7**, 113–119 (1999).
34. L. Scudiero, S.C. Langford, J.T. Dickinson, *Tribol. Lett.* **6**, 41–55 (1999).
35. K.L. Johnson, *Contact Mechanics*, (Cambridge University Press, UK, 1985).
36. W. Maw, F. Stevens, S.C. Langford, J.T. Dickinson, *J. Appl. Phys.* **92**, 5103–09 (2002).
37. F. Stevens, S.C. Langford, J.T. Dickinson, *J. Appl. Phys.* **99**, 023529(1–8) (2006).
38. E. Gnecco, R. Bennewitz, E. Meyer, *Phys. Rev. Lett.* **88**, 215501(1–4) (2002).
39. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, *Phys. Rev. Lett.* **92**, 134301(1–4) (2004).
40. R. Bagnold, *The Physics of Blown Sand and Desert Dunes*, (Methuen, London, 1941).
41. P. Vettiger, G. Cross, M. Despont, U. Drechsler, U. Durig, B. Gotsmann, W. Haberle, M.A. Lantz, H.E. Rothuizen, R. Stutz, G.K. Binnig, *IEEE Trans. Nanotechnol.* **1**, 39–55 (2002).
42. E. Eleftheriou, T. Antonakopoulos, G.K. Binnig, G. Cherubini, M. Despont, A. Dholakia, U. Dürig, et al., *IEEE Trans. Magn.* **39**, 938–45 (2003).
43. B. Bhushan, K.J. Kwak, M. Palacio, *J. Phys.: Cond. Matter.* **20**, 365207(1–36) (2008).
44. B. Gotsmann, U. Duerig, J. Frommer, C.J. Hawker, *Adv. Funct. Mater.* **16**, 1499–505 (2006).
45. E. Gnecco, E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale* (Springer, Berlin, 2007).
46. M. D’Acunto, *Nanotechnology* **17**, 2954–62 (2006).
47. B.J. Briscoe, D.C.B. Evans, *Proc. R. Soc.* **380**, 389–407 (1982).
48. R. Kaneko, T. Miyamoto, E. Hamada, in *Handbook of Micro/Nanotribology*, ed. by B. Bhushan (CRC, Boca Raton, 1995).
49. M. D’Acunto, in *Advances in Contact Mechanics*, ed. by R. Buzio, U. Balbusa (Transworld Research Network, India, 2006).
50. O.M. Leung, M.C. Goh, *Science* **255**, 64–7 (1992).
51. Z. Elkaakour, J.P. Aime, T. Bouhacina, C. Odin, T. Masuda, *Phys. Rev. Lett.* **73**, 3231–34 (1994).
52. B.D. Beake, G.J. Leggett, P.H. Shipway, *Polymer* **42**, 7025 (2001).
53. F. Dinelli, G.J. Leggett, P.H. Shipway, *Nanotechnology* **16**, 675 (2005).
54. M. Surtchev, N.R. de Souza, B. Jérôme, *Nanotechnology* **16**, 1213 (2005).
55. B. Gotsmann, U.T. Duerig, S. Sills, J. Frommer, C.J. Hawker, *Nano Lett.* **6**, 296–300 (2006).
56. H.A. Kramers, *Physica* **7**, 284–304 (1940).

57. N. van Kampen, *Stochastic Processes in Physics and Chemistry*, (North-Holland, Amsterdam, 1983).
58. H. Risken, *The Fokker-Planck Equation. Methods of Solution and Applications*, (Springer, Berlin, Heidelberg, 1984).
59. R.F. Bader, Phys. Rev. B **61**, 7795–802 (2000).
60. D.L. Stein, C.R. Doering, R.G. Palmers, J.L. van Hemmen, R.M. McLaughlin, J. Phys. A, Math. Gen. **23**, L203–8 (1990).
61. A. San, M. Wortis, Phys. Rev. E **70**, 031102(1–11) (2004).
62. D'Acunto, unpublished.
63. R.P. Sear, Phys. Rev. E **70**, 021605(1–6) (2004).
64. S. Jun, H. Zhang, J. Bechhoefer, Phys. Rev. E **71**, 011908(1–6) (2005).
65. A. Ulman, *Introduction to Ultrathin Organic Films: From Langmuir-Blodgett to Self-Assembly*, (Academic, San Diego, CA, 1991).
66. R.G. Nuzzo, D.L. Allara, J. Am. Chem. Soc. **105**, 4481–3 (1983).
67. L. Haussling, H. Ringsdorf, F.J. Schmitt, W. Knoll, Langmuir **7**, 1837–40 (1991).
68. H.M. Schessler, D.S. Karpovich, G.J. Blanchard, J. Am. Chem. Soc. **118**, 9645–51 (1996).
69. H.C.D. Tidwell, S.I. Ertel, B.D. Ratner, B.J. Tarasevich, S. Atre, D.L. Allara, Langmuir **13**, 3404–13 (1997).
70. V.N. Bliznyuk, M.P. Everson, V.V. Tsukruk, J. Tribol. **120**, 489–96 (1998).
71. E. Barrena, S. Kopta, D.F. Ogletree, D.H. Charych, M. Salmeron, Phys. Rev. Lett. **82**, 2880–3 (1999).
72. B. Bhushan, H. Liu, Phys. Rev. B **63**, 245412(1–11) (2001).
73. M. Salmeron, Tribol. Lett. **10**, 69–79 (2001).
74. J. Hautman, M.L. Klein, J. Chem. Phys. **91**, 4994–5001 (1989).
75. K. Mougín, H. Haidara, in *Fundamental of Friction and Wear on the Nanoscale*, ed. by E. Gnecco, E. Meyer (Springer, Berlin, 2006).
76. A. Lio, D.H. Charych, M. Salmeron, J. Phys. Chem. B **101**, 3800–5 (1997).
77. K. Edinger, A. Golzhauser, K. Demota, C. Woll, M. Grunze, Langmuir **9**, 4–8 (1993).
78. D.J. Lavrich, S.M. Wetterer, S.L. Bernasek, G. Scoles, J. Phys. Chem. B **102**, 3456–65 (1998).
79. B.L. Weeks, A. Noy, A.E. Miller, J.J. De Yoreo, Phys. Rev. Lett. **88**, 255505(1–4) (2002).
80. Y.S. Leng, S. Jiang, J. Chem. Phys. **113**, 8800–6 (2000).
81. X. Nie, P. Zhang, A.M. Wiener, Y.T. Cheng, Nano Lett. **5**, 1992–6 (2005).
82. H. Liu, B. Bhushan, Ultramicroscopy **91**, 185–202 (2002).

Contact Potential Difference Techniques as Probing Tools in Tribology and Surface Mapping

Anatoly Zharin

Summary. Contact potential differences techniques have been adapted for continuous nondestructive monitoring of changes in the electron work function of a rubbing surface. The method can be used to investigate tribological materials for a wide range of conditions, including changes in load, sliding speed, and environment, with or without lubrication. It relies on the sensitivity of the work function to the various events, which accompany friction, for example, plastic deformation, creation of new surface of material, adsorption, oxidation, phase changes, and redistribution of alloy components. At present, this is the only method sensitive to both surface and near-surface defects and permits study of one of the two interacting surfaces during sliding. For metals and alloys, the thickness of a layer contributing to the electron work function measurement is equal to several atomic distances, that is, even traditional contact potential differences measurements is really related to nanoscale. Kelvin probe force microscopy allows to determine not only the surface topography as does atomic force microscopy, but in addition also delivers images of the surface work function on a nanometer scale. Modern contact potential differences techniques cover the range from macro/micro to nanoscales. The current paper focused on an in situ contact potential difference measurement during the sliding of materials.

Key words: Contact potential difference, Electron work function, Friction, Friction fatigue, In situ Kelvin probe, Kelvin probe, Surface mapping, Wear.

19.1 Introduction

The existence of contact potential difference (CPD) in a dry environment has been well known since the beginning of the modern history of electricity. In 1801, the Italian physicist Alessandro Volta was the first to demonstrate the CPD effect.

CPD techniques of electronic work function (EWF) measurement are an excellent nondestructive monitoring technique. Lord Kelvin offered the basis of CPD measurements in 1898 [1]. CPD techniques were developed considerably parallel with quantum theory of solids. As researchers were trying to

correlate experimental data with theory, the EWF was explained according to the fundamental quantum mechanical parameters of solids. However, a strong influence of surface conditions on the experimental results was found and the technique was practically forgotten. Later, problems in the measurement of surface conditions have gained a special importance with the development of solid-state electronics. However, systems of surface analysis began to appear during the same years. Such systems were complicated devices attached to ultrahigh vacuum systems. These systems have overshadowed the CPD technique. An analysis of published papers has shown that surface analysis systems yield interesting results when conducting fundamental experiments with pure model surfaces. Results are not reliable for most of engineering surfaces. It is explained that surface analysis systems, in most cases, do not analyze the surface, but instead analyze artifacts on the surface. According to our experience, CPD does give reliable information about the surface. Moreover CPD technique is good enough to provide an in situ friction surfaces monitoring.

An extreme sensitivity of EWF to practically all surface perturbations is the drawback of the CPD technique. In particular, there is a big problem of measurement of absolute value of EWF. In contrast, our practice shows that using “black box” approach it is possible to get unique results. The best experiments with CPD technique should be registering EWF during a perturbation. In our case, it is on-line EWF registration during mechanical or friction testing as well as surface mapping.

In the present paper, the results of more than 30 years worth of the author’s works with CPD and its application for tribological studies and surface mapping are described.

19.2 Electron Work Function as a Parameter for Surfaces Characterization

The forces of interaction with a crystal lattice do not allow electrons to escape from the metal, that is, there is a potential barrier on a metal surface. Only those electrons with sufficient energy to overcome these forces can escape from the metal [2]. Two terms, the dipole and the quantum-mechanical exchange-correlation, contribute to the surface potential barrier [3]. A classic analogue of the exchange-correlation term is an image force. The dipole term arises because the wave function of an electron has a “tail” at the metal surface. The density of electrons near the metal surface is not equal to zero and is not compensated by a positively charged ionic skeleton of the metal lattice. Therefore, there is a dipole layer on the metal surface.

The electrons in the metal have energy even at absolute zero. This energy is termed the Fermi level. Therefore, for the removal of an electron from the metal, it should receive energy equal to the difference between the surface potential barrier and the Fermi level. This energy is termed as the EWF.

The EWF is one of the fundamental characteristics of the condensed (solid) metallic state. It correlates with some physicochemical and mechanical characteristics of metals. This characteristic applies to the study of the phenomena of deformation and destruction of metals, and to the study of adsorption and desorption.

The surface changes considerably under friction. Many of the parameters vary; among these are roughness, temperature, chemical and phase composition, defect density, and others. In addition, a variety of physical and chemical processes are in force. The formation and destruction of oxide and lubricating films, and adsorption and desorption of lubricants and gases take place at the rubbing surface. Practically all these processes will influence the EWF, directly or indirectly.

The chemical composition of alloys influences the position of the Fermi level considerably and, accordingly, affects the surface EWF. This issue has been well-described in the literature [4]. Generally, the chemical composition of the bulk does not vary under friction. However, considerable changes in the chemical composition at the rubbing surface can occur. Therefore, the Kelvin technique in general allows for the study of tribochemical processes.

The Kelvin technique's applicability for monitoring a rubbing surface that has a lubricant may be a problem. Apparently, the following processes will contribute to the EWF change with a deposition of lubricant: the adsorption of lubricant molecules, the chemical interaction of lubricant with oxides and the material, the formation of a dipole layer on a metal-lubricant boundary, and others. A screening property of the lubricant layer can contribute to the measurement results also. Pekar et al. [5] have found that the EWF of the metal surface covered with an oxide or semiconductor layer should essentially not vary if the layer thickness is much less than the Debye length of screening. The Debye length of screening depends on the density of free charge carriers (electrons, ions, and holes) in the substance. Metal oxides and the majority of lubricants are dielectrics or semiconductors with a low density of free charge carriers. In these cases, the Debye length of screening has a high value (sometimes centimeters) [6]. As it is known, there is a high density of free electrons in metals. Therefore, for metals and alloys the thickness of a layer contributing to the EWF measurement is equal to several atomic distances. It follows, then, that when studying surfaces covered by dielectric film (lubricant, oxide, etc.), one will only study the metal behavior through the dielectric film. In the case of lubricated friction, the influence of the change of the lubricant film thickness to the EWF can be neglected.

The adsorption processes on the metallic surface strongly change its EWF. It is governed by the redistribution of electrons at the formation of an adsorptive bond between a free atom or molecule and the surface dipole layer [7]. The EWF change of the metallic surface by adsorbed atoms or molecules is directly proportional to their effective electrical dipole moment and the degree of surface occupation. The high sensitivity of the EWF to the adsorption has led to wide use of this parameter for the study of adsorption processes of

organic and inorganic molecules and atoms, and for a study of the oxidation. The adsorption processes must be taken into account for studies of lubricated friction. The adsorption of molecules creates a double electric layer on the “metal–lubricant” boundary. Certainly, taking into consideration the contributions of all processes happening on a rubbing surface simultaneously is rather complicated. Therefore, it is convenient to consider the contribution of the double electric layer as constant. In practice, one can create a condition, through the control of lubrication (e.g., flowing), where this contribution will be a constant.

Mechanical perturbations can influence the EWF through the Fermi level. The compactness of the metal lattice varies under mechanical compression or tension. Thus, the position of Fermi level varies and creates a EWF change [8]. The EWF will grow with an increase in compression for the majority of metals. The typical EWF change is about several $\mu\text{eV}/\text{bar}$ [9]. Thus, the EWF varies insignificantly with changes in lattice compactness in the elastic regime.

Crystalline defects in metals influence the EWF to a greater degree. The greatest influence on EWF is rendered by linear defects (boundary and screw dislocations). The atoms in the neighborhood of a dislocation are under a considerable hydrostatic pressure created by its stress field [10]. Latishev et al. [11] estimated the local EWF change in a dislocation core on the surface theoretically. They found that at a dislocation core the EWF is reduced by approximately 0.3 eV. Therefore, the integrated value of EWF can vary considerably with a change in the defect density. Mints et al. [12] have found theoretically that the dislocations can result in a decrease of an integrated surface EWF by 10^{-2} – 10^{-1} eV.

Recently, DeVecchio and Bhushan [13] showed that the change in surface potential during friction and wear at ultralow loads might be due to chemical and structural changes in the first few nanometers of the sample.

19.3 Measurements of Contact Potential Difference

The EWF can be measured by direct and indirect techniques. Direct techniques induce electron emission, and their accurate sensing requires operation in a vacuum. The indirect techniques do not require high vacuum conditions. The techniques of EWF determination on CPD are most convenient from the point of view of rubbing surface monitoring and surface mapping. Therefore, we shall consider it in detail.

Let us consider the contact phenomena in the case of two metals with a thin vacuum gap (Fig. 19.1a) [14]. The EWF (φ , Fig. 19.1) is the difference between the Fermi level (E_F) of the metal and the surface potential barrier. In Fig. 19.1a, the surface potential barrier is shown by a dotted line, while the solid horizontal lines show the Fermi level. This figure corresponds to the initial time, when the metals separated by a distance d_0 , at which an effective exchange of electrons stipulated by a thermionic emission is possible.

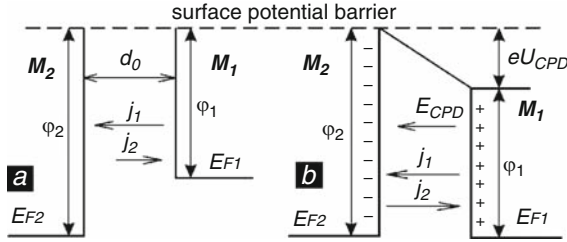


Fig. 19.1. The contact phenomena in the case of two metals with a thin vacuum gap

The regularities of the thermionic emission are described by the Richardson-Dushman equation: $j = AT^2 \exp(-F/kT)$, where T is the temperature, k is a Boltzmann constant, and A is a Richardson constant. If $\varphi_2 > \varphi_1$, then $j_1 > j_2$ and there will be an electron transfer from the right to the left (as it is shown in a Fig.19.1a). The first metal will be charged positively and the second metal will be charged negatively. The electric field E_{CPD} and, accordingly, the potential difference U_{CPD} will appear in a gap. An equilibrium condition implies an equality of Fermi levels. The thermoemission currents are aligned, that is, $j_1 = j_2$. This equality means that the potential barrier $\phi_1 + eU_{CPD}$ should be equal to the potential barrier φ_2 . In other words, $\phi_1 + eU_{CPD} = \varphi_2$. Thus, after equilibrium establishment the CPD becomes $U_{CPD} = (\varphi_2 - \varphi_1)/e$. The case above is only a model. In practice, equilibrium (equality of Fermi levels) is attained by an external electric circuit.

19.3.1 Kelvin-Zisman Probe

The condenser technique is the most widely applied CPD measurement. The metal plates M_1 and M_2 form a parallel-plate capacitor of capacitance C (Fig. 19.2a). The stipulated CPD charge Q on the capacitor will be $Q = CU_{CPD}$. The gap between the capacitor plates periodically changes due to the vibration of one of the plates (introduced by Zisman) [15]. Therefore, the capacitance periodically varies and an alternating current varies with the frequency of plate vibration. A source of compensation voltage U_{comp} is included, the output of which can be adjusted so that it will compensate the CPD. In this case:

$$Q = C(U_{CPD} + U_{comp}) = 0 \tag{19.1}$$

The fundamental equation that characterizes Kelvin probe current response is given as

$$i = \frac{dQ}{dt} = \frac{d(CU_{CPD})}{dt} = C \frac{dU_{CPD}}{dt} + U_{CPD} \frac{dC}{dt} \tag{19.2}$$

The Kelvin-Zisman method results in current flow by vibrating the probe relative to the surface of interest. The Kelvin probe is generally fixed in position

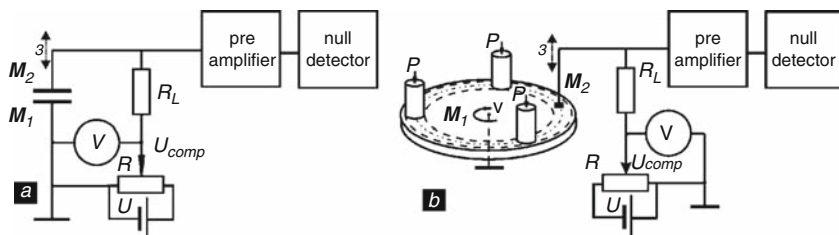


Fig. 19.2. Functional diagram of the generalized Kelvin technique (a) and its modification for friction testing (b)

above a surface in quasi-equilibrium, so the dU_{CPD}/dt term is assumed to be negligible.

Because of the time-varying capacitance, the alternating voltage appears on the high-Ohmic resistor and applies to the input of an electrometric preamplifier (an amplifier with a high input resistance) and to the indicator of a zero signal. The compensation voltage is equal to the CPD when Q is set to zero (19.1), from which the CPD is determined. This is the basis of the large number of devices for the CPD measurement.

Another step in the development of the CPD measurement technique was the application of automatic compensation of a measurable value [16,17]. After amplification, the output signal of the vibrating-reed capacitor is rectified by a phase-sensitive detector and fed back to the vibrating-reed capacitor for CPD compensation. An automatic record of the EWF changes is thus performed.

19.3.2 Nonvibrating Probe

For a topology study on a surface without determination of the EWF value, a technique using nonvibrating capacitor was offered, that is, a kind of Kelvin technique that allows measurement of EWF changes without vibrating the capacitor plates [6, 18]. The basis of such an approach is as follows.

The nonvibrating CPD method is similar to the Kelvin-Zisman method. However, the probe is translated, rather than vibrated, over a surface. If the surface does not exhibit large geometry changes, then the dC/dt term in (19.2) is negligible and the probe signal results from the dU_{CPD}/dt term of (19.2). The advantages of this method are that it can be applied to any rotational machinery for in-situ monitoring of chemical or physical changes.

Figure 19.3 shows the physical arrangement for this kind of probe operation. The probe also forms a capacitor between the reference electrode and surface-of-interest; however, the probe is not vibrated. The probe is positioned a fraction of a millimeter above a moving surface. Motion (scanning) of the probe produces a continuous stream of data at a series of points. Output signal is generated when probe passes over surface with geometrical features [according to the first term of (19.2)] or different chemistry, deformation, and adsorption [according to the second term of (19.2)].

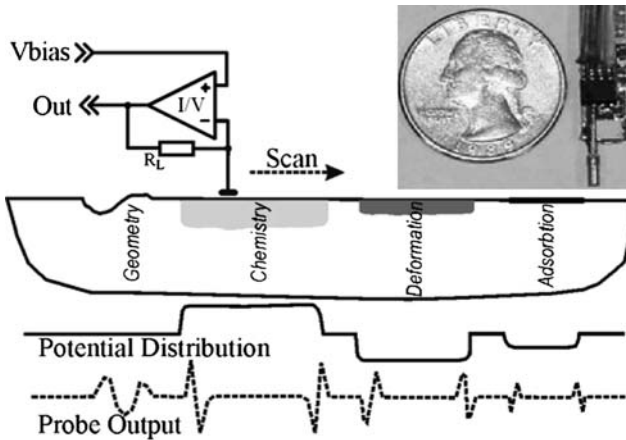


Fig. 19.3. The schematic diagram of the nonvibrating probe functionality and photo of the probe

The question is the separation contributions from geometrical and potential spots on the surface. This problem was solved by applying the bias voltage to the probe. Output signal from geometrical features is changed with bias value and from the potential spots keeping stable. Using computer control and data processing, geometrical and potential contributions were separated. Another technical problem solved by the computer data processing was increasing signal-to-noise ratio. Data accumulation from several scans over the same line let us eliminate noise. Finally as small as $10\ \mu\text{m}$ in diameter reference electrode were used for surface mapping.

Albeit simple, technique of a nonvibrating capacitor has been shown to be reliable [19–21]. One needs only a high-Ohmic current-to-voltage converter. Therefore, it can be used as a built-in sensor for surface monitoring. The nonvibrating probe does not provide the EWF value but its derivative. It can be used for monitoring the local spots on a surface and for surface topology studies, but not for long-term changes in integrated EWF of the surface. This has been previously used to study wear of hard disks and wear of metals in pin-on-shaft experiments.

Harvey et al. [22] reported recently an application of the electrostatic charge sensitive technology to monitor wear. This approach is very similar to the nonvibration probe.

19.3.3 Ionization Probe

Another modification of the CPD technique, which is promising for tribological applications, is the ionization technique. In this technique, the air gap between the capacitor plates is ionized by a source of radioactive particles. Ions are moved in the electric field (E_{CPD} , Fig. 19.1) to the opposite charged

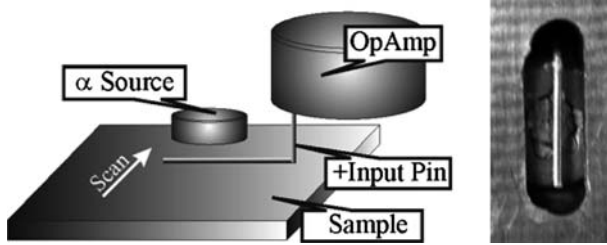


Fig. 19.4. The schematic diagram of the ionization probe and photo of the reference electrode arrangement

plates, that is, there is an ionic current between the capacitor plates. The increased potential is a result of the ionic current. Thus, the potential difference between plates will be equal to the CPD and can be measured with an electrometric voltmeter. Currently, the first steps in its usage for tribology are underway at the Georgia Institute of Technology (Danyluk and Zharin) (Fig. 19.4).

19.3.4 Atomic Force Microscope in Kelvin Mode

CPD techniques described earlier are dealing with macro- and microlayers of the surfaces but with nanoscale thickness (at least for metals). Recent advances in MEMC technology emphasized the demands of tribology testing in nanoscale. The promising tool for nano-objects is the atomic force microscopy (AFM). Since the invention of noncontact AFM, the electrostatic force has been used to measure electrical properties of various samples without contact. Using an AFM, Martin et al. [23] performed the first surface-potential measurements adapting the concept of Kelvin probing to the local scale accessible by scanning force microscopy. Further developments by Nonnenmacher et al. [24] introduced the modulation method for measuring surface-potential differences between various metallic electrodes. Consecutive improvements by Lu et al. [25] proved to be advantageous for Kelvin probe force microscopy (KPFM) because they drastically increased the measurement sensitivity. KPFM is currently used by many groups to measure CPD variations on a nanometer scale by AFM.

The CPD between the conductive tip and the sample generates electrostatic force and the mechanical deflection of the cantilever (Fig. 19.5a). This force F takes the form of

$$F = \frac{1}{2}(U_{\text{CPD}})^2 \frac{dC}{dz} \quad (19.3)$$

where z is the tip-sample distance.

External potential U_{comp} can be applied to zero the force. If an AC voltage $U_{\text{AC}} \sin(\omega t)$ with adjustable offset U_{comp} is applied to the conducting AFM tip, the electrostatic force interaction between the two electrodes becomes

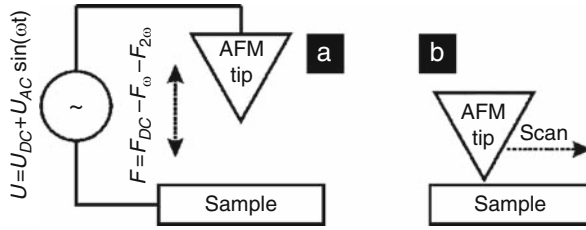


Fig. 19.5. The schematic diagram of the KPFM (a) and the AFM in contact operation mode for simulation of tribotester (b)

$$F = \frac{1}{2} (U_{CPD} + U_{comp} + U_{AC} \sin(\omega t))^2 \frac{dC}{dz} \tag{19.4}$$

The spectral components of force at DC and with frequencies ω and 2ω are

$$F_0 = \frac{1}{2} \left((U_{CPD} + U_{comp})^2 + \frac{U_{AC}^2}{2} \right) \frac{dC}{dz} \tag{19.5}$$

$$F_\omega = (U_{CPD} + U_{comp}) U_{AC} \sin(\omega t) \frac{dC}{dz} \tag{19.6}$$

$$F_{2\omega} = \frac{1}{4} U_{AC}^2 \cos(2\omega t) \frac{dC}{dz} \tag{19.7}$$

As a result of the F_ω and $F_{2\omega}$, the tip starts vibrating. The vibration amplitude at frequency ω is proportional to F_ω and can be easily monitored during operation of the AFM. If the DC offset U_{comp} applied to the tip is equal in magnitude and opposite in sign to U_{CPD} , the spectral component at frequency ω of the tip vibration amplitude as well as of the force acting on the tip will become zero. The principle of operation of KPFM is based on nullifying tip vibration amplitude by applying adjustable DC voltage. The 2ω component of the tip vibration amplitude can be used to control tip-sample gap similar to that in noncontact AFM mode to obtain surface topography.

It is interesting to point out that practically the same approach was used to measure CPD at macro scale by Yousef et al. in 1965 [26].

The contact AFM mode could be adjusted for simulation friction tester for tribology studies (Fig. 19.5b). Bhushan and coworkers at the Ohio State University performed the first CPD measurements adapting the concept of microtribology studies [13]. They used contact AFM mode to simulate wear test consequently with CPD and topology measurements on the same sample. It was shown that even in the cases where there was little or no damage to the surface, as observed by the topography scans of an AFM, there was a change in surface potential of the sample. On the basis of data obtained, the authors concluded that the change in surface potential during friction and wear at ultralow loads might be due to chemical and structural changes in the first few nanometers of the sample. This approach is currently used for detecting wear precursors (precursors to measurable wear) in the field of nanotribology

(see, e.g., [27–30]). This technique seems to be very attractive for tribology studies.

19.4 Typical Electron Work Function Responses

19.4.1 Surface Deformation

The EWF is repeatedly referred to in studies of surface conditions of metals and alloys (more than 2,000 citations in Ref. [31]), including responses of EWF on different mechanical perturbations. Recently, Li et al. [32, 33] made a lot of efforts to revise the data in the field.

We carried out additional experiments on the influence of deformation the EWF for simple cases of loading [34, 35]. The goal of such experiments was to provide a basis for interpretation of tribological data.

The test machine “Instron” was used. The Kelvin probe was mounted on a lateral surface of the sample. The sample elongation and EWF were recorded simultaneously. Typical results for the EWF change under a simple single axial tension for a case of soft steel are shown in Fig. 19.6a. It follows from the figure, in which up to the yield point EWF varies insignificantly. However, there is a sharp EWF decrease. The nucleation and movement of dislocations in surface layers can explain such EWF decrease. This is consistent with Vishniakov’s [36] conclusion that the EWF-integrated value can vary considerably with a change in defect density. In our case, it varies with changes in dislocation density on the surface. It is known [36] that the dislocation density in the material can be increased up to some critical value only. After this, the equilibrium between the nucleation and coalescence of dislocations is achieved. To our minds, it can explain the saturation of the EWF

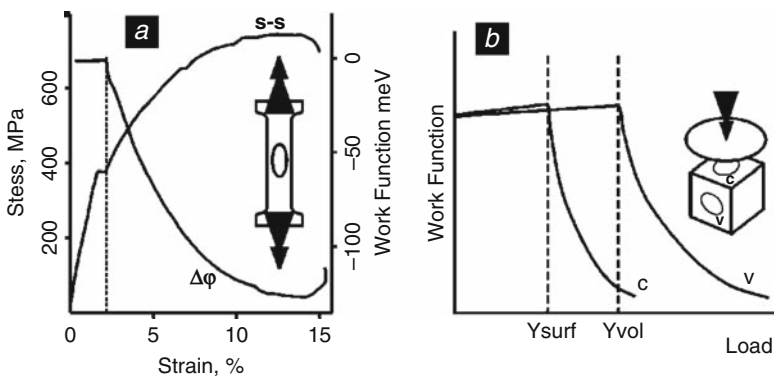


Fig. 19.6. Results of simultaneous measurements of the regular stress–strain diagram and the work function for medium carbon steel (a) and schematic representation of experiments on contact deformation (b)

under the further deformation in a plastic area. It is necessary to point out that similar EWF behavior at deformation takes place also for materials not having a yield point. This fact is promising for the use of the Kelvin technique for general mechanical testing.

Generally, the behavior of EWF dependence on deformation (both tensile and compressive) is identical for the majority of metals. There is a small increase in EWF in the elastic regime caused by the change in the Fermi level due to the crystal lattice dilatation. It decreases quickly at the hardening stage. The speed of the EWF change decreases as dynamic recovery is approached [36]. Thus, the EWF can characterize the mechanical processes on a metal surface both in the elastic and plastic regimes.

Problems of deformation processes in thin surface layers are of particular interest in tribology. For the majority of metals and alloys, the thickness of the metal layer, which contributes to the EWF, has a value approximately equal to that of the between atomic distance. The other feature of the technique is that the measurements are produced integrally on a rather large segment of the surface equal to the dimension of the reference electrode (in our case, $\sim 5 \text{ mm}^2$).

A sample was mounted in the test machine and was compressed through a glass plate within 30s. Then it was removed from the test machine and the EWF of the lateral and contact surfaces were measured. The roughness parameter (Ra) of the contact surface was also measured. This procedure was carried out with incremented load and dependencies of the EWF of lateral and contact surfaces and the roughness of the contact surface on the applied load were found. In Fig. 19.6b, such dependences and the method of loading are shown schematically. In this case, EWF measurements of the lateral surface are indicative of the behavior of the bulk of the material.

In Fig. 19.6b, it is shown that the character of the dependency of EWF on the contact (c) and lateral (v) surfaces, on the applied load, is similar. In both cases, a sharp decrease in the EWF is observed at some load value. The decrease of the curve for the contact surface (Y_{surf}) is observed under a smaller load than for the lateral one (Y_{vol}). It is possible to explain it by the deformation of the contact surface asperities and by the beginning of plastic processes (reaching yield point) under smaller loadings than for the bulk of material. The surface roughness does not vary until some critical value under small loads (Fig. 19.7). The roughness begins to decrease under a further load increase for the case in Fig. 19.7a. The contact surface roughness does not decrease, but increases under the sample loading with rough plates (Fig. 19.7b). From Fig. 19.7, it is clear that the EWF changes occur a little bit earlier than roughness changes begin. Hence, the beginning of the EWF decrease in this case signals a yield point in the surface layer even before the plastic deformations are measured by traditional methods.

Thus, the Kelvin technique can be applied for a study of deformation of thin surface layers. In particular, it can be used for the determination of a contact deformation type and yield points at contact deformation. There is

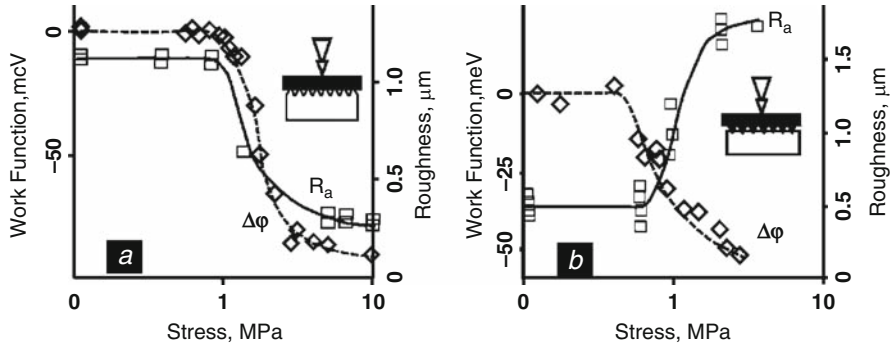


Fig. 19.7. Results of simultaneous measurements of the surface roughness and the work function depending on load for cases of “flat on rough” (a) and “rough on flat” (b) loading

an opportunity of experimental verification and improvements of the existing contact theories of rough surfaces.

In the Kelvin technique, the average distance between plates of a vibrating-reed capacitor greatly exceeds the dimensions of the asperities of a surface and the variations of the roughness values do not affect the results generally [37]. However, the dependence of the surface EWF on roughness should be considered in connection with a physical condition of the surface. The presence of surface defects, their character, and density will influence the EWF. Nazarov et al. [38] have applied the CPD technique to monitor the processing of mirror metallic surfaces. They have mentioned that, by using a conventional technique, it is possible to monitor roughness up to $0.1\ \mu\text{m}$ only. The CPD technique can only be applied for surfaces with roughness less than $0.1\ \mu\text{m}$.

From the above description, it follows that the basic contribution to the EWF under friction would be introduced by the deformation. More exactly EWF changes with the density change of dislocation cores on a surface, density change of point defects on the surface, and change of atomic roughness. The next most significant is the contribution from a change in chemical composition of the surface. It is possible to interpret the results from friction, basically, from the point of view of defect dynamics.

The adsorption of both lubricant molecules and molecules of ambient gases also contribute to the EWF considerably. However, the influence of adsorption processes can be minimized by creating constant conditions of lubrication and ambient gases when conducting long-time experiments.

19.4.2 Friction

Application of the CPD techniques in the field of tribology was in the great demands over the years. Practically all reported tribology related application of CPD techniques were related to postfriction surface studies (*see, e.g.,*

[39, 40]). Using a vibrating CPD probe, Zharin et al. [41, 42] performed the first EWF measurements of friction surface in situ adapting the concept of CPD probing to the on-line rubbing surface monitoring.

Figure 19.2b shows a generalized modification of the Kelvin technique suitable for friction testing on-line [41, 43, 44]. The device for the CPD measurement in the dynamics of the friction should meet high requirements concerning a noise reduction and stability of operation since high level of vibrations, acoustic, and electromagnetic noise originate from the friction test rig. It is necessary to provide an effective guard from mechanical, electrical, and magnetic disturbances. The basic requirement for friction testing machines is the presence of a reliable grounding for the rotating sample and the presence of a segment of friction track not covered by a counter sample.

The essential point is the relationship between the reference electrode dimensions and the friction track width. At the large reference electrode, the device will “feel” a material located near a friction track that can give an error in measurement. In order to eliminate this, reference electrode should be selected to have a narrower friction track (2 and 7 mm accordingly for results described below).

The generalized circuit of the Kelvin probe was developed with the above requirements in mind (Fig. 19.8a) [6, 44]. In Fig. 19.8a, the probe is vibrated by a vibrator and is located at a distance less than 0.5 mm from the sample (rubbing surface of the rotating disc). Together they form a capacitor. The alternating current induced by the periodic change in the capacitance is converted to a voltage signal by an I/V converter. The signal comes to the phase detector together with the reference signal from the vibrator. On detection, the signal goes to the noninverting input of I/V converter and, finally, to the reference electrode to compensate CPD. The circuit maintains an output voltage equal to the CPD. When necessary, an integrator with a time constant much exceeding the period of revolution of the specimen can be incorporated between the recorder input and the probe output to record the integrated

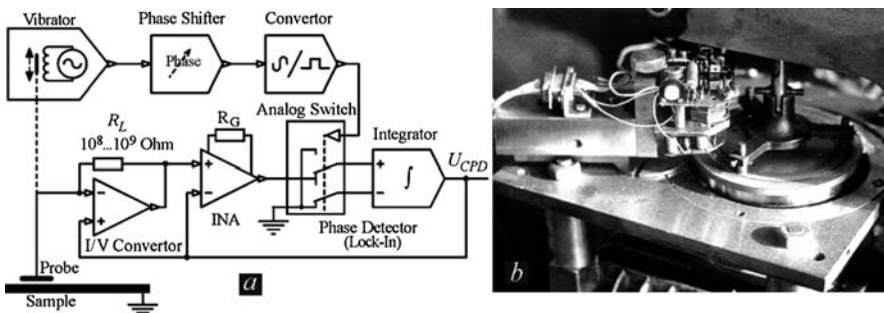


Fig. 19.8. Schematic diagram of measurement system (a) and photograph of Kelvin probe mounted on the tree pin on disk friction tester (b)

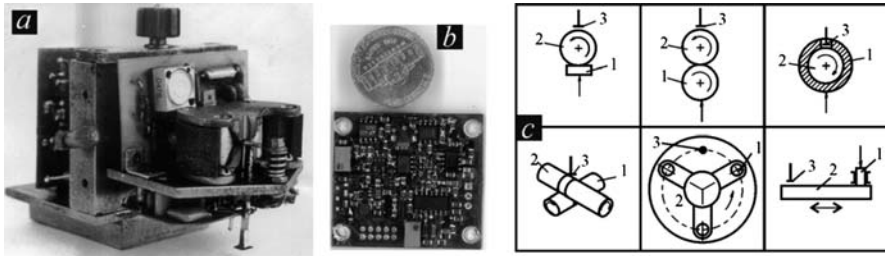


Fig. 19.9. Photographs of stand-alone Kelvin probe for tribology studying (a) and modern CPD board with all electronics shown on Fig. 19.8a (b). Possible arrangements for the method of implementation with different friction testing machine geometries. (1) Counterbody; (2) specimen; (3) reference electrode (c)

EWF of the rubbing surface. Figure 19.8b shows a photograph of CPD probe mounted on the three pin-on-disk friction testing machine.

The experience of designing the Kelvin probes has shown that the modern electronics allows one to make them as a small device entirely located near the rubbing surface (Fig. 19.9a and b). Practically any friction machine may be equipped with the Kelvin probe when the surfaces do not fully overlap (Fig. 19.9c). This includes reciprocating machines also [45]. Some technical problems of the CPD probe design have been considered in [41, 43–45].

The Kelvin technique can be used in two experimental modifications [6, 46]:

- Monitoring of the EWF with a small time constant of the signal integration and its synchronization with the measurement of sample position, allowing the investigation of the topology of the EWF;
- Monitoring the EWF with the time constant of the signal integration exceeding the period of revolution of the sample, allowing investigation of integral EWF changes over the entire rubbing surface.

In the former case, one can study the rubbing surface topology evaluation with sliding time, as shown in Fig. 19.10a. The data on the Fig. 10a were obtained with Kelvin probe. The nonvibrating probe is a much better way for studying the rubbing surface topology evaluation. The basic requirement is relatively high speed of sample movement. Figure 10b shows the rubbing surface work function topology evaluation as a function of time of friction registered by the nonvibrating probe. It is possible to study the evaluation “hot spots” or spots of damages on the rubbing surface with time of friction.

In most of the experiments conducted, we used the second approach, allowing investigation of integral EWF changes over the entire rubbing surface. In our opinion, the properties of the entire rubbing surface are responsible for the tribological behavior of materials.

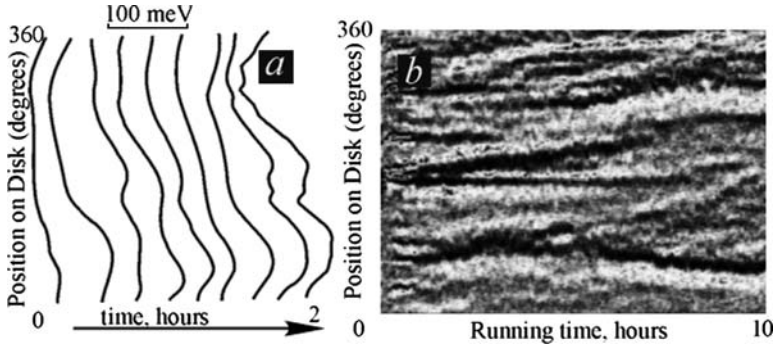


Fig. 19.10. Rubbing surface work function topology evaluation as a function of time of friction registered by the Kelvin probe (a) and nonvibrating probe (b)

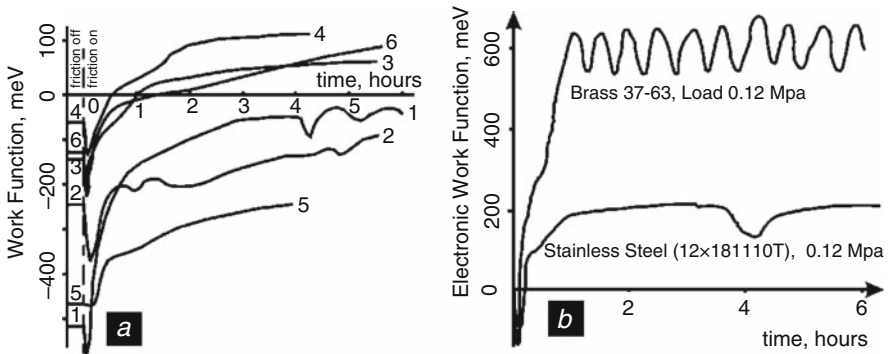


Fig. 19.11. Variation of the integral value of rubbing surface work function with time for some materials (1, brass; 2, bronze (Al); 3, bronze (Sn); 4, commercial copper; 5, stainless steel; 6, medium carbon steel) at the load of 0.05 MPa (a) and 0.12 MPa. In case (b), initial values of work function are zeroed for both metals

19.4.3 Experimental Examples of Kelvin Technique Application

Figure 19.11 shows the variation with time of the rubbing surface EWF for some materials. At the initial friction contact, a sharp decrease is followed by a sharp increase of the EWF relative to the initial value. Our interpretation is that the initial transients correspond to changes in the original oxide or other surface layers and running-in of the sliding surfaces. After a certain time, which depends on material properties, the friction regime stabilizes. From Fig. 19.11, one can find running-in finishing and the beginning of steady state friction conditions [47].

Examples of rubbing surface EWF responses on some perturbations are shown in Fig. 19.12. Figure 19.12a shows the EWF responses of different materials with the same lubricant (the initial EWF values of materials are equalized). The speed of the lubricant layer formation can be determined

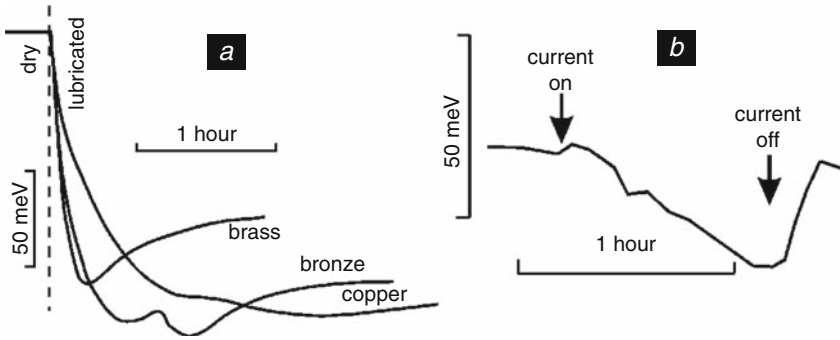


Fig. 19.12. Rubbing surface work function responses using the same lubricant for different materials (a) and rubbing surface work function response when applying an electric current through the friction contact (b)

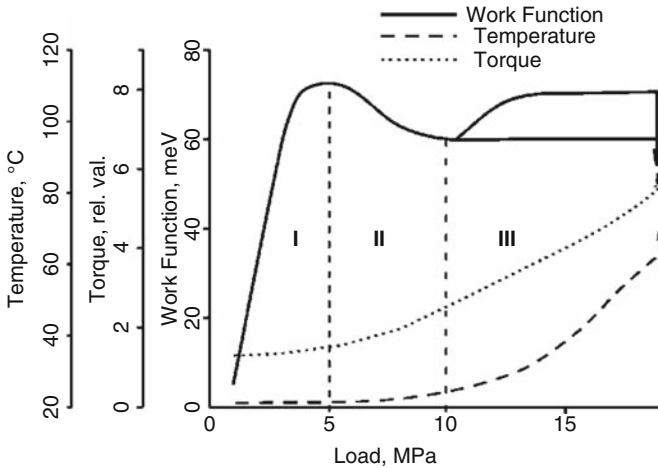


Fig. 19.13. Bronze rubbing surface electron work function, torque, and surface temperature vs. normal load curves

from such results. It is possible that this could be a fruitful way to study lubricant properties. This technique is applicable for sliding electric contact investigation as well. Figure 19.12b shows the responses of the rubbing surface EWF when the electric current passes through the friction contact.

The CPD technique was used for two major purposes: (1) to determine the critical points with respect to changes in normal load, with relevance to the materials selection and optimization, and (2) to determine the kinetics of friction processes, including periodic changes which may be related to those in fatigue.

The studies of a wide range of materials have shown a similar qualitative dependence of the rubbing surface EWF on the normal load (Fig. 19.13).

Generally, there are three specific zones [35, 46]. There is an increase in the EWF with load in zone I. At a higher load, there is a qualitative change of the curve. In zone II, EWF decreases. With further loading in zone III, there is very little change until the beginning of scoring, which is detected by increases in friction force and bulk temperature of the sample. During scoring, the value of the EWF decreases sharply. Hence, according to the character of the EWF changes, one can record changes in the material surface layers that do not manifest themselves in the external friction parameters (e.g., friction force and bulk temperature).

It has been found that during the long-run trials the wear rate is very low under loads corresponding to zones I and II. For zone III, the damaged spots on the surface and high wear rate are observed. The additional investigations of specimens via independent methods as well as studies concerning changes of the EWF during simple deformation (Figs. 19.6 and 19.7) allow us to interpret the results obtained on the base of dislocation interactions. We suggest that the first zone corresponds to mainly elastic and early stages of plastic deformation without a significant increase in dislocation concentration near the surface. In the second zone, plastic deformation dominates, but the density of dislocations increases with load (in this zone the EWF decreases with increasing load, as shown in Fig. 19.13). In the third zone, the plastic processes also dominate, but the dislocation concentration do not changes significantly; that is, there is a dynamic process involving the generation and annihilation of dislocations and the creation of micropores and microcracks. Examples of rubbing surface electron work function (EWF), torque and surface temperature, vs. normal load curves for two bronzes and aluminum base alloy are shown in Fig. 19.14.

Studies of a wide range of materials have shown that the transition load-separating zone II and zone III could serve as an objective experimental criterion to estimate the serviceability of the tribological materials. This transition is probably related to the critical transition between mild and severe wear for a given material. The criterion developed is fruitful for developing

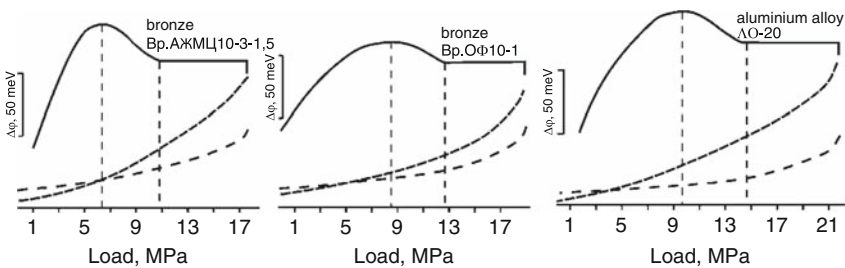


Fig. 19.14. Examples of rubbing surface electron work function, torque, and surface temperature vs. normal load curves for two bronzes (Al, Fe, Zn and Sn, P additives) and aluminum base alloy (used Russian notation)

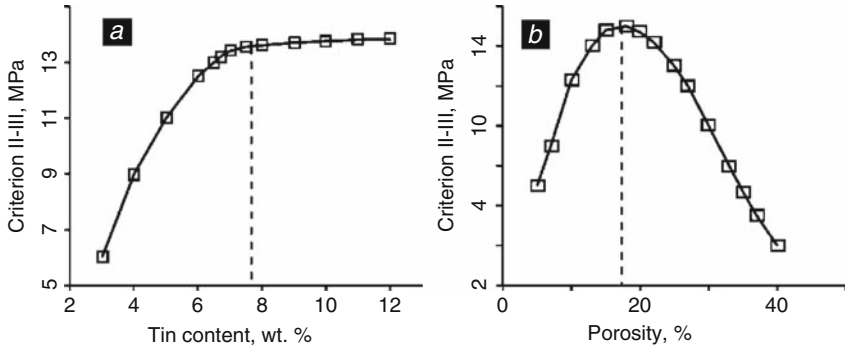


Fig. 19.15. Examples involving powder metallurgy materials (a) and porosity optimization (b)

new materials for tribological applications [6]. Examples involving materials and porosity optimization are provided in Fig. 19.15.

19.5 Periodic Electron Work Function Changes During Friction

19.5.1 Phenomenology

Within the third zone described earlier, the kinetics of steady state friction is characterized by regular periodic changes of the rubbing surface EWF integral value (*see, e.g.,* Fig. 19.11b) [42]. The period, amplitude, and harmonic contents of such changes are determined by the properties of materials and testing conditions [48]. The existence of the EWF periodic changes was confirmed independently at the Ohio State University [49, 50].

The EWF periodic changes were observed most clearly in the case of friction of a 63–37 brass (disk) (Fig. 19.16). The amplitude of EWF periodic changes exceeds 100 meV, and the period is over 30 min. Several thousand passes of the pin on the sample surface occurred over one period. The shape of the EWF periodic changes is close to sinusoidal. The periodic changes were observed in the cases of reciprocated and lubricated friction also.

The studies of a wide range of materials have shown that the form of rubbing surface EWF changes can be divided into the following groups:

EWF does not vary with friction within the limit of measurement errors. This group is represented by materials with low strength properties working under dry friction and most materials working at lubricated friction under low loading. Kasai et al. [50] have found that periodic changes disappear under decreased atmospheric pressure, that is, in vacuum condition.

The rubbing surface EWF changes have a periodic character with a shape close to sinusoidal. This group is represented by single-phase solid solutions on a copper base at dry friction.

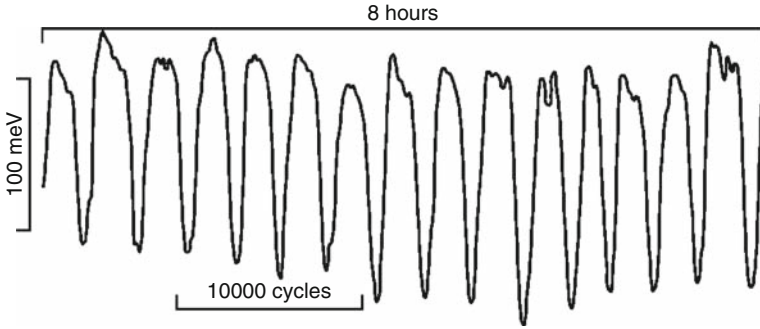


Fig. 19.16. Record of 63–37 brass rubbing surface electron work function periodic changes over 8 h. (Pins were made of phosphor bronze. The contact pressure is 0.12 MPa and the period of disk rotation is 0.42 s^{-1} .)

The rubbing surface EWF changes with a destroyed periodicity; that is, superposition of periodic EWF changes with different periods and amplitudes. This group is represented by multiphase copper-based solid solutions working under dry friction.

The rubbing surface EWF changes have a periodic character with a “saw tooth” shape, that is, the EWF rather slowly increases and quickly decreases. This group is represented by the most cases of steels at dry friction and most of materials under boundary lubrication. A “spike-like” behavior [50] should be included in this group.

Thus, the period, amplitude, and harmonic contents of the EWF changes depend on chemical and phase composition of materials, their mechanical characteristics, and testing conditions. The experiments were carried out on devices that have an integration time that exceeds the time of one revolution of the disk, that is, the EWF value was integrated over the rubbing surface. Therefore, periodic changes of EWF reflect physical processes happening on the whole rubbing surface.

The question is whether the periodic changes of the rubbing surface EWF integral value are damping in time. The statistical analyses have been carried out to examine this and related points.

Data for the statistical analyses were collected during long duration tests (over 50 days) of brass (disc) – phosphoric bronze (pins) at a normal pressure 0.15 MPa under the same conditions. The records of the periodic changes of the rubbing surface EWF integral value were statistically processed and then the results were validated using statistical criteria. From the statistical analyses, it follows that the periodic changes of the rubbing surface EWF integral value have the following properties: (a) they are identical, that is, they have the same statistical distribution during repeated tests under similar testing conditions (Fig. 19.17a), (b) they are nondamping (Fig. 19.17b), (c)

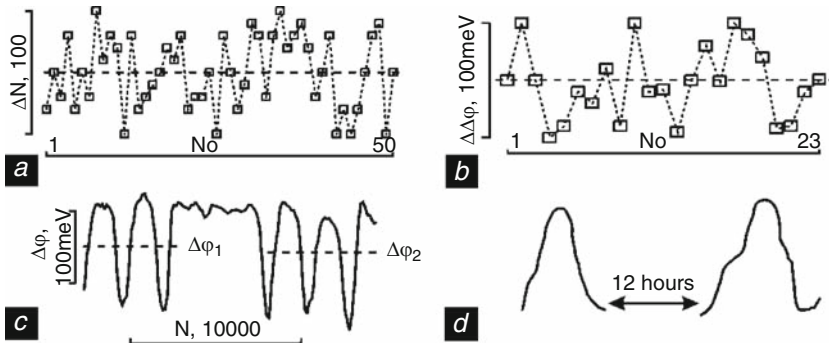


Fig. 19.17. Statistical regularities of rubbing surface work function periodic changes

they are self-restorable (Fig. 19.17c), and (d) they conserve the same phase of period after the interruption and resumption of sliding (Fig. 19.17d).

It follows from statistical studies that the rubbing surface integral value of the EWF periodic changes is stable and nondamping. On the contrary, it implies that there is a synchronization of the periodic changes over the rubbing surface.

The property (d) means that the surface maintains its conditions at the conclusion of sliding. It allows the investigation of surface and subsurface layers in the particular points of the rubbing surface EWF period by means of regular techniques.

Studies of surface and subsurface conditions at four characteristic points during period of EWF change were carried out. These characteristic points are selected: positive and negative peaks and the midpoints of the EWF period increase and decrease. The disks were 4 mm thick and had a diameter 100 mm. They are made of brass 37–63. The pins were made of phosphor bronze. The samples were mounted in the friction-testing machine equipped with the Kelvin probe and tested before they reached a required point of EWF change, and then they were prepared for further studies.

Figure 19.18 shows scanning electron micrographs of the worn surfaces of brass 37–63 at these points. The worn surface has the highest visible “defectness” at the point of the negative peak. On further sliding, the worn surface smoothes out. At the midpoint of the EWF increase (Fig. 19.18b), there are only individual areas containing visible “defectness” similar to those shown in Fig. 19.18a. The surface was smoothest at the positive peak (Fig. 19.18c). Then at the midpoint of the EWF decrease, the separate damaged areas appeared on the worn surface, as can be seen in Fig. 19.18d.

Figure 19.18 also shows the microhardness distribution below the worn surfaces at the above-mentioned four points. The minimum subsurface work hardening is observed at the point of negative peak (Fig. 19.18a). As a result of further sliding, work hardening increases, reaching the values shown in Fig. 19.18b and c. In these three cases, the shape of the microhardness

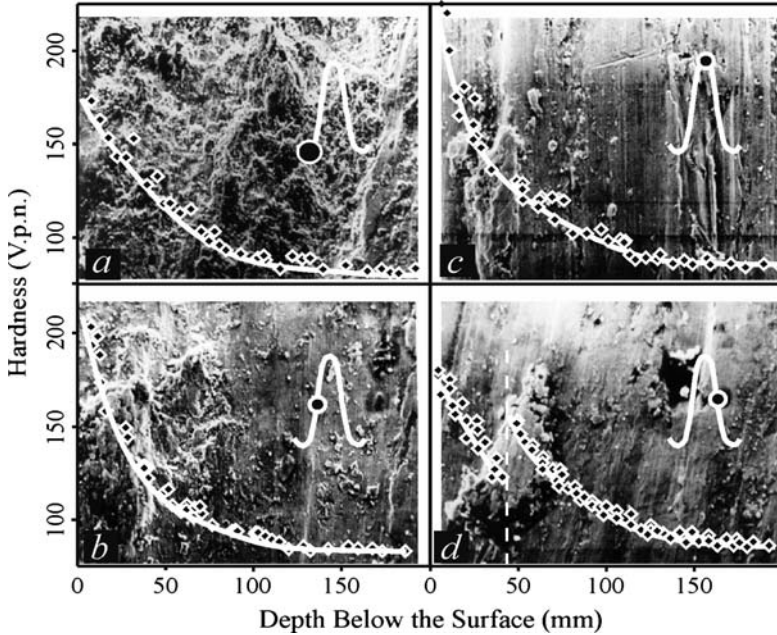


Fig. 19.18. Microphotographs (SEM) and microhardness distribution of brass worn surfaces at four characteristic points based on the period of the work function periodic changes. All have the same magnification ($\times 400$). (a) The negative peak, (b) the midpoint of the work function period increase, (c) the positive peak, (d) the midpoint of the work function period decrease

distribution remains unchanged and is approximated by a power-law curve. The shape of the curve of microhardness distribution changes radically for the midpoint case of decreasing period (Fig. 19.18d). At a depth of more than $\sim 40 \mu\text{m}$, the microhardness distribution is described by the power-law curve, that is, the form of the distribution remains the same as for the cases described earlier. At depths of less than $\sim 40 \mu\text{m}$, the microhardness distribution is linear.

The alloy components distribution below the worn surface was studied by looking at cross-sections (normal to the worn surface) of the specimens corresponding to the same four points. The concentration nonhomogeneities of the alloy components were detected on the specimens corresponding to the midpoint of the decreasing EWF only and at the same depth below the surface where cracks were observed. Figure 19.19 shows distributions of the brass 37–63 components below the worn surface: part (a) at the location of the subsurface crack, and part (b) at the crack-free region. The crack was identified by a simultaneous reduction of the concentration of both alloy components. In the region free of the subsurface cracks and at the same depth where they were observed (approximately $40 \mu\text{m}$), there was a redistribution of alloy components, that is, an increase in zinc and a decrease in copper concentrations.

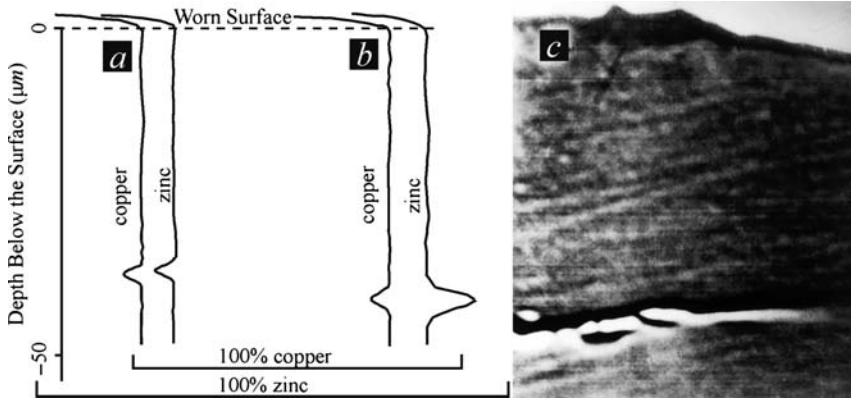


Fig. 19.19. Distribution of brass 37–63 components below the worn surface. (a) At the location of subsurface crack, (b) at the crack-free region, and (c) SEM Zn $K\alpha$ image of the worn surface cross-section at the midpoints of the work function period decrease

The increase in the zinc concentration was significant, reaching $\sim 8\%$. The above observations are additionally illustrated by an SEM Zn $K\alpha$ image of the worn surface cross-section (Fig. 19.19c), where the near crack region looks lighter due to the higher zinc concentration.

The friction tests were carried out using relatively low sliding speed. An increase in zinc concentration was observed significantly below the worn surface. That is why, it is reasonable to suppose that the redistribution of alloy components is not related to thermal effects in the subsurface layers. The behavior can be explained by the transfer of zinc atoms through the crystal defects into the defect accumulation zone.

Similar studies were carried out with specially prepared model copper alloys: Cu + 6%Sn and Cu + 8%Sn. It should be noted that the first alloy belongs to α -solid solutions and the second one contains δ -phase inclusions. Similar to the case of brass, these materials showed cracks parallel to the worn surface at the midpoint of the EWF decrease, but the subsurface cracks in bronze, Cu + 6%Sn, were $\sim 35 - 40 \mu\text{m}$ deep and in Cu + 8%Sn were $\sim 70 - 75 \mu\text{m}$ deep. Figure 19.20 shows panoramic optical micrographs (separate $200\times$ micrographs are combined) of the cross-sections made normal to the worn surface and perpendicular to the sliding direction (here and after, (a) is Cu + 6%Sn and (b) is Cu + 8%Sn). The length of the subsurface cracks is considerable (the width of the micrographs Fig. 19.15a and b are approximately 2 mm). In some cases, very long cracks were observed; they are extended practically to the full width of the friction track (~ 8 mm). The same cross-sections are shown in detail in Fig. 19.21.

The micrographs shown in Figs. 19.20 and 19.21 demonstrate that there is no significant plastic flow in the near surface. This was usually observed

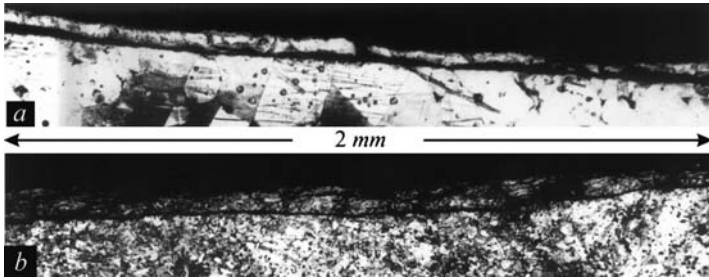


Fig. 19.20. Panoramic optical micrographs (the width corresponds to ~ 2 mm) of the cross-sections made normal to the worn surface and perpendicular to the sliding direction at the prefracture stage. (a) Cu + 6%Sn, (b) Cu + 8%Sn

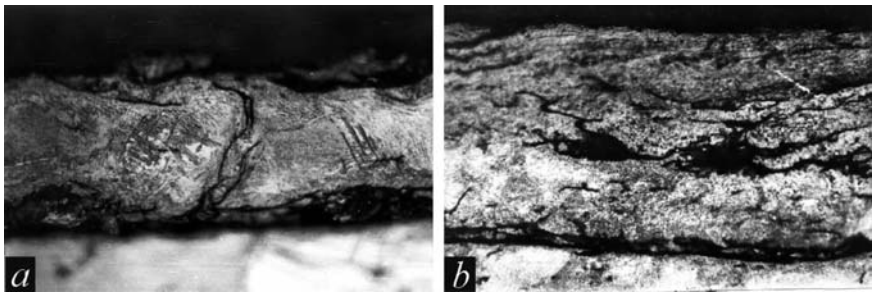


Fig. 19.21. Detailed micrographs of the same cross-sections as in Fig. 19.20 (a) Cu + 6%Sn, (b) Cu + 8%Sn

in numerous subsurface damage studies (*see, e.g.,* [51]). Another feature of the micrographs is that the damaged zone below the worn surface looks like extended system of voids but not like a crack. The metallographic structure of the material band between the surface and damaged zone was not revealed by the metallographic etching; the bulk of material under the band shows a nice metallographic structure. The latter can be explained by a subgrain structure formation in this band. The modification of the microhardness distribution law as it is shown in Fig. 19.18d could be an indirect confirmation of the above observation. (Fig. 19.22)

Samuels et al. [52] have found a generally similar structural arrangement. Their schematic representation of an abraded brass surface includes the subgrain band on the surface (Fig. 19.23a). It is reasonable to extend their schematic representation by adding voids under the subgrain band, as shown in Fig. 19.23b. In this case, stresses appearing on the rubbing surface are transferred to deeper layers by the subgrain band, without dislocation nucleation and accumulation within the individual subgrain. The main processes governing subsurface fatigue are the nucleation of voids and their growth, while crack propagation processes play an insignificant role.

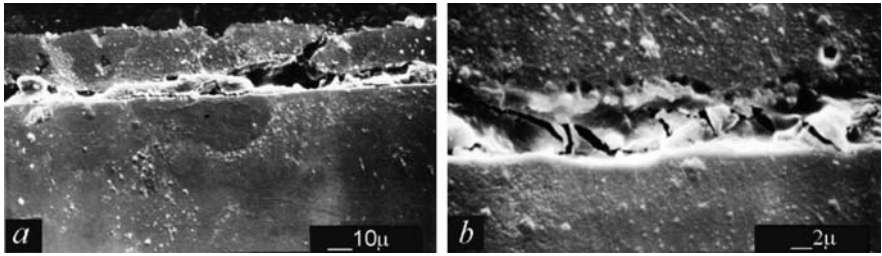


Fig. 19.22. Micrographs of the subsurface features in details (Cu+6%Sn, unetched)

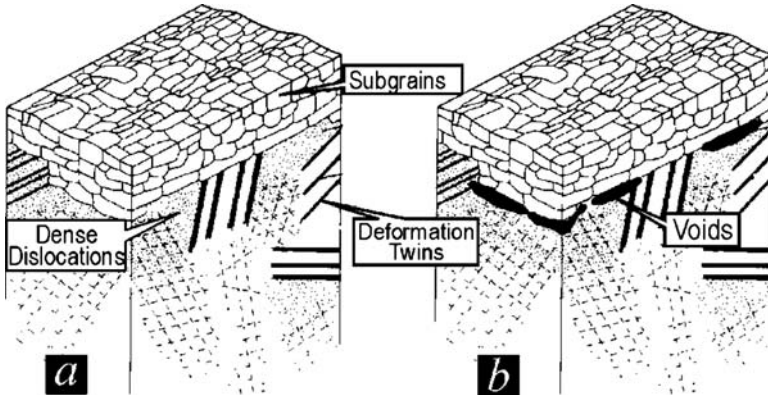


Fig. 19.23. Worn surface schematic representation for the prefracture stage of subsurface fatigue

Two things are responsible for the nucleation of voids and their growth. The first one is the appearance of considerable shear stresses on the boundary between the subgrain band and the bulk material, as, in general, the reordering of near surface dislocations into subgrain boundaries leads to changing of the near surface band volume. The second one is void nucleation and growth. To our minds, the second case is more reasonable as it is indirectly confirmed by the alloy components redistribution (Fig. 19.19) that could be explained by a movement of a number of the dislocations, together with their impurity atoms (zinc), to the accumulation zone.

Let us consider the kinetics of processes in the rubbing surface from the point of view of the subsurface fatigue processes described earlier. Initially the pins smooths the outermost surface layers and creates a surface with certain physical properties and with a certain EWF. Simultaneously, the dislocation density in the subsurface layers grows due to repeated stresses. When the dislocation density reaches a critical value, reordering of the near surface dislocation dense into subgrain boundaries will occur. The reordering process can start when both conditions (certain critical dislocation density and certain critical stress level) are reached simultaneously. That is why, below the

subgrain band, where stresses are not sufficient; dislocation density is continuing to increase without reordering. Further dislocation density increase leads to void nucleation and growth under the subgrain band according to the known mechanisms. For example, in materials with low stacking fault energy like the tested brass, voids can be nucleated from the flat dislocation pile-ups and the dislocation movement can grow the voids [53]. The continuation of the process described earlier can lead to unstable conditions when a relatively thick surface band is almost separated from the bulk material, but still remains under friction because the mechanical properties of the subgrain band are high enough to keep its integrity. The density of defects on the outermost surface layers may remain relatively low due to, for example, dislocation annihilation on the free surface.

The failure of one or several weak spots of the subgrain band will increase the load on the remaining segments of the band. Subsequently, further destruction will behave like an avalanche, that is, the destruction process will be synchronized over a considerable segment of the surface.

After subgrain band destruction and removal, one of the walls of the subsurface voids will become a rubbing surface. The smoothed outermost surface will be replaced with a highly defective surface. This new surface is smoothing by the pins again and its EWF is increasing. In parallel, conditions for the next cycle of the subsurface fatigue are accumulating. By the next cycle, physical properties of the surface change significantly so that the differences in the EWF between the surfaces before and after the fracture become significant. From the point of view of the outermost worn surface, the processes described earlier can be illustrated by the microphotographs in Fig. 19.18. Figure 19.18a shows the surface just after subgrain band removal. After that, the surface layer is smoothed down (Fig. 19.18b and c). Figure 19.18d shows a prefracture stage when spots of destruction appear.

The above model is consistent with Kasai et al.'s [49, 50] observations of metallic and oxidized surfaces at different points of the EWF period.

It has been demonstrated that the harmonic content of the rubbing surface EWF periodic changes depends on the physical properties of the material [54]. In Fig. 19.24, the rubbing surface EWF periodic changes and their spectrums are shown for the case of the single-phase (a) and double-phase (b) bronzes, described earlier. The single-phase material clearly exhibits strictly regular periodic EWF changes and its spectrum has only one maximum. The periodic changes contain additional harmonic and an additional maximum in the spectrum in the case of a double-phase material (b). This additional maximum is apparently due to the difference in the fatigue fracture mechanism between single- and double-phase materials. Double-phase materials have additional obstacles, as secondary phase inclusions keep defects (dislocations) from movement. It can be assumed that the additional process of the fatigue damages accumulation (in secondary phases, solid inclusions, etc.) causes additional harmonics in the spectrums of the EWF periodic changes.

The earlier description can be confirmed also by comparing the microphotographs in Figs. 19.20 and 19.21. Case (a) is related to the single-phase bronze

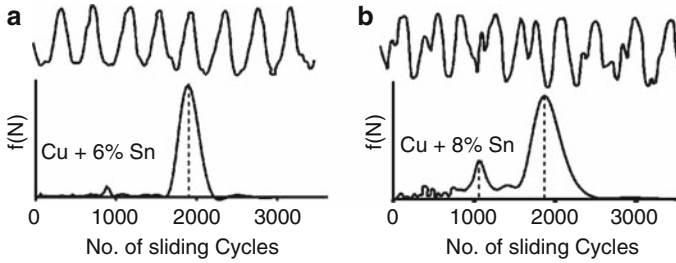


Fig. 19.24. The rubbing surface work function periodic changes and their spectrums. (a) Single-phase bronze, (b) double-phase bronze

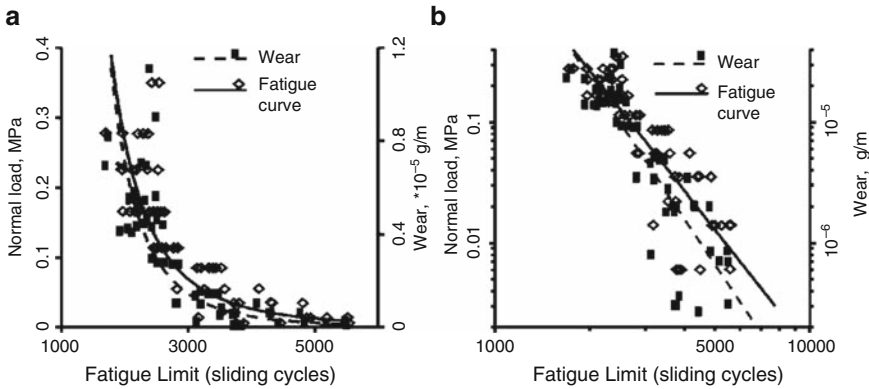


Fig. 19.25. Fatigue curve and wear per cycle evaluated from periodic work function changes of the rubbing surface using normal (a) and logarithmic (b) axes

and case (b) is related to the double-phase bronzes. The band includes some additional minor damage [case (b)], which do not appear in case (a). These results may be considered an experimental confirmation of the fact that several fatigue mechanisms may operate simultaneously in subsurface layers of multiphase materials during sliding.

Additional process of damages accumulation (in secondary phases, solid inclusions, etc.) causes additional harmonics in the spectrums of the EWF periodic changes. Power and layout of the spectral peak correlates to the probability of fracture at given number of sliding cycles.

On the basis of on the fatigue-like nature of the rubbing surface EWF periodic changes integral value, one expects a correlation with known characteristics of bulk fatigue fracture. Curves similar to bulk fatigue curves (Weller’s curves) were plotted using the work function periodicity data (Fig.19.25) [6, 54, 55]. Since it is possible to plot curves similar to those of Weller using data on the rubbing surface periodic EWF changes, approaches used for the conventional fatigue processes can be applied to evaluate fatigue fracture

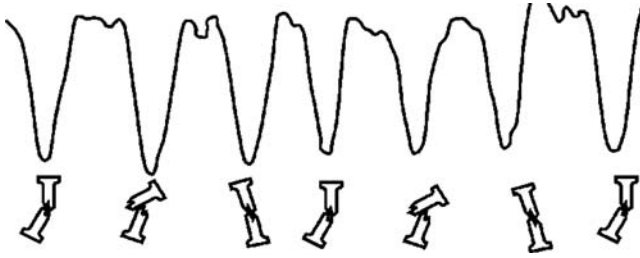


Fig. 19.26. Comparison periodic EWF change and bulk fatigue testing. Every cycle of periodic EWF change corresponds to the broken sample at bulk fatigue testing

parameters for the rubbing surfaces. In contrast, one could evaluate volume fatigue behavior of materials from its testing on friction fatigue (Fig. 19.26).

19.6 Surface Mapping Examples

CPD is an extremely sensitive indicator of surface modification, adsorption, membrane polarization, and redox state, etc. Consequently, CPD is increasingly used in the study of metal, semiconductor, organic semiconductors, biological systems, heterogeneous catalysis, surface reactivity, organic monolayer, corrosion, and tribology. In modern implementations, CPD probes are widely used for surfaces mapping. The CPD arises in turn from the work function difference or, for nonmetals, surface potential difference existing between probe and sample. In the scanning Kelvin probe (SKP), the spatial distribution of CPD values is determined by mechanically scanning the vibrating probe over the sample surface.

Let us classify those SKP techniques employing probe diameters within $1 - 1,000 \mu\text{m}$ as microscopic and those employing probe diameters more than 1 mm as macroscopic. KPFMs typically related to the nanoscale and operate through the detection and zeroing of an oscillating force as described earlier. SKP typically operate through the detection and zeroing of an alternating current produced by probe vibration, although “off-null” measurements are also possible both with vibrating and nonvibrating probes.

Design of the macroscopic SKP in general case is very similar to described earlier for a case of friction surface monitoring. It is not a problem to make a vibrating Kelvin probe with probe diameters above 1 mm because the capacitance of the probe – sample gap – is big enough to get a reliable signal for measurements. On the contrary, the spatial resolution of scanned image 1 mm and even more is good enough for studying around tribological application or surfacing (coating, spraying, etc) on macroscale.

In the case of microscopic SKP, the capacitance of the Kelvin capacitor is becoming much less and subjected to variation due to scanning procedure. To avoid these, lots of efforts were made to fix the probe – sample distance small

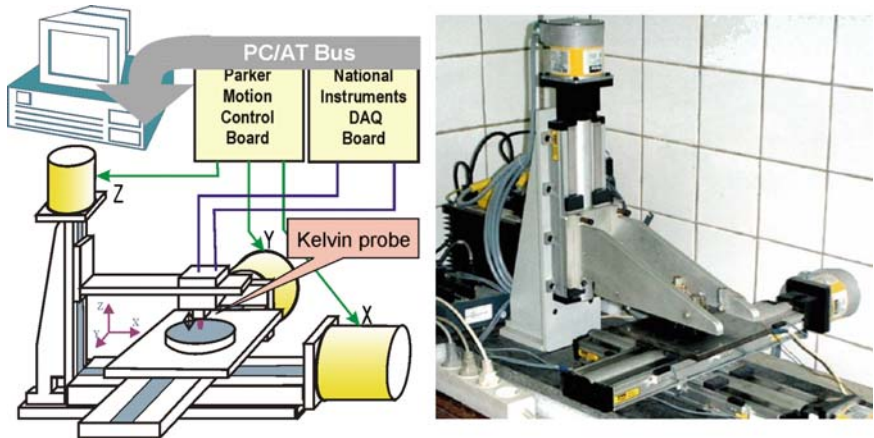


Fig. 19.27. The schematic diagram of the SKP with vibrating probe (2 mm diameter) for macroscopic application functionality and photo of the system

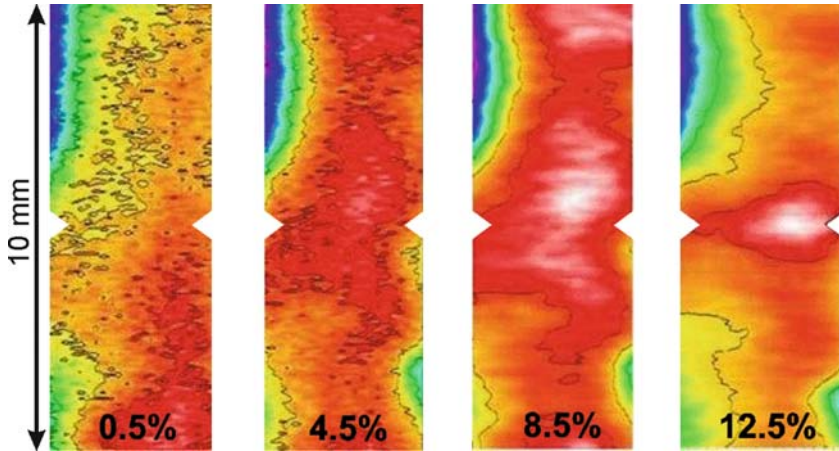


Fig. 19.28. The EWF maps of the lateral surface of the sample with different strain

and constant (see, e.g., [56,57]). Gap regulating systems for the SKP have previously been performed through the use of an additional external modulation voltage over the sample/probe gap. Another possibility that has been used is the analysis of displacement current harmonics, which has subsequently improved by the addition of a compensation voltage.

In Fig. 19.27, the schematic diagram of the SKP for macroscopic applications functionality and photo of the system are shown.

In addition to experiments on the influence of deformation of the EWF for simple cases of loading (Fig. 19.6a), we were mapping lateral surface of the sample. In Fig. 19.28, the EWF maps are shown for different strain. One can see the defects localization before the sample braking.

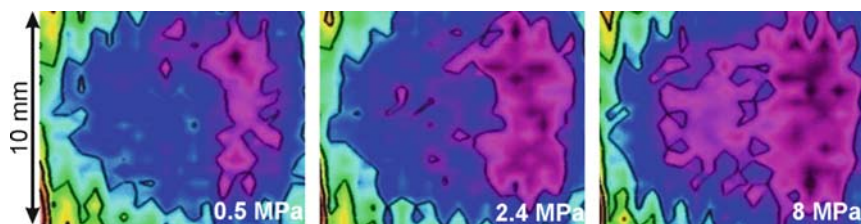


Fig. 19.29. The EWF maps of the contact surface of the sample after different stress applied

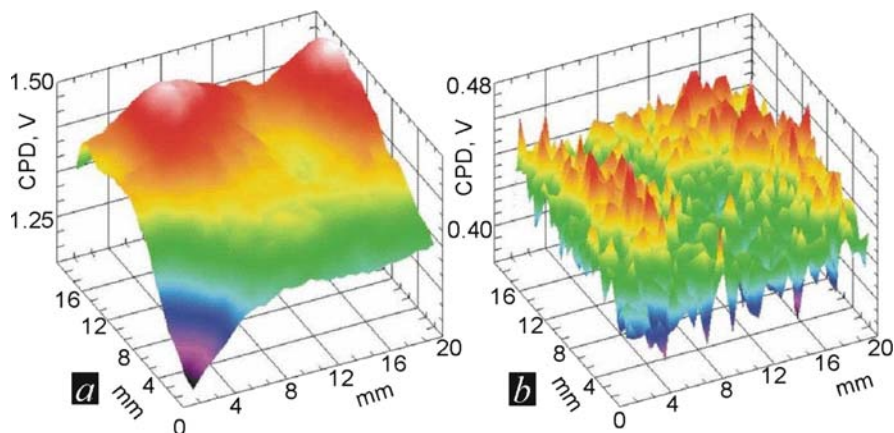


Fig. 19.30. The CPD distribution on the glass surface (a) and on the same surface after spraying of AgCaF (the thickness is 30 nm) (b)

For a case of surface deformation (Fig. 19.6b), the contact surface of the sample was scanned by the SKP as shown in Fig. 19.27. In Fig. 19.29, the EWF maps are shown for different stresses applied. The deformed spot is growing along with the deformation.

The SKP appeared to be a nice tool for homogeneity studies of precision surfacing. In Fig. 19.30 shown glass surface before (a) and after spraying 30 nm of AgCaF layer. In the case (a) the value of the CPD is rather high and, to our mind, is related to surface potential but not a EWF. The sprayed layer is more conductive and the CPD value is compatible with EWF difference range.

A problem related to the SKP arises from the time consumption for the scanning. In a case of traditional SKP, one need to vibrate the probe and adjust the backing potential for zeroing of an alternating current produced by probe vibration. That is why the detail mapping of the macro scale sample (centimeters range) some time requires several hours. The use of translation instead of vibration results in a dramatic improvement in the data acquisition rate and as a result in the time of scanning. The nonvibrating probe described earlier can acquire data at millions of samples per second, making it suit-

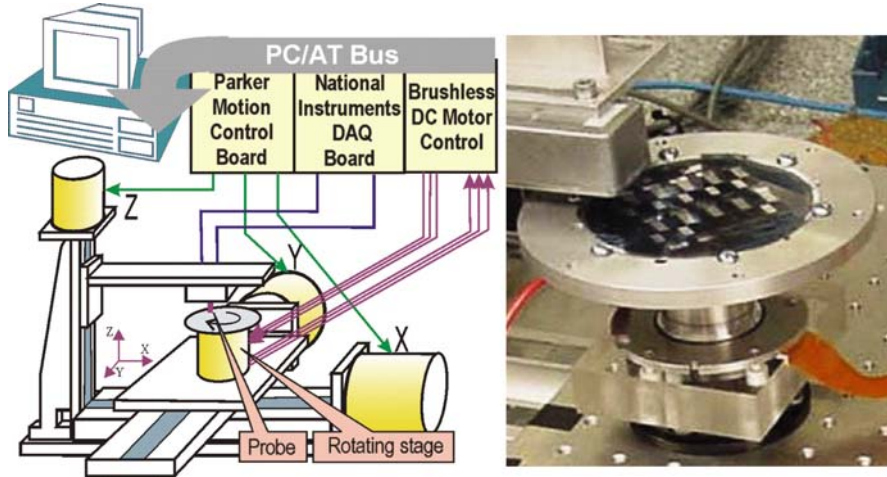


Fig. 19.31. The schematic diagram of the system for surfaces mapping with nonvibrating probe (a) and photo of the rotating stage (b)

able for high-speed mapping applications. In fact, the faster relative motion between the nonvibrating probe and the sample results in an increase in signal strength. The nonvibrating probe with a needle as a reference electrode (estimated diameter is less than $10\ \mu\text{m}$) has a capability to map the macroscale samples for a minutes or less.

The schematic diagram of the system for surfaces mapping with nonvibrating probe is shown in Fig. 19.31. In the system, the same as in Fig. 19.27, XYZ stage was used with an additional rotation stage, as regular linear actuators cannot provide a high enough speed of sample's movement for nonvibrating probe operation. During rotation of the sample, the stream of probe output signal together with synchronizing pulses was acquired and plotted by a computer.

As an example of surface mapping using the nonvibrating probe in Fig. 19.32, the maps of two metallic samples with good polishing and the diamond-like coating are shown. In spite of precision surfacing, the potential distribution shows some nonhomogeneous regions. It is interesting to point out a fingerprint resolved, which is under the coating.

From the point of serviceability, the spots of a contrast on the potential maps could be a future corrosion nucleation centers.

19.7 Closure

The described technique is applicable for investigations and analysis of a wide range of processes involving rubbing surfaces, including running-in processes, formation and depletion of lubricating layers, responses of a surface to variations of friction conditions, fatigue processes, etc.

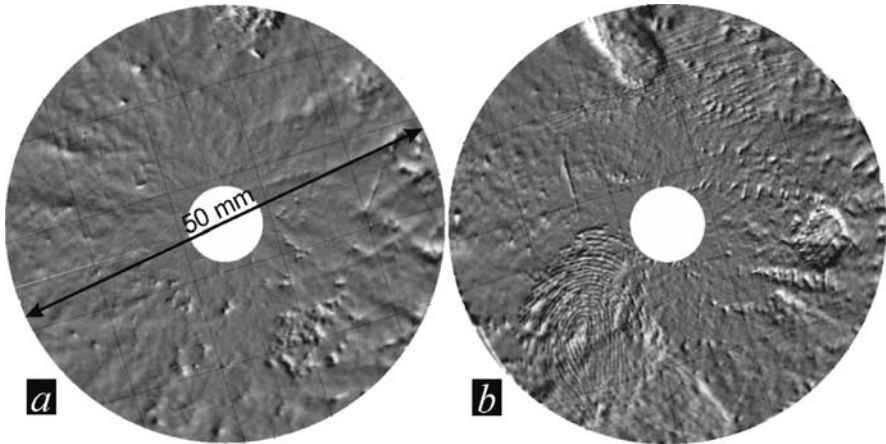


Fig. 19.32. The potential maps of two metallic samples with good polishing and the diamond-like coating

CPD techniques are practically unique in its sensitivity, versatility, and simplicity for practical implementation.

Acknowledgements

I would like to acknowledge the contributions of my colleagues Nicolay Shipitsa, Efim Fishbein, and Brendon Steele. I would like to thank Steve Danyluk and David Rigney for support and collaboration.

References

1. Lord Kelvin, Contact electricity of metals, *Phil. Mag.* **46**, 82–120 (1898).
2. J. Friedel, The physics of clean metal surfaces, *Ann. phys.* **1**(6), 257–307 (1976).
3. G. Mahan, W.L. Schaich, Comment of the theory of work function, *Phys. Rev.* **B10**(6), 2647–2654 (1974).
4. S. Yamamoto, K. Susa, U. Kawabe, Work function of binary compounds, *Japan J. Appl. Phys.* **2**, 209 (1974).
5. S.I. Pekar, O.F. Tomasevich, Thermionic emission from metals, covered by thick semiconductor layer, *USSR J. Tech. Phys.* **17**(12), 1339–1342 (1947).
6. A.L. Zharin, *Contact Potential Difference Technique and Its Application in Tribology*, (Bestprint, Minsk, 1996) (in Russian).
7. N.V. Cherepnin, *Sorption Phenomenon in Vacuum Technique*, (Soviet Radio, Moscow, 1973), (in Russian).
8. A.A. Andreev, Ia. Polige, Work function change under cold deformation of molybdenum and tungsten in ultrahigh vacuum, *Proc. USSR Acad. Sci.* **152**(5), 1986–1088 (1963) (in Russian).
9. P. Craig, Direct observation of stress - induced shifts in contact potentials, *Phys. Rev. Lett.* **22** (1969) 14 700.

10. E.M. Gutman, *Mechanochemistry of Metals and Corrosion Protection* (Metallurgy, Moscow, 1974) (in Russian).
11. A.N. Latishev, M.I. Molotski, K.V. Chibisov, An interaction of the chemisorbed particles with dislocations, *Proc. USSR Acad. Sci.*, **224**(4), 880–882 (1975) (in Russian).
12. R.I. Mints, V.P. Melekchin, M.B. Partenski, Exoelectrons emission relation with work function in metals, *USSR J. Metal Phys.*, 40(4), 886–889 (1975) (in Russian).
13. D. DeVecchio, B. Bhushan, Use of a nanoscale Kelvin probe for detecting wear precursors, *Rev. Sci. Instrum.* **69**, 3618–3624 (1998).
14. J.A. Chalmers, Contact potentials, *Phil. Mag.* **33**, 399–430 (1942).
15. W.A. Zisman, A new method of measuring contact potential difference in metals, *Rev. Sci. Instrum.* 367–370 (1932).
16. H. Palevsky, R.K. Swank, R. Grenchik, Design of dynamic condenser electrometer, *Rev. Sci. Instrum.* **18**, 297–314 (1947).
17. R. Simon, Work function of iron surfaces produced by cleavage in vacuum, *Phys. Rev.* **116**(3), 613–617 (1959).
18. S. Danyluk, K. Hamall, L.A. Reid, A.L. Zharin, The non-vibrating capacitance probe for wear monitoring, (2007), US patent RE39,803 E.
19. E. Zanoria, K. Hamall, S. Danyluk, A.L. Zharin, Surface wear monitoring with a non-vibrating capacitance probe, *J. KSTLE* **11**, 40–46 (1995).
20. E. Zanoria, K. Hamall, S. Danyluk, A.L. Zharin, The non-vibrating Kelvin probe and its application for monitoring surface wear, *J. Test. Evaluation*, *JTEVA* **25**(2), 233–238 (1997).
21. E. Zanoria, S. Danyluk, C.S. Bhatia, A.L. Zharin, Kelvin probe measurements of wear of a magnetic hard disk, *Adv. Inform. Storage Syst.* **7**, 181–191 (1996).
22. T.J. Harvey, S. Morris, L. Wang, R.J.K. Wood, H.E.C. Powrie, Real-time monitoring of wear debris using electrostatic sensing techniques. *Proc. IMechE Part J: Eng Tribol.* **221** (2007).
23. Y. Martin, D.W. Abraham, H.K. Wickramasinghe, High-resolution capacitance measurement and potentiometry by force microscopy, *Appl. Phys. Lett.* **52**, 1103–1105 (1988).
24. M. Nonnenmacher, M.P. Oboyle, H.K. Wickramasinghe, Kelvin probe force microscopy, *Appl. Phys. Lett.* **58**(25), 2921–2923 (1991).
25. J. Lu et al., Surface potential studies of self-assembling monolayers using Kelvin probe force microscopy, *Surf. Interface Anal.* **27**(5–6), 368–373 (1999).
26. Y.L. Yousef, S. Mischriki, S. Aziz, Measurement of contact potential by electrostatic excitation of low-frequency vibration, *J. Phys. E.* 12, 873–875 (1965).
27. B. Bhushan, A.V. Goldade, Kelvin probe microscopy measurements of surface potential change under wear at low loads, *Wear* **244**, 104–117 (2000).
28. B. Bhushan, A.V. Goldade, Measurements and analysis of surface potential change during wear of single-crystal silicon at ultralow loads using Kelvin probe microscopy, *Appl. Surf. Sci.* **157**, 373–381 (2000).
29. B. Bhushan, Nanotribology and nanomechanics, *Wear* **259**, 1507–1531 (2005).
30. B. Bhushan, M. Palacio, B. Kinzig, AFM-based nanotribological and electrical characterization of ultrathin wear-resistant ionic liquid films, *J. Colloid. Interf. Sci.*, **317**, 275–287 (2008).
31. V.S. Fomenko, *Handbook of Thermionic Properties* (Plenum, New York, 1966).

32. W. Li, D.Y. Li, In situ measurements of simultaneous electronic behavior of Cu and Al induced by mechanical deformation. *J. Appl. Phys.* **99**(7), 2005–2012 (2006).
33. W. Li, M. Cai, Y. Wang, S.I. Yu, Influences of tensile strain and strain rate on the electron work function of metals and alloys. *Scripta Materialia* **54**(5), 921–924 (2006).
34. A.L. Zharin, V.A. Genkin, E.I. Fishbein, N.A. Shipitsa, A.L. Terekhov, E. Barkun, Determination of contact deformation mode from the electron work function, *Soviet J. Friction Wear* **11**, 144–146 (1990).
35. A.L. Zharin, E.I. Fishbein, N.A. Shipitsa, Effect of contact deformation upon surface electron work function, *Soviet J. Friction Wear* **16**(3), 66–78 (1995).
36. Vishniakov Ia. D., *Modern technique for investigation of deformed crystal structure*, (Metallurgy Press, Moscow, 1975) (in Russian).
37. G.V. Dydko, Contact potential difference measurements by condenser technique, *USSR J. Exp. Tech.* **5**, 128–130 (1961).
38. U.V. Nazarov, B. Postagonov, G.I. Geigopov, N.V. Domashka, The basis of nanotechnology, *Russian Proc. Mashinconstruction* **1**, 29–31 (1990) (In Russian).
39. A.A. Markov, *Electron work function changes during friction, in Electric phenomena during friction, cutting and lubrication of solids.* (Nayka, Moscow, 1973) 28–34 (in Russian).
40. W. Li, D.I. Li, Exploring the application of the Kelvin method in studying the history prior to wear and the onset of wear”, *Wear* **253**(7), 746–751 (2000).
41. A.L. Zharin, G.P. Shpenkov, *Device for Friction Pair Monitoring*, (1978) USSR Patent no. 615379.
42. A.L. Zharin, G.P. Shpenkov, Macroscopic effects of delamination wear, *Wear* **56**, 309–313 (1979).
43. A.L. Zharin, Techniques of friction monitoring, *Soviet J. Frict. Wear* **14**, (3) 111–120 (1993).
44. A.L. Zharin, D. Rigney, Application of the contact potential difference technique for on-line rubbing surface monitoring (review), *Tribol. Lett.* **4**, 205–213 (1998).
45. A.L. Zharin, V.A. Guenkin, Study of friction processes with reciprocating displacement, *Soviet J. Frict. Wear* **11**, 128–131 (1990).
46. A.L. Zharin, *Application Macro- and Micro Kelvin Probe in Tribological Studies In book: Fundamentals of Tribology and Bridging the Gap Between the Macro- and Micro/Nanoscales*, (Kluwer, Netherland, 2001) pp. 445–466.
47. A.L. Zharin, V.A. Genkin, E.I. Fishbein, N.A. Shipitsa, A.L. Terekhov, “Method for Run-in of Friction Assembly Materials, *Soviet J. Frict. Wear* **10**, 530–534 (1989).
48. A.L. Zharin, V.A. Guenkin, On rubbing surface electron work function periodicity, *Soviet J. Frict. Wear* **2**(1), 91–95 (1981).
49. T. Kasai, D. Rigney, A.L. Zharin, Changes detected by a non-contacting probe during sliding, *Scripta Mater.* **39**, 561–567 (1998).
50. T. Kasai, X. Fu, D. Rigney, A.L. Zharin, Application of a non-contacting Kelvin probe during sliding, *Wear* **225–229**, 1186 (1999).
51. A. Alpas, H. Hu, J. Zhang, Plastic deformation and damage accumulation below the worn surface, *Wear* **162**, 188–195 (1993).
52. L.E. Samuels, E.D. Doyle, D.M. Turley, *Fundamentals of Friction and Wear of Materials*, ASM Materials Science Seminar, 13 (1980).

53. J. Hirt, I. Lote, *Theory of Dislocations* (Atomizdat, Moscow, 1972) (in Russian).
54. A.L. Zharin, N.A. Shipitsa, E.I. Fishbein, Some features of fatigue at sliding friction, *Soviet J. Frict. Wear* **14**(4), 13–22 (1993).
55. A.L. Zharin, V.A. Guenkin, O.V. Roman, Connection of periodic changes of electron work function of a rubbing surface with fatigue damage, *Soviet J. Frict. Wear* **7**(2), 112–120 (1986).
56. I.D. Baikie, P.J.S. Smith, D.M. Porterfield, P.J. Estrup, Multitip scanning bio-Kelvin probe. *Rev. Sci. Instr.* **70**(3), 1842–1850 (1999).
57. K. Wapner, B. Schoenberger, M. Stratmann, G. Grundmeier, Height-regulating scanning Kelvin probe for simultaneous measurement of surface topology and electrode potentials at buried polymer/metal interfaces. *J. Electrochem. Soc.* **152**(3), E114–E122 (2005).
58. H. Rivere, *Work Function. Measurements and Results, in Solid State Surface Science* **1**, (Dekker, NY. 1969).

Part III Industrial Applications

Modern Atomic Force Microscopy and Its Application to the Study of Genome Architecture

Kunio Takeyasu, Hugo Maruyama, Yuki Suzuki, Kohji Hizume,
and Shige H. Yoshimura

Summary. Recent development of atomic force microscopy (AFM) has been accomplished by various technical and instrumental innovation including high-resolution imaging technology in solution, fast-scanning AFM, and general methods for cantilever modification and force measurement. These modern AFM technologies have made it possible to conduct biological studies under physiological conditions. Application of the recognition imaging mode that can simultaneously obtain a topographic image together with a recognition signal is now successful by using protein-(antibody-) coupled cantilever, and revealed the specific protein bindings on the chromatin. AFM can also be combined with biochemical and cytochemical methods. Recent AFM researches involving series of reconstitution experiments have shown that the efficiency of the chromatin reconstitution by salt-dialysis method is drastically increased simply by using longer (>100 kb) and supercoiled DNA. This suggests that the physical properties of DNA are critical for the higher-order chromatin folding. Since double-stranded DNA, like other polymer chains, carries certain elasticity and flexibility, the length of DNA could affect the stability of nucleosome and chromatin fiber. This notion is well supported by the fact that in eukaryotic chromosome, the averaged length of a single chromatin loop is ~100 kb. On the contrary, the large-scale structure of chromatin fiber is affected by a local protein binding. Indeed, histone H1 is essential for the reconstitution of 30 nm fibers. Type II topoisomerase (Topo II) has been known as a major component of chromosomal scaffold and an essential protein for mitotic chromosome condensation. AFM has shown that Topo II binds to bear DNA and clamps two DNA strands even in the absence of ATP, and promotes chromatin compaction depending on the existence of histone H1. Namely, H1-induced 30-nm chromatin fibers were converted into large complex by the effect of Topo II. On the basis of these results, a chromatin packing model triggered by Topo II that clamps DNA strands can be proposed. AFM analyses have also shown the similarities and differences between the eukaryotic and prokaryotic genome organizations. In spite of the presence of different structural proteins, the higher-order stepwise hierarchies from 30 nm fiber to 80 nm beaded structure are shared in eukarya, bacteria, and archaea, but the fundamental structural units are very different, that is, nucleosomes in eukarya and some archaea, and nonnucleosome unit in bacteria.

Key words: Archaea, Atomic force microscopy (AFM), Bacteria, Carbon nanotube, Fast-scanning AFM, Force measurement, Genome architecture, Nucleosome, Recognition imaging.

20.1 Introduction: History of AFM Applications to Biological Macromolecules

The application of atomic force microscopy (AFM) to biological samples dates back to the late 1980s (Fig. 20.1). The first notable application was made in the early 1990s, for the observation of double-stranded DNA [1–4]. This achievement greatly encouraged many biological researchers to jump into nanoworld in the late 1990s. Here we review the history of how AFM has been utilized to reveal nanoscale structures of DNA–protein complexes, and what types of technical developments have accelerated the advancement of the application.

20.1.1 Nanometer Scale Imaging of DNA–Protein Complexes

Various DNA–protein complexes have subsequently been analyzed. AFM is routinely used to identify the binding site of a DNA-binding protein, with a resolution of several tens of base pairs, and to determine the degree of protein polymerization on linear DNA of several kilobase pairs [5, 6]. The higher-order structures of DNA including supercoiling [7–10], stem-loop structures [11], and enhancer-promoted DNA loops [12, 13] as well as nucleosomes and chromatin [14–22] have also been elucidated.

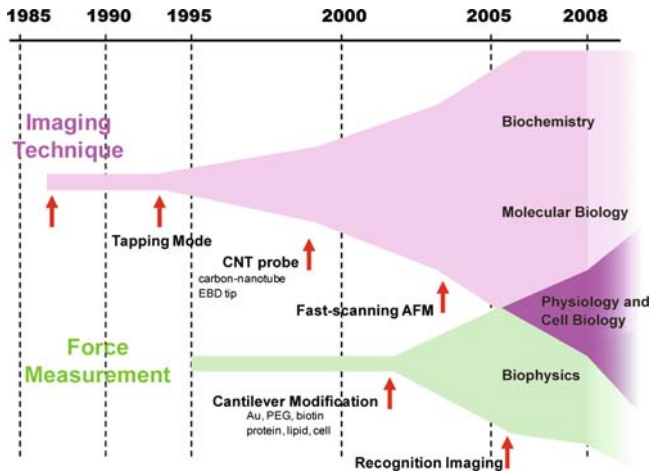


Fig. 20.1. History of AFM in biology. AFM was developed in 1986. The development of AFM for biological samples (imaging and force measurements) has been accelerated by several important technical improvements (vertical arrows). Now it is used in various fields of biology

From the late 1990s to the present, higher-order architectures of DNA–protein complexes, chromatin, and chromosomes have been dissected by AFM. A combination of biochemical reconstitution of nucleosome fibers and AFM observation elucidated the molecular basis of nucleosome dynamics, nucleosomal fiber folding, and chromatin condensation [14–18, 20, 23, 24]. The addition of linker histone H1 mediates the formation of 30-nm fibers [15], and further addition of various chromosomal proteins causes compaction of the chromatin fibers [23, 24] (see Sect. 20.3). In addition to these structural studies, there have been several reports on the function of various chromatin-modifying enzymes. The effects of chromatin-remodeling factors [25, 26] and histone poly(ADP-ribose) polymerase [27] on chromatin structure have been investigated by AFM.

Along with the biochemical reconstitution approach, the chemical dissection of native chromosomes and nuclei, followed by AFM observation, has revealed various higher-order architectures of chromatin fibers. The sequential removal of the cell membranes, cytoplasm, and nucleoplasm from cultured cells on a coverslip leads to the exposure of interphase chromatin [28]. Using this method and nanometer-scale imaging, the chromatin structures inside the interphase nucleus have been revealed [29]. The depletion of cytoplasmic and nucleoplasmic materials from HeLa cells exposes the interphase chromosome, composed of an ~80-nm-diameter granular unit that forms an ~80-nm-wide beaded fiber. In contrast, cytoskeletal fibers and bundles are always thinner, with a width of 10–60 nm.

20.1.2 Visualization of Various Biological Macromolecules

AFM has been used to visualize biological molecules other than DNA, including nuclear pore complex, membrane receptors, and molecular chaperones. The nuclear pore complex (NPC) is a large protein complex in the nuclear envelope of eukaryotic cells. The nuclear envelope from *Xenopus* oocyte is extensively used in structural studies of the NPC because of its large size (ϕ 0.1 μ m). Observation of the nuclear envelope by AFM revealed the architecture of NPC as well as the network of nuclear lamina. It has been reported that the NPC changes structure depending on the intracellular Ca^{2+} concentration [30–32], and on other intracellular environmental factors [33, 34] and extracellular stimuli [35, 36].

The visualization of membrane proteins had been extensively studied using trimeric bacteriorhodopsin in purple membrane [37]. In combination with force measurements (see Sect. 20.1.3), the force required to unfold the membrane protein and pull the polypeptide out of the lipid bilayer was measured [38]. Other membrane proteins studied by AFM include F-type ATPase [39, 40], Rho in native disc membrane [41], the cyanobacterial photosystem [42], and S-layer protein [43].

The detailed subunit arrangement of multi-subunit protein complexes has also been extensively studied by AFM, although extremely high resolution

is required. For example, the subunit arrangement of yeast condensin complex was revealed by structural comparison of subcomplexes [44]. In another example, the subunit stoichiometry of a multi-subunit membrane protein was elucidated by a novel method using AFM and antibody [45]. Binding of the antibody to a specific subunit clearly indicated the arrangement of the subunits in the complex.

20.1.3 Challenges Toward Technical Advancement

The spatial resolution of AFM depends mainly on the sharpness of the probe tip. Significant improvement in spatial resolution was achieved in the late 1990s. Attachment of a carbon nanotube to the tip of the AFM cantilever dramatically improved the spatial resolution [46, 47], allowing the twisting of two DNA double strands to be clearly visualized [47] (Fig. 20.2). Since the attachment of the nanotube was performed by manual handling of both the nanotube and the AFM cantilever in a scanning probe microscope (SEM)

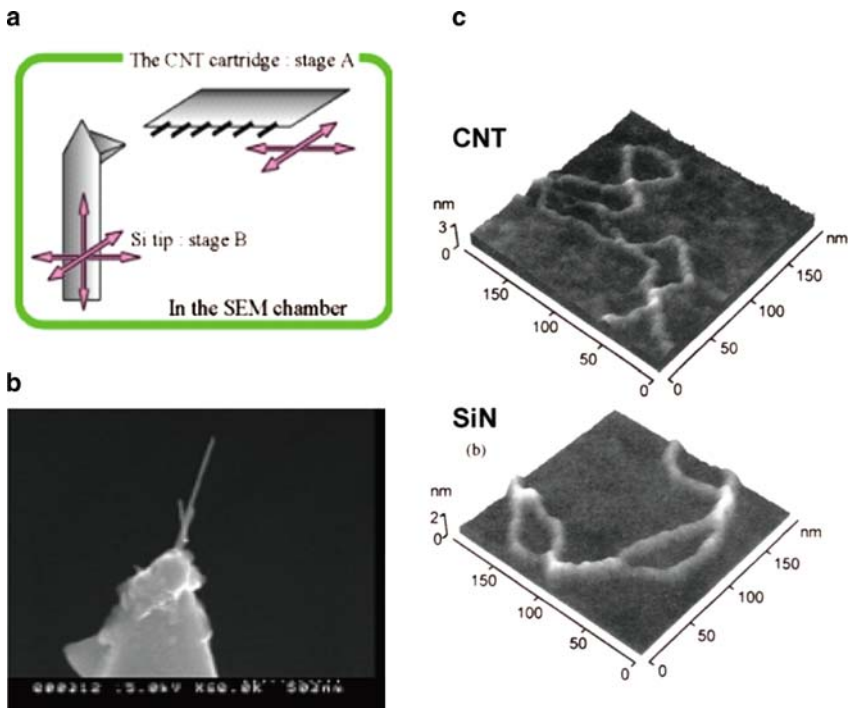


Fig. 20.2. Development of the carbon nanotube probe. (a) A carbon nanotube is transferred from a knife edge to the tip of an AFM probe in a SEM chamber. (b) A SEM image of a nanotube probe (c) DNA image obtained with a nanotube probe (*upper*) or a normal silicon cantilever (*bottom*). Only the nanotube probe can clearly visualize two DNA strands crossing over

chamber, the yield was very low. Improvements in the attachment process have allowed the commercial production of nanotube probes. Further improvement in spatial resolution has resulted from the use of noncontact AFM in solution [48], allowing nanometer resolution in visualizing the GroEL septamer and bacteriorhodopsin.

Improvement of the time resolution had been an important issue in biological applications of AFM. The movements of RNA polymerases and restriction enzymes were observed by time-lapse imaging by AFM in solution in the late 1990s, with frame rates of several seconds. An extraordinary improvement in the device was made by Ando's group at Kanazawa University in 2001 [49]. The development of "fast-scanning AFM" enabled the analysis of biological process at nanometer and subsecond resolution. Applications of this device include the kinetic analysis of protein conformational changes [50], visualization of the enzymatic reaction of restriction endonucleases [51], and the dynamic behavior of chromatin fibers (*see* Sect. 20.2.1). A dissection of the conformational changes of a molecular chaperone, GroEL/ES, by fast-scanning AFM has shown that GroEL has at least two distinct open-conformations in the presence of nucleotide: an ATP-bound prehydrolysis open-form and an ADP-bound posthydrolysis open-form. In addition, the ADP-bound open-form is less stable than ATP-bound open-form and can easily be converted to closed state by releasing the ADP [50].

Several recent biological applications of AFM include the measurement of surface physical properties of a sample, and the interactive forces between biological macromolecules or between domains within a molecule [52]. When the cantilever pushes down on a sample surface or pulls one of two interacting molecules up, it deflects upward or downward depending on the force applied to it. This force is proportional to the deflection: $F = k\Delta x$, where k and Δx are the spring constant of the cantilever and the deflection, respectively. Thus, by measuring the cantilever deflection, interactions between biological molecules can be characterized. The application of single-molecule force measurements is described in Sects. 20.2.2 and 20.2.3.

20.2 Trends in Biological AFM

High-resolution AFM imaging in solution has become routine, and even dynamics of biological macromolecules can be monitored under a fast-scanning AFM. Force measurement between proteins of interest can be done under "physiological conditions." Single-cell manipulation techniques have also been developed with the applicability to physiological and pathological studies.

20.2.1 Analyses of Biological Macromolecules in Motion

The development of tapping mode AFM in liquids has made it possible to study structural and mechanical properties of biological molecules under

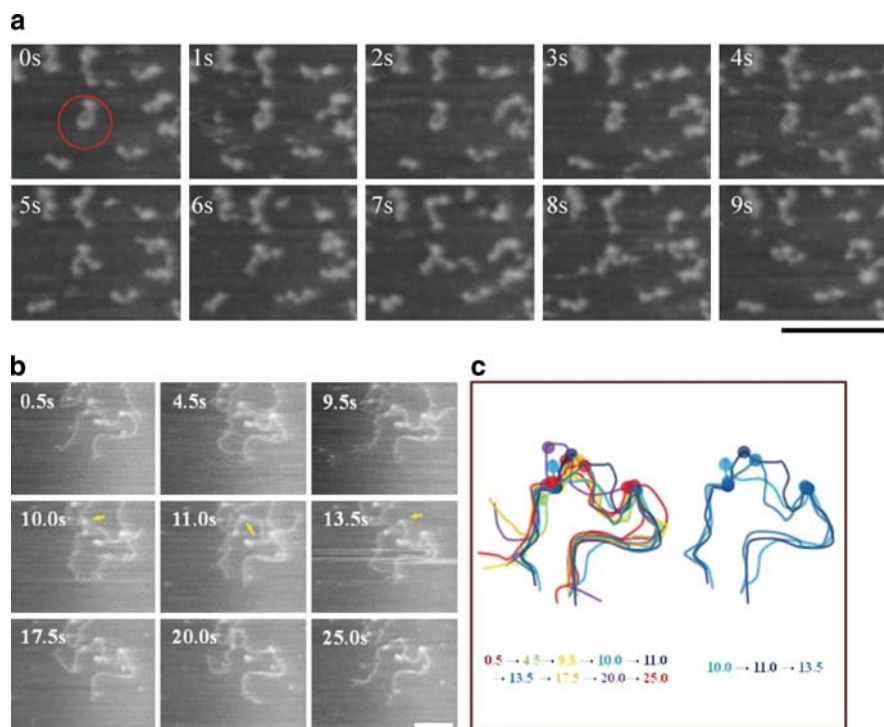


Fig. 20.3. Dynamics of biomacromolecules revealed by fast-scanning AFM. (a) Time-lapse images of “open-and-close” behavior of *Bacillus subtilis* RecN [structural maintenance of chromosomes (SMC)-like protein] obtained in the absence of ATP at 1 frame per second. The time of observation (in s) is shown at the *upper left corner*. The focused molecule is indicated by a red circle. (b) Time-lapse images of nucleosome sliding on reconstituted chromatin fiber obtained at 2 frames per second. The time of observation (in s) is shown at the *upper left corner*. The nucleosome that shows high mobility on mica is indicated by yellow arrows. The images are cropped from an original scan size of 800×600 nm. Scale bar: 100 nm. (c) Movement of nucleosomal arrays on mica. The nucleosomal arrays were traced and depicted in different colors. The overlaid images are shown

“physiological conditions” [53, 54]. However, such investigations have been restricted by the inherent slow scan rate necessary for nanoscale resolution. The scanning rates of commercially available AFMs are usually several seconds to minutes per frame, which is too slow to monitor many physiological processes. Newly developed fast-scanning AFMs have a miniaturized cantilever and scan stage to reduce the mechanical response time of the cantilever and to prevent the onset of resonant motion during high speed scanning [49, 55]. Fast temporal resolutions of 1–3 frames per second allow the dynamics of biomolecules to be followed more closely on the subsecond timescale [49–51, 56] (Fig. 20.3).

Using fast-scanning AFMs, the dynamic behavior of enzyme–DNA complexes in solution can be visualized at the single molecule level in subseconds [51]. A number of DNA-modifying enzymes require communication between two spatially separated sites on the DNA molecule for their activity [57]. These sites can come into contact with each other by 3D diffusion-mediated chance interaction between enzymes bound at the two sites or by active translocation of the intervening DNA by a site-bound enzyme. In the latter case, the translocated DNA would be extruded into an expanding loop. The restriction enzyme EcoP15I is a type III restriction enzyme that requires the specific asymmetric recognition sites of two of these enzymes to be in a head-to-head orientation for DNA cleavage [58, 59]. EcoP15I was previously thought to function by ATP-dependent translocation along the DNA, between the sites accompanied by the extrusion of loops. However, several studies on type III restriction enzymes reached the conflicting conclusion that no loops were formed [60]. Time-lapse imaging by fast-scanning AFM would permit direct imaging of the enzymatic reaction and test the proposed mechanical model. An analysis of a single EcoP15I–DNA complex demonstrated that the intervening DNA loop in the complex can be translocated unidirectionally in an ATP-dependent manner, forming an extruded loop and accompanied by diffusive loop formation [51].

Application of fast-scanning AFM to enzymatic reactions has provided new insights into the mechanisms of protein–DNA interactions. These results raise the possibility that many other biological processes could be analyzed in real time with high temporal and spatial resolution. However, success with time-lapse analysis by AFM requires conditions that satisfy two conflicting requirements; (1) attachment of the specimen molecules of interest onto the substrate surface and (2) freedom of movement of the molecules. To image the dynamics of biomolecules, the molecules must retain enough mobility to allow their functional movement. At the same time, the molecules must be stably adsorbed to the mica surface and remain within the scanning area during imaging by the AFM tip. Fast-scanning AFM can capture an entire image about 1,000 times faster than a conventional instrument, so the requirement for sample immobilization is relaxed. However, the necessity to bind the sample to the mica surface cannot be eliminated.

20.2.2 Measurement of Pico-Newton Mechanical Forces in Biological Systems

The force measurement mode has been used to measure the elasticity of living cells [61–65]. When an AFM cantilever approaches and pushes against the cell surface, a large indentation in the cell and its surface is usually observed when the probe first contacts the cell surface. This indentation can be plotted against the force of the cantilever and fitted to the Hertz model equation [66] to estimate the Young’s modulus, which describes the elasticity of the sample. The Young’s moduli of areas over the nucleus (1.8 ± 1.5 kPa) are

smaller than those of the cell peripheries (5.1 ± 2.9 kPa). The actin network may be responsible for the elasticity of the cell [67], that is, the degradation of cellular actin by cytochalasin D leads to a decrease in the measured Young's modulus [68, 69]. Immunostaining of actin filaments in HeLa cells indicated that these filaments mainly exist in the peripheral areas of the cell. Only a few filaments were observed over the nucleus [70]. Thus, the distribution of actin might be one reason why areas over the nucleus appear to be softer than peripheral areas. Elasticity measurements have shown that both the plasma membrane and the nuclear envelope are "flexible" enough to absorb a large deformation formed by an AFM probe. Penetration of the plasma membrane and the nuclear envelope are possible when a probe with a sharp tip (tip angle of $\sim 25^\circ$) deeply indents the cell membrane, causing the membrane to come close to a hard glass surface [65]. These types of experiments will provide useful information for the development of single-cell manipulation techniques that are applicable to the evaluation of physical properties of the cell under physiological and pathological conditions.

20.2.3 Cantilever Modification and Application to Force Measurements

How the molecule of interest is attached to the cantilever is one of the most important factors determining the accuracy of single-molecule force measurements. A variety of methods for chemically modifying the AFM cantilever have been developed and applied to single molecule measurements of intermolecular interactions [52, 71–74]. Specific interactions between glutathione and glutathione *S*-transferase (GST), or between Ni-NTA (nickel-nitrilotriacetic acid) and (His)₆, have been utilized to trap GST-/(His)₆-tagged proteins to the AFM cantilever in the desired orientation. Methods to covalently bind glutathione or Ni-NTA to an AFM cantilever via a polyethylene glycol (PEG) spacer, and to attach GST or (His)₆-fused protein to this cantilever, have been developed [72, 74].

Glutathione contains a cysteine residue in the middle of the tripeptide (NH₂-Gln-Cys-Gly-COOH). The thiol group of the cysteine residue can react with a maleimide group on the PEG-linked AFM cantilever. The rupture force between glutathione and GST is ~ 150 pN [74]. Similarly, Ni-NTA can also be attached to an AFM cantilever via a PEG linker, with a rupture force of 150–200 pN against (His)₆ [72, 75]. The major advantage of using the glutathione and GST system (and also Ni-NTA and (His)₆ system) is that once the function of the GST (or (His)₆)-fused protein is confirmed by *in vitro* bulk assay, it can be expected to have the same activity after attachment to the cantilever because no additional modification is required for the attachment. Antibody can also be attached to an AFM cantilever via PEG-linker coupled to the carboxyl group of the antibody. By this strategy, several coupling methods have been developed [76–78]. For example, the cantilever was amino-functionalized by 3-aminopropyltriethoxysilane (APTES), then

incubated with hetero-bifunctional PEG linker with *N*-hydroxysuccinimidyl (NHS) at one end and the maleimide group at the other end. 5-Carboxyl-1-pentanethiol was then reacted with the PEG-coupled cantilever to produce carboxyl group at the end, to which antibody was coupled by EDC (1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride).

Using a glutathione-coupled cantilever, the unbinding force between importin- β and nucleoporin, a component of the nuclear pore complex (NPC), has been analyzed and provided new insights into the NPC transport system [79]. In another application, the interaction between lamin B receptor (LBR) and chromatin was examined. The rupture forces between LBR and naked DNA/histone octamer/reconstituted chromatin were measured to be 30–50 pN [78]. The measurement of adhesion forces between the heparin-binding haemagglutinin adhesion and heparin was achieved using the site-directed Ni-NTA (His)₆ system. Bimodal distribution of the heparin-binding haemagglutinin adhesion –heparin adhesion frequency and increment in adhesion force with contact time indicated that the interaction involves multiple intermolecular bridges [80].

20.2.4 Recognition Imaging: Integration of Force Measurements and Imaging

The recent development of recognition imaging using the TRECTM mode has enabled identification of a specific molecule in the AFM image [25, 81–83]. It is possible to simultaneously obtain a topographic image and the position of a specific interaction caused by attractive forces between the specimen and the protein (antibody)-coupled cantilever. With this technique, histone proteins in reconstituted chromatin have been recognized and distinguished from DNA [25, 83]. CENP-A protein in histone core particles released from native centromeric protein has also been directly identified [84]. Topographical identification of vascular endothelial-cadherin binding sites on mouse myocardium cell membranes is another application of this mode [81]. The LBR (in the inner nuclear membrane) binding sites on chromatin have also been mapped at the single molecule level. The topographical image and recognition signal show that LBR preferentially interacts with the nucleosome in chromatin (Fig. 20.4).

20.3 Eukaryotic Genome Architecture

Fully developed modern AFM technology is now highly attractive for the researchers in the field of molecular and cell biology. One of the long-standing questions to be addressed directly using AFM is how the genome organization is achieved. DNA has various local structures such as bent, supercoiled, single-stranded, and stem loop structures. These structures can be detected by biochemical assays such as chemical or enzymatic treatment followed

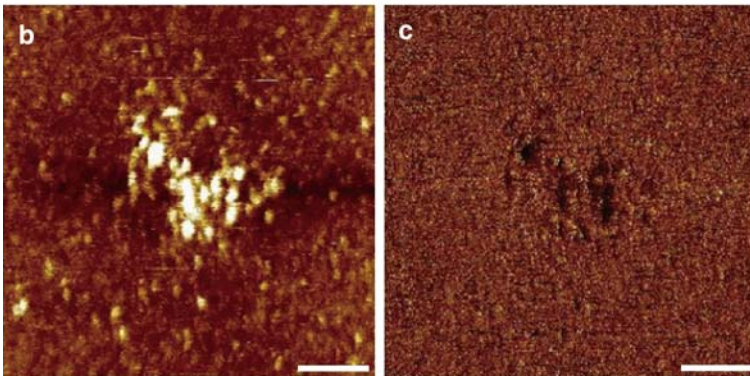
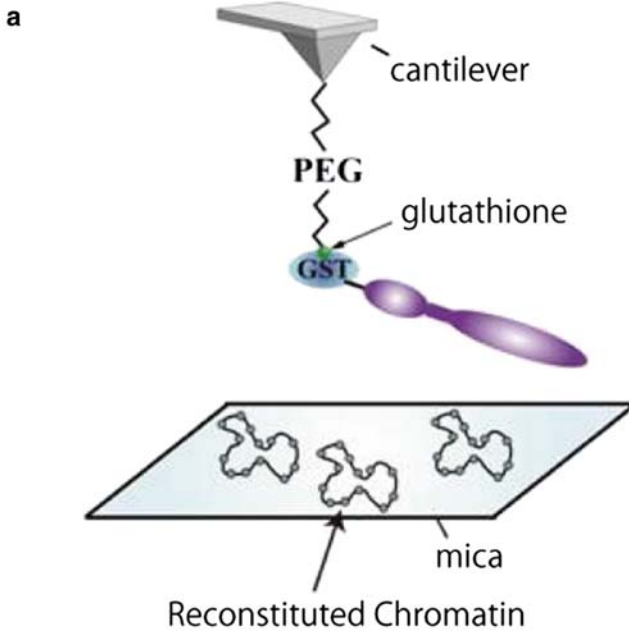


Fig. 20.4. Application of recognition imaging to identify the specific interaction of biomolecules. (a) Schematic illustration of the experimental system. The purified glutathione *S*-transferase (GST)-fused lamin B receptor (LBR) was attached to the glutathione-coupled cantilever via a polyethylene glycol (PEG) linker. (b) Beads-on-a-string chromatin architecture was observed in a topographic image. (c) Interactions between functionalized LBR-cantilever and chromatin were recorded as black signals in the recognition image. By overlapping the topographic image with the recognition image, the recognition signal can be mapped on the nucleosomal structure. Scale bar: 200 nm

by agarose gel electrophoresis. However, in terms of “structure,” biochemical analyses are somewhat limited due to their reliance on inherent indirect detection systems. On the contrary, AFM has the potential to visualize such structures directly.

20.3.1 Biophysical Properties of DNA and DNA-Binding Proteins

DNA supercoiling results from DNA winding around DNA-binding proteins and/or from the topological constraint imposed on closed circular DNA. An AFM study has demonstrated that a local strain imposed by initiator binding can induce a drastic shift of DNA conformation from a supercoiled to a relaxed state [10]. Without introduction of a DNA strand break or a local melting of the DNA double strand, the superhelical strain of a closed circular DNA can be drastically redistributed over several kilobases from writhing to twisting on protein binding, which, in turn, induces an apparent relaxation of circular DNA. A given DNA strand is rather flexible and can absorb a superhelical strain up to 1 helical turn per ~ 150 base pairs without changing its apparent shape. However, when the helical strain becomes larger, the apparent structure of the DNA strand dramatically changes to form supercoiling. This means that, when the torsional strain is close to the border size, a subtle change in the strain could induce a sudden conformational change of DNA between supercoiled and relaxed states. This physical property of DNA plays an important role in the dynamic structural transition of specific genes on replication initiation, transcriptional regulation, and recombination [10].

Conformational changes of giant DNA molecules depend on the electrical charge and topology of the DNA [85–88]. DNA conformation is affected by polycations. Polylysine [89] and spermidine [87,90] induce a large transition in DNA structure (a first-order phase transition) from the coiled state into the folded compact state (aggregates). A combination of fluorescence microscopy and AFM has revealed that different lengths of DNA exhibit different physical properties. Short DNA (<100 nm) is hard to bend, whereas medium length (>100 nm but $<4\ \mu\text{m}$) DNA is flexible, so that end-to-end joining, for example, is attained easily. Long DNA ($>4\ \mu\text{m}$) becomes less flexible than medium length in terms of end-to-end joining, due to “self-avoiding walk” [91]. The relationship between the counter length and the linear end-to-end distance between both ends of DNA (end-to-end distance) is affected by DNA topology. Comparison between circular and linear 106 kb DNA showed that the hydrodynamic radius of linear DNA, calculated from the Brownian motion of a fluorescent spot of DAPI-stained DNA molecule, is markedly smaller than that of circular DNA [85].

The earlier-mentioned physical properties of DNA are critical for the next step of DNA folding, induced by DNA-binding proteins. On binding, DNA-binding proteins induce various complex conformational changes in the DNA. AFM studies have played an important role in detecting these structures

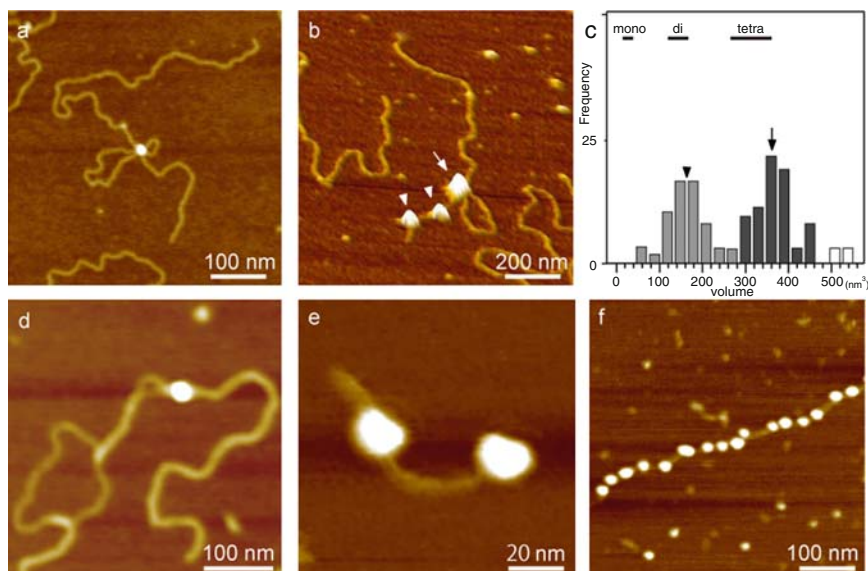


Fig. 20.5. AFM images of DNA–protein complexes. (a) DNA loop in the promoter region of the β -globin gene mediated by transcription factors. (b) DNA-loop mediated by telomere protein, TRF2. The proteins involved in the loop formation (*arrow*) and those not involved in (*arrowhead*) have different sizes. (c) The statistical analysis of the size of the protein in (b) indicates that proteins involved in the loop formation are TRF2 tetramer and those not involved in are dimer. (d) DNA-bound topoisomerase II α . (e) dinucleosome reconstituted from purified core histone and tandem repeat of positioning signals (from the 5S rRNA gene). (f) Polynucleosome reconstituted by an extract from fly embryo

(Fig. 20.5). For example, eukaryotic transcription factors (NF1 [92] and Oct-1 [93]), DNA damage recognition complex (XPC-HR23B [94]), and bacterial nucleoid proteins [HU (see Sect. 20.4)] are known to bend DNA on binding. The ability of HMGB1 (*high-mobility group box protein 1*) to bend DNA or to stimulate DNA end joining is modulated by acetylation at Lys2 [95]. Yeast centromeric DNA (*CEN* DNA)-binding factor 3 (CBF3) binds to DNA fragments containing the *CEN* sequence and bends the DNA by 55° [96]. When a transcription factor, Bach1/MafK dimer, binds to the two specific sites in the locus control region (LCR) of the β -globin gene, a DNA loop is formed via Bach1/MafK tetramerization [13]. A telomere protein, TRF2, binds to the telomeric DNA region at several places simultaneously as dimers, and often mediates DNA loop (so called T-loop) formation by forming a tetramer at the junction [97].

These AFM studies have allowed us to analyze the site of protein binding, the angle of the DNA bent at the protein-binding site, and the size and the degree of oligomerization of proteins on DNA. Namely, AFM is powerful

in solving the molecular mechanisms underlying a variety of DNA–protein interactions.

20.3.2 Fundamental Structures of Eukaryotic Genomes

Genome DNA interacts with a number of proteins and forms a hierarchical higher-order structure called “chromatin.” The higher-order structure of chromatin has been one of the main study focuses in molecular biology, because it is closely related to the mechanism of gene functions such as transcriptional regulation and DNA repair. To understand the molecular mechanisms controlling gene function, it is essential to understand the structural properties of chromatin and their stability and flexibility. Toward this goal, AFM has been combined with various biochemical reconstitution techniques, and successfully applied to the elucidation of the initial steps of genome folding.

Beads-on-a-String Structure with Core Histones

In all eukaryotic cells (except sperm), genome DNA first interacts with core histones and forms a “beads-on-a-string” structure; the bead is the fundamental component of chromatin. These beads, called “nucleosomes,” have a diameter of 11 nm [98]. Initial observations of reconstituted “beads-on-a-string” chromatin fiber were reported in 1993 [99]. By measuring the difference in contour length of reconstituted “beads-on-a-string” chromatin fibers and DNA used for the reconstitution, it was found that 146 bp DNA wraps around a core histone octamer, which is consistent with data from X-ray crystallography [98] and other AFM observations [14].

The N- and C-termini of core histones, or “histone-tails,” are believed to play a significant role in higher-order arrangement of the nucleosome array [100]. All four core histones carry the N-terminal tail regions, which are rich in lysine residues. These tails extrude from the nucleosome core [98] and undergo posttranscriptional modifications such as acetylation, phosphorylation, and methylation [101–103].

Trypsin digestion removes these histone tails [104]. The tail-less histones can form nucleosomes, but these nucleosomes exhibit behavior different from nucleosomes with normal histones, primarily due to changes in the number of charges required for nucleosome formation and nucleosome–nucleosome interaction (see below).

Linker Histones Turn Beads-on-a-String into 30-nm Chromatin Fibers

Linker histone H1 is the best-characterized protein capable of changing the structure of the nucleosomal array. Analyses of chromatin reconstitution and AFM imaging have revealed that the nucleosome–nucleosome distance becomes shorter following the addition of H1, resulting in compaction of the

nucleosomes [15, 16, 18, 22]. In the case of a long nucleosomal array reconstituted on a 100-kb plasmid [14, 105], the formation of 30-nm fibers was promoted by linker histone H1 without the need for other proteins such as histone chaperones (Fig. 20.6). The width of the fiber depends on the salt concentration: ~ 30 nm in a buffer containing 100 mM NaCl and ~ 20 nm in 50 mM NaCl [15]. This is probably because nucleosome–nucleosome interactions become weaker at lower salt concentrations [15].

Nucleosomal arrays reconstituted with tail-less histones cannot form a uniform 30-nm fiber in the presence of histone H1, indicating that the histone tails are critical for proper formation of H1-induced fibers [106]. Trypsin digestion also removes the tail of linker histones [107]. Tail-less linker histones induce compaction of normal nucleosomes, but not of nucleosomes lacking the N-terminal tails of H3 [17].

In reconstituted chromatin, the extent of H1-induced compaction of nucleosomes is greater on hypermethylated DNA than on nonmethylated DNA [16]. On the contrary, hyperacetylated nucleosomes are converted into 20-nm fibers by H1, which is thinner than the nonacetylated 30-nm fiber (Hizume et al., unpublished data). Nucleosomal arrays reconstituted with tail-less histones can be “compacted” by addition of H1, but are not converted into a uniform fiber, indicating that the histone tails are critical for proper formation of H1-induced fibers [106]. These results suggest that electrical charges on the DNA/core histone contribute greatly to the formation of 20–30-nm fibers, whereas core histone tails are likely to interact with histone H1 in the formation of H1-induced fibers. Of the many proteins that interact with chromatin (see below), only linker histone H1/H5 can promote 30-nm fiber formation.

Nonhistone Chromosomal Proteins and Chromatin Structure

Other than histones, various other proteins are associated with chromatin and regulate its structure and function. By combination of AFM and biochemical analyses, the structural changes of the chromatin induced by the chromosomal proteins have been extensively examined. For example, a transcriptional regulator, positive coactivator 4 (PC4) facilitates activator-dependent transcription by RNA polymerase II. PC4 inhibits RNA Pol-II phosphorylation, and thus activates Pol-II-mediated transcription [108]. From the analysis using circular dichroism spectroscopy, PC4 has been found to mediate chromatin condensation as histone H1 does. In contrast, AFM observation has shown that PC4 induces irregular aggregation of nucleosomes, which is quite distinct from H1-induced fiber, when mixed with reconstituted chromatin [23].

Type II topoisomerase (Topo II), which is a component of nuclear scaffold and is essential for chromosome condensation during mitosis, also binds to chromatin, according to pull-down assay using the reconstituted chromatin with H1-GST. AFM analysis has revealed that Topo II induces loop formation in 30-nm chromatin fiber reconstituted with histone H1 [24] (Fig. 20.6).

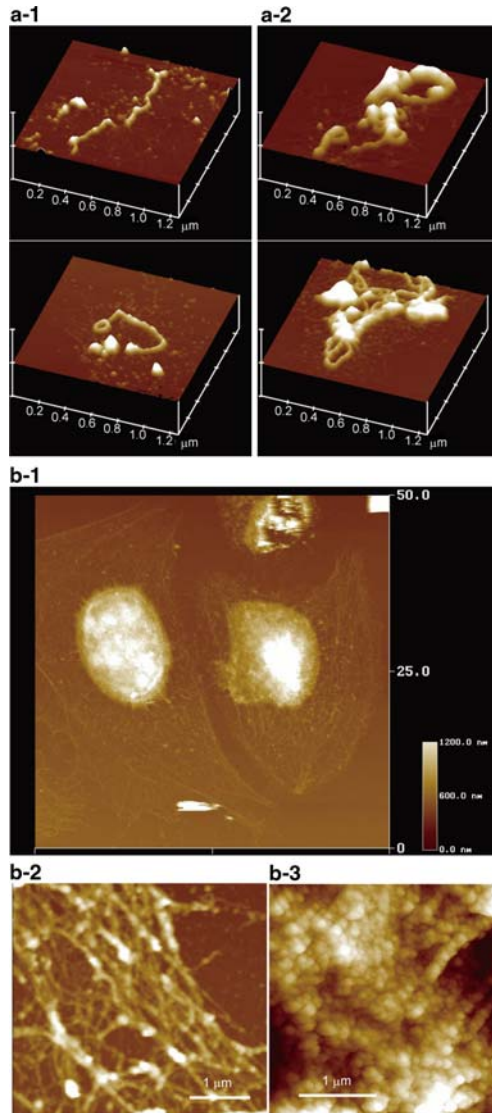


Fig. 20.6. AFM images of higher-order structures of chromatin and cytoskeleton. **(a-1)** Chromatin fibers were reconstituted from 186-kb plasmid DNA, core histones, and linker histone H1, and their structures were observed by AFM in air. H1-induced 30-nm fibers were detected. **(a-2)** Addition of topoisomerase II to the reconstituted 30-nm fibers induced loop structures. **(b-1)** AFM image of HeLa cells depleted of plasma membrane and cytoplasm. HeLa cells on a coverslip were sequentially treated with buffer A containing 0.5% Triton X-100, buffer B containing 250 mM $(\text{NH}_4)_2\text{SO}_4$, and 10 U/ml DNaseI **(b-1)**. Cytoskeletal fibers **(b-2)** and chromatin fibers **(b-3)** were clearly observed in enlarged images. The interphase chromatin is composed of granular fibers with a width of 78.1 ± 8.1 nm

Another nucleosome-binding protein, poly (ADP-ribose) polymerase 1 (PARP-1), is an NAD(+)-dependent enzyme. Glycerol-gradient fractionation of the mixture of chromatin and PARP-1 has demonstrated that PARP-1 binds to chromatin and modulates the structure [109]. AFM observation of the PARP-1/chromatin complex has indicated that PARP-1 promotes localized compaction of chromatin into supranucleosomal structures independent of the N-terminal tails of the core histones [27].

In summary, AFM analyses have revealed that the complexes of chromatin/PC4, Topo II, and PARP-1 are resulted in via distinct molecular mechanisms. Namely, PC4 forms aggregation of nucleosomes regardless of the presence of histone H1 [23]. Topo II induces chromatin compaction in a histone H1-dependent manner [24]. PARP-1 and H1 bind competitively to overlapping sites on nucleosomes [109].

20.3.3 Chromosome Structure in the Mitotic Phase

Mitotic chromosomes have been a long-standing target of structural analyses using cytological staining and optical and electron microscopy. These techniques become even more powerful when combined with AFM. AFM combined with a conventional optical microscope has been used to visualize 30-nm chromatin fibers of mitotic chromosomes isolated from Chinese hamster lung cells, and the synaptonemal complex isolated from rat spermatocytes [110]. AFM combined with *in situ* hybridization has been used to visualize specific chromosomal regions (telomere and centromere) at 200 nm resolution [111].

Giemza staining detects ~300 bands on human mitotic chromosomes. Chromatin fibers with 50–100-nm beaded structures have been identified in Giemza-positive regions, and fibers with 30–50-nm beads have been identified in Giemza-negative regions [112]. Since Giemza-positive regions contain AT-rich sequences [113], AT-rich regions are expected to be more highly compacted than GC-rich regions.

20.3.4 Chromatin Structure Inside Nuclei

Chromatin is packed into higher-order structures in nuclei, and the regulation of this packing is critical for transcriptional activity: loosely packed chromatin possesses higher transcriptional activity, and more tightly packed chromatin is silenced.

AFM analysis of chromatin structure inside the nucleus requires the plasma membrane, cytosol, and nuclear envelope to be removed to expose the chromatin. Chromatin fibers can be exposed by treatment with a buffer containing 1% Triton X-100, followed by a buffer containing 250 mM ammonium sulfate [28]. HeLa cells treated this way exhibit 80-nm beaded structures inside the nuclei [29] (Fig. 20.6). A comparison of chromatin structure inside the nuclei of HeLa cells, chicken erythrocytes, and fission yeast cells has lead to

the identification of 30–40-nm fibers and 80–100-nm beaded structures as common structural units [114]. The current model of the genome DNA hierarchy is shown in Fig. 20.7, which illustrates the stepwise formation of higher-order folding of the genome with different nuclear proteins.

Both somatic and sperm chromatin have been analyzed. Sperm nuclei contain predominantly DNA and protamine instead of histones, and the DNA region is six times more tightly packed than in normal cell nuclei. In the spermatozoa of *Sminthopsis crassicaudata*, the chromatin region containing DNA and histones has 120–180-nm nodules, but the region containing DNA and protamine has smaller 50–80-nm nodules [115]. These findings are interesting, but their biological significance remains unclear.

One advantage of AFM over electron microscopy is that AFM and fluorescence microscopy images can be merged. However, application of this combined method to reconstituted chromatin [20] and mitotic chromosomes [112, 116] has not yet yielded significant results. In the future, it is hoped that AFM analysis will be able to detect specifically labeled nanometer-scale genomic regions.

20.4 Prokaryotic Genome Architecture

Prokaryotes lack a nuclear envelope and intracellular organelles. They comprise two domains of life, bacteria and archaea that are phylogenetically distinct [117–119]. As in eukaryotic cells, genome DNA in prokaryotic cells is folded into higher-order structures. However, the protein components of prokaryotic chromosomes are very different from eukaryotic cells. Therefore, examination of the differences and the similarities between prokaryotic and eukaryotic genome architectures will provide clues for elucidation of the principles of genome-folding mechanism. AFM has been useful toward this goal.

20.4.1 Bacterial DNA-Binding Proteins

Most bacterial genomes are circular and are packed into cells as structures a few micrometers in diameter in the form of a “nucleoid” [120]. Studies on proteins involved in the genome architecture have revealed that bacteria lack histones [121]. *Escherichia coli* is one of the best-studied bacteria, and significant effort has been made to understand its nucleoid structure. *E. coli* has a set of DNA-binding proteins including RNA polymerases and about 300 types of transcription factors [121]. The amount of each nucleoid protein in a cell changes depending on growth conditions and influences the genome structure [122, 123]. Among them, HU (heat-unstable nucleoid protein) is one of the most abundant protein in *E. coli* [122]; IHF (integration host factor), H-NS (histone-like nucleoid structural protein), Fis (factor for inversion stimulation), Hfq (host factor for phage RNA Q β replication), StpA (suppressor

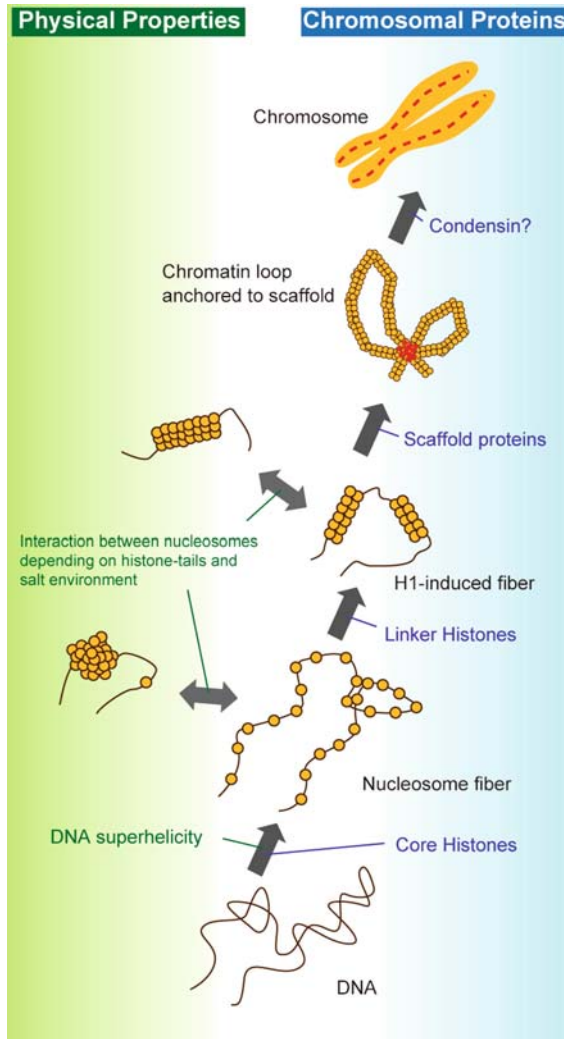


Fig. 20.7. Schematic presentation of the hierarchical architecture of eukaryotic chromatin/chromosomes. Physical properties of DNA, chromosomal proteins, and local salt concentrations are critical for the formation of higher-order structures. First, DNA interacts with core histones, forming a beads-on-a-string structure, in which negative supercoiling of DNA promotes formation of nucleosomes. Second, linker histone H1 converts the beads-on-a-string into a 20–30-nm fiber. Salt concentrations affect the interaction between nucleosomes. By increasing the concentration of NaCl from 50 to 100 mM, well-separated beads-on-a-string convert to aggregates, and H1-induced fibers thicken from 20 to 30 nm in diameter. As shown in Fig. 20.6a, topoisomerase II, a component of the nuclear scaffold, makes loop structures in H1-induced fibers. Finally, mitotic chromosomes are formed with help from condensin complexes and other cellular components

of T4 td mutant phenotype A), and Dps (DNA-binding protein from starved cells) are the major small nucleoid proteins [124]. These basic proteins do not have homology with eukaryotic histones, but AFM studies have shown that they bind to DNA in a distinctive manner to form different higher-order structures [125].

HU and Dps proteins are widely spread throughout the bacterial domain, whereas the other nucleoid proteins are specific to γ -proteobacteria including *E. coli* [126]. An AFM study [127] has revealed that at low concentration (<100 nM), individual HU dimers induce very flexible bends in DNA, and at higher concentrations, a rigid nucleoprotein filament is formed in which HU appears to arrange helically around the DNA. The bending angle of HU-DNA complex exhibits broad distribution rather than preferential angle, indicating that HU-DNA complex is a flexible structure.

AFM study has identified another type of proteins that form dynamic cross-links between DNA duplexes, leading to the formation of loops [128]. In this manner, many members of this class of proteins are capable of imposing an additional level of higher-order organization on the bacterial chromosome [129]. For example, H-NS can form topological barriers because of this DNA-bridging function [130, 131]. It is known that the distribution of nucleoid proteins over the entire genome is uneven and restricted to certain chromosomal loci [132–134]. Such uneven localization may be related to structural boundaries on the nucleoid [135]. H-NS binds to DNA of a specific sequence at low concentrations, but without sequence specificity at high concentrations [124, 136]. This mode of DNA binding (sequence specificity at low concentration and less specificity at high protein concentration) seems to be common for bacterial nucleoid proteins.

Bacteria also possess the structural maintenance of chromosomes (SMC) family of proteins such as MukB in *E. coli* [137] and BsSMC in *Bacillus subtilis* [138]. AFM imaging has shown the existence of higher-order rosette-like structures of BsSMC *in vitro*, which might explain the formation of the SMC centers *in vivo* [138]. These SMC proteins have little sequence similarity to the eukaryotic SMCs, but share common structural and functional characteristics (Fig. 20.3) such as DNA-binding domains and ATP-binding domains connected by a long coiled-coil and a flexible center hinge [138, 139]. The enzymatic and DNA-binding mechanisms of these proteins are similar to the eukaryotic condensins, and their ATPase activities appear to play a key role in DNA-binding and nucleoid condensation [140, 141].

20.4.2 Bacterial Genome Structure and Dynamics

Bacterial nucleoids (chromosome) also have stepwise higher-order architectures [123, 142]. A modification of the technique of Murphy and Zimmerman can be applied to the study of nucleoid structural hierarchy by AFM [143]. *E. coli* cells are harvested in the log phase and treated with lysozyme and detergent to disrupt the cell wall and cell membrane. When lysed using these

conditions, the thinnest fibers released are 30 nm in diameter [123] (Fig. 20.8). Loops composed of 80-nm beaded structures are also detectable [123]. These fibers are further condensed in stationary phase cells, or cells exposed to environmental stress [123,144]. Interestingly, 30-nm fibers and 80-nm beaded structures are commonly found in nucleoids of other bacteria, such as *Staphylococcus aureus* [145] and *Clostridium perfringens* [126], which have somewhat different nucleoid proteins from *E. coli*. This implies the existence of stepwise DNA-folding mechanisms for achieving higher-order architectures, regardless of the protein components involved. A physicochemical mechanism is likely responsible for this hierarchy. In addition, these higher-order structures are generally sensitive to RNase A. Enzymatic digestion with RNase A, but not with RNase H, produced 10-nm fibers, suggesting that the thinnest fibers in the absence of RNA are 10 nm in diameter [142].

The nucleoid changes its structure dynamically. Examples are Dps-dependent nucleoid condensation in stationary phase or oxidatively stressed cells. Dps was first identified as a starvation-induced protein with nonspecific DNA-binding activity [146]. Dps helps the protection of DNA against irradiation, oxidation, and nuclease attack [147]. When Dps binds to DNA, and it produces large aggregates *in vitro* (Fig. 20.8) [148]. The amount of Dps in cells dramatically increases in cells approaching stationary phase [122] or on oxidative stress [142]. AFM analysis has shown that the *E. coli* nucleoid undergoes tight condensation near stationary phase, and that Dps is required for this condensation [123]. Oxidative stress also induces Dps-dependent compaction of the *E. coli* and *S. aureus* nucleoid [145].

20.4.3 Archaeal DNA-Binding Proteins, Genome Structure, and Dynamics

Although archaea constitute a domain of life distinct from eukarya and bacteria, the components of archaeal DNA replication [149] and transcription [150,151] apparatuses are more closely related to those of eukaryotes than to bacteria. There are two major phyla in archaea: *Euryarchaeota* and *Crenarchaeota*. Although most euryarchaeal species have histone genes [152], some euryarchaeota such as *Thermoplasma* lack histone genes and instead have genes homologous to HU [153]. *Crenarchaeota* species lack both histone and HU genes, but all have Alba protein (acetylation lowers binding affinity). The mechanisms controlling the genome structures of these species, which have different DNA-binding proteins, remain unknown and expected to be critical issues in understanding the genome architectures. AFM should be a powerful tool for providing insights into this matter.

Archaeal histones consist solely of a conserved histone-fold motif and form tetramer [154,155], whereas eukaryotic core histones have additional sequences that extend N- and/or C-terminal from the histone folds and form octamer [98]. These differences causes an apparent difference in properties of archaeal and eukaryotic nucleosomes; namely, eukaryotic histones wrap 146 bp of DNA

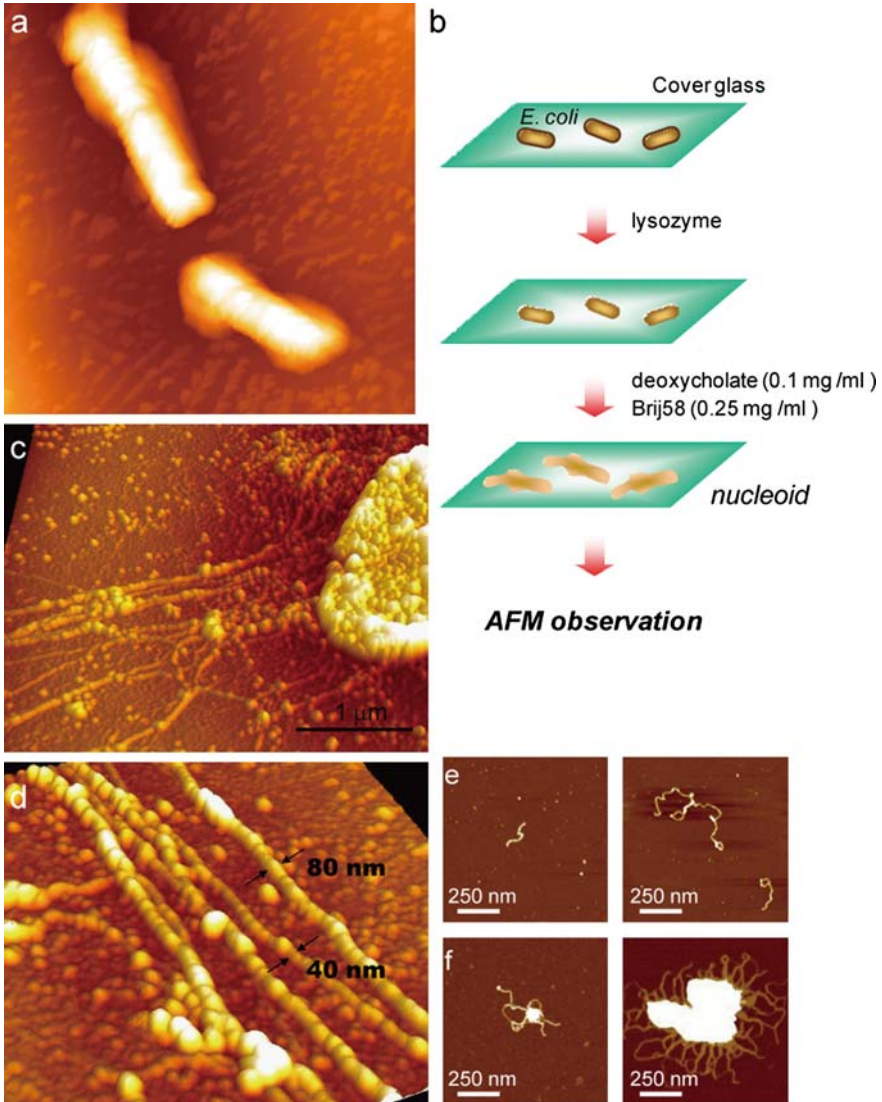


Fig. 20.8. Fundamental structures of *Escherichia coli* nucleoid. (a) AFM observation of intact *E. coli*. (b) Procedure of on-substrate-lysis of *E. coli* cell. Log phase cells were attached to a cover glass and then mildly treated with lysozyme and detergent. (c, d) AFM images of the lysed *E. coli* cell. (e, f) In vitro actions of nucleoid proteins. 14-kb plasmid was mixed with HU (e) and Dps (f) at a ratio of 1 protein per 1 bp and observed by AFM. 10-nm fibers (e) and aggregates (f) were observed

[98, 156], whereas archaeal histones wrap about 120 bp DNA per nucleosome [157]. In this situation, the surface interaction between DNA and histone fold in archaea is different from that in eukaryote, and accordingly reflected in the Micrococcal nuclease digestion assay *in vitro* [158, 159], where ~60 and ~90 bp DNA can be protected in archaeal nucleosome, whereas ~150 bp DNA is protected in the eukaryotic nucleosome.

Thermococcus kodakaraensis is a hyperthermophilic archaeon in the *Euryarchaeota* lineage and has histone genes. It is an ideal candidate for the study of chromosome architecture because the cell is easily disrupted and chromosomes can be exposed by a hypotonic treatment. This leads to disruption of the cell membrane and release of the chromosome fibers from the cell. AFM revealed both beads-on-a-string and thicker fibrous structures in the *T. kodakaraensis* chromosome (Fig. 20.9). When chromosome of *T. kodakaraensis* is purified, partially digested with Micrococcal nuclease and separated on sucrose density gradient sedimentation, chromosome fragments with different protein compositions can be separated. AFM analysis shows that these chromosome fragments have distinct structures with different proteins (Fig. 20.9), suggesting that the archaeal chromosomes are not homogeneous in terms of protein composition and structure. It is known that bacterial chromosome is not a uniform structure but rather form "macrodomains," which contain different proteins [160]. This may also be the case for archaeal chromosome. Thus, AFM analysis after separation of different kinds of chromosome fragments by a sucrose density gradient sedimentation can be a general method to identify different structures on chromosome.

Condensation of some archaeal chromosomes (e.g., *T. kodakaraensis*) occurs in the stationary phase. By staining DNA with DAPI, or by AFM observation, the extent of chromosome condensation can be analyzed. Compared to cells in exponential phase, chromosomes of stationary phase cells are less spread out. This suggests that chromosome condensation in stationary phase is common between bacteria and archaea. Genes coding for SMC proteins also exist in archaeal genomes. It is possible that SMC proteins are involved in prokaryotic chromosome condensation.

As with bacterial nucleoids, RNase A treatment of *T. kodakaraensis* chromosomes results in exposure of 10-nm fibers. This is consistent with the result that RNase A treatment changes the density of bacterial nucleoids [161] or eukaryotic chromosomes [162]. RNase A also disrupts eukaryotic heterochromatin structure [163]. Thus, involvement of RNA in higher-order folding of the genome may also be a common mechanism among the three domains of life. AFM observation of RNase A-treated chromosomes may be a general method for analyzing the fundamental structures of chromosomes.

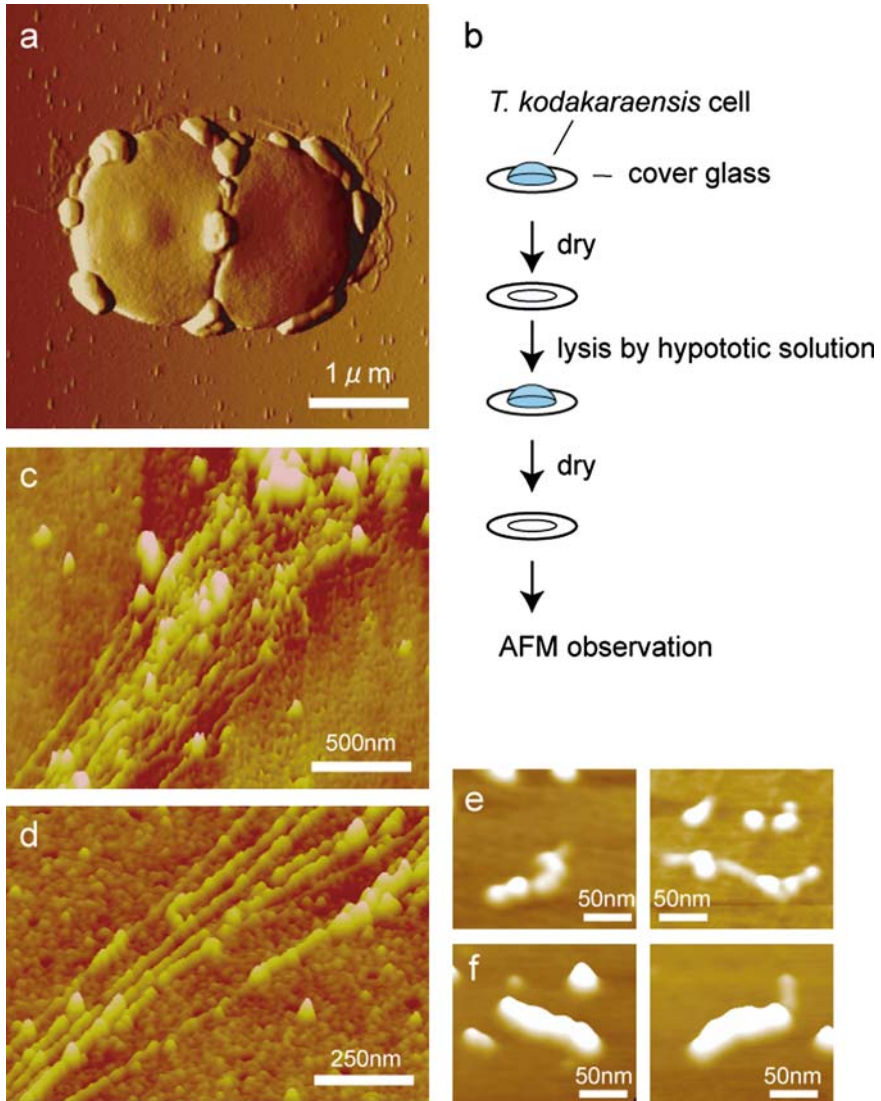


Fig. 20.9. Fundamental structures of *Thermococcus kodakaraensis* chromosome. (a) AFM image of intact *T. kodakaraensis* cell. (b) Procedure of on-substrate-lysis of *T. kodakaraensis*. Cells were attached onto a cover glass. After removing liquid medium, MilliQ water was added to lyse the cells. (c, d) AFM images of lysed *T. kodakaraensis* cells. (e, f) Structures of *T. kodakaraensis* chromosome fragments separated by sucrose density gradient sedimentation. (e) Beads-on-a-string-like structure was identified in the fractions containing histone proteins. (f) Sausage-like structure was detected in the fractions containing nonhistone proteins

20.5 Conclusion/Perspectives

In the last three decades, the biological applications of AFM have expanded widely due to technical developments and advances in instrumentation. These advances include high-resolution imaging in solution, fast-scanning AFM technology, and cantilever-modification methods. Researchers now have the necessary tools and techniques for structure/function analyses of biological macromolecules in solution at the nanometer scale, subsecond time resolution, and pico-Newton force level. These AFM technologies have uncovered many aspects of genome folding mechanisms including (1) physical properties of DNA and their importance in the first steps of genome DNA folding and (2) the existence of hierarchy in the higher-order architecture of genome DNA.

In eukaryotes, AFM has been powerful in elucidating (1) the mechanisms and dynamics of nucleosomes on long DNA, (2) the role of linker histones in the second step toward higher-order architecture (such as the formation of 30-nm chromatin fibers), (3) the significance of histone tails in the formation of 30-nm chromatin fibers, and (4) the mechanisms of chromatin-loop formation. It should be noted that, among many proteins, only linker histone H1/H5 can promote 30-nm fiber formation. Sperm chromatin is distinct from that of other cells in that the major chromatin protein is protamine instead of histone. The biological significance of this important distinction is unclear, and must be clarified in the future.

The protein components of prokaryotic chromosomes are very different from eukaryotic cells. For example, all bacteria lack histones and the most possess HU, but some archaea possess histone genes with a significant diversity between eukaryotic histones. Nevertheless, it is intriguing that genome DNA of all three domains (eukarya, bacteria, and archaea) is folded into very similar higher-order structures stepwise from ~ 10 to 30-nm fibers, and then into much thicker structures via loop formation. Involvement of RNA in higher-order folding of genomes may also be a common mechanism among the three domains of life. Elucidation of the physicochemical principles that govern the hierarchical architecture is urgently required and is expected to be key for revealing the structure and function of the genome in general.

References

1. P.K. Hansma, V.B. Elings, O. Marti, C.E. Bracker, Scanning tunneling microscopy and atomic force microscopy: Application to biology and technology. *Science* **242**, 209–216 (1988).
2. D.P. Allison, L.A. Bottomley, T. Thundat, G.M. Brown, R.P. Woychik, J.J. Schrick, K.B. Jacobson, R.J. Warmack, Immobilization of DNA for scanning probe microscopy. *Proc. Natl. Acad. Sci. USA* **89**, 10129–10133 (1992).
3. R.D. Edstrom, X.R. Yang, G. Lee, D.F. Evans, Viewing molecules with scanning tunneling microscopy and atomic force microscopy. *FASEB J.* **4**, 3144–3151 (1990).

4. F. Zenhausern, M. Adrian, B. ten Heggeler-Bordier, R. Emch, M. Jobin, M. Taborelli, P. Descouts, Imaging of DNA by scanning force microscopy. *J. Struct. Biol.* **108**, 69–73 (1992).
5. S. Kasas, N.H. Thomson, B.L. Smith, H.G. Hansma, X. Zhu, M. Guthold, C. Bustamante, E.T. Kool, M. Kashlev, P.K. Hansma, Escherichia coli RNA polymerase activity observed using atomic force microscopy. *Biochemistry* **36**, 461–468 (1997).
6. S. Nettikadan, F. Tokumasu, K. Takeyasu, Quantitative analysis of the transcription factor AP2 binding to DNA by atomic force microscopy. *Biochem. Biophys. Res. Co.* **226**, 645–649 (1996).
7. Y.L. Lyubchenko, DNA structure and dynamics: An atomic force microscopy study. *Cell. Biochem. Biophys.* **41**, 75–98 (2004).
8. A.L. Mikheikin, A.Y. Lushnikov, Y.L. Lyubchenko, Effect of DNA supercoiling on the geometry of holliday junctions. *Biochemistry* **45**, 12998–13006 (2006).
9. V. Viglasky, F. Valle, J. Adamcik, I. Joab, D. Podhradsky, G. Dietler, Anthracycline-dependent heat-induced transition from positive to negative supercoiled DNA. *Electrophoresis* **24**, 1703–1711 (2003).
10. S.H. Yoshimura, R.L. Ohniwa, M.H. Sato, F. Matsunaga, G. Kobayashi, H. Uga, C. Wada, K. Takeyasu, DNA phase transition promoted by replication initiator. *Biochemistry* **39**, 9139–9145 (2000a).
11. T. Ohta, S. Nettikadan, F. Tokumasu, H. Ideno, Y. Abe, M. Kuroda, H. Hayashi, K. Takeyasu, Atomic force microscopy proposes a novel model for stem-loop structure that binds a heat shock protein in the Staphylococcus aureus HSP70 operon. *Biochem. Biophys. Res. Commun.* **226**, 730–734 (1996).
12. K. Virnik, Y.L. Lyubchenko, M.A. Karymov, P. Dahlgren, M.Y. Tolstorukov, S. Semsey, V.B. Zhurkin, S. Adhya, “Antiparallel” DNA loop in gal repressosome visualized by atomic force microscopy. *J. Mol. Biol.* **334**, 53–63 (2003).
13. S.H. Yoshimura, C. Yoshida, K. Igarashi, K. Takeyasu, Atomic force microscopy proposes a ‘kiss and pull’ mechanism for enhancer function. *off. J. Electron. Microsc. (Tokyo)* **49**, 407–413 (2000b).
14. K. Hizume, S.H. Yoshimura, K. Takeyasu, Atomic force microscopy demonstrates a critical role of DNA superhelicity in nucleosome dynamics. *Cell Biochem. Biophys.* **40**, 249–262 (2004).
15. K. Hizume, S.H. Yoshimura, K. Takeyasu, Linker Histone H1 per se Can Induce Three-Dimensional Folding of Chromatin Fiber. *Biochemistry* **44**, 12978–12989 (2005).
16. M.A. Karymov, M. Tomschik, S.H. Leuba, P. Caiafa, J. Zlatanova, DNA methylation-dependent chromatin fiber compaction in vivo and in vitro: requirement for linker histone. *Faseb J.* **15**, 2631–2641 (2001).
17. S.H. Leuba, C. Bustamante, K. van Holde, J. Zlatanova, Linker histone tails and N-tails of histone H3 are redundant: scanning force microscopy studies of reconstituted fibers. *Biophys. J.* **74**, 2830–2839 (1998a).
18. S.H. Leuba, C. Bustamante, J. Zlatanova, K. van Holde, Contributions of linker histones and histone H3 to chromatin structure: scanning force microscopy studies on trypsinized fibers. *Biophys. J.* **74**, 2823–2829 (1998b).
19. S.H. Leuba, G. Yang, C. Robert, B. Samori, K. van Holde, J. Zlatanova, C. Bustamante, Three-dimensional structure of extended chromatin fibers as revealed

- by tapping-mode scanning force microscopy. *Proc. Natl. Acad. Sci. USA* **91**, 11621–11625 (1994).
20. T. Nakai, K. Hizume, S.H. Yoshimura, K. Takeyasu, K. Yoshikawa, Phase transition in reconstituted chromatin. *Europhysics Letters* **69**, 1024–1030 (2005).
 21. T. Sakaue, K. Yoshikawa, S.H. Yoshimura, K. Takeyasu, Histone core slips along DNA and prefers positioning at the chain end. *Phys. Rev. Lett.* **87**, 078105 (2001).
 22. M.H. Sato, K. Ura, K.I. Hohmura, F. Tokumasu, S.H. Yoshimura, F. Hanaoka, K. Takeyasu, Atomic force microscopy sees nucleosome positioning and histone H1-induced compaction in reconstituted chromatin. *FEBS Lett.* **452**, 267–271 (1999).
 23. C. Das, K. Hizume, K. Batta, B.R. Kumar, S.S. Gadad, S. Ganguly, S. Lorain, A. Verreault, P.P. Sadhale, K. Takeyasu, et al. Transcriptional coactivator PC4, a chromatin-associated protein, induces chromatin condensation. *Mol. Cell. Biol.* **26**, 8303–8315 (2006).
 24. K. Hizume, S. Araki, K. Yoshikawa, K. Takeyasu, Topoisomerase II, scaffold component, promotes chromatin compaction in vitro in a linker-histone H1-dependent manner. *Nucleic Acids Res.* **35**, 2787–2799 (2007).
 25. R. Bash, H. Wang, C. Anderson, J. Yodh, G. Hager, S.M. Lindsay, D. Lohr, AFM imaging of protein movements: Histone H2A-H2B release during nucleosome remodeling. *FEBS Lett.* **580**, 4757–4761 (2006).
 26. G.R. Schnitzler, C.L. Cheung, J.H. Hafner, A.J. Saurin, R.E. Kingston, C.M. Lieber, Direct imaging of human SWI/SNF-remodeled mono- and polynucleosomes by atomic force microscopy employing carbon nanotube tips. *Mol. Cell Biol.* **21**, 8504–8511 (2001).
 27. D.A. Wacker, D.D. Ruhl, E.H. Balagamwala, K.M. Hope, T. Zhang, W.L. Kraus, The DNA binding and catalytic domains of poly(ADP-ribose) polymerase 1 cooperate in the regulation of chromatin structure and transcription. *Mol. Cell. Biol.* **27**, 7475–7485 (2007).
 28. E.G. Fey, G. Krochmalnic, S. Penman, The nonchromatin substructures of the nucleus: the ribonucleoprotein (RNP)-containing and RNP-depleted matrices analyzed by sequential fractionation and resinless section electron microscopy. *J. Cell. Biol.* **102**, 1654–1665 (1986).
 29. S.H. Yoshimura, J. Kim, K. Takeyasu, On-substrate lysis treatment combined with scanning probe microscopy revealed chromosome structures in eukaryotes and prokaryotes. *J. Electron. Microsc. (Tokyo)* **52**, 415–423 (2003).
 30. C. Perez-Terzic, J. Pyle, M. Jaconi, L. Stehno-Bittel, D.E. Clapham, Conformational states of the nuclear pore complex induced by depletion of nuclear Ca²⁺ stores. *Science* **273**, 1875–1877 (1996).
 31. D. Stoffler, K.N. Goldie, B. Feja, U. Aebi, Calcium-mediated structural changes of native nuclear pore complexes monitored by time-lapse atomic force microscopy. *J. Mol. Biol.* **287**, 741–752 (1999).
 32. Z.M. Wang, M.L. Messi, O. Delbono, Patch-clamp recording of charge movement, Ca(2+) current, and Ca(2+) transients in adult skeletal muscle fibers. *Biophys. J.* **77**, 2709–2716 (1999).
 33. H. Oberleithner, H. Schillers, M. Wilhelmi, D. Butzke, T. Danker, Nuclear pores collapse in response to CO₂ imaged with atomic force microscopy. *Pflugers Arch.* **439**, 251–255 (2000).

34. A. Rakowska, T. Danker, S.W. Schneider, H. Oberleithner, ATP-Induced shape change of nuclear pores visualized with the atomic force microscope. *J. Membr. Biol.* **163**, 129–136 (1998).
35. L. Kastrup, H. Oberleithner, Y. Ludwig, C. Schafer, V. Shahin, Nuclear envelope barrier leak induced by dexamethasone. *J. Cell. Physiol.* **206**, 428–434 (2006).
36. V. Shahin, Y. Ludwig, C. Schafer, D. Nikova, H. Oberleithner, Glucocorticoids remodel nuclear envelope structure and permeability. *J. Cell. Sci.* **118**, 2881–2889 (2005).
37. D.J. Muller, J.B. Heymann, F. Oesterhelt, C. Moller, H. Gaub, G. Buldt, A. Engel, Atomic force microscopy of native purple membrane. *Biochim. Biophys. Acta.* **1460**, 27–38 (2000).
38. M. Kessler, H.E. Gaub, Unfolding barriers in bacteriorhodopsin probed from the cytoplasmic and the extracellular side by AFM. *Structure* **14**, 521–527 (2006).
39. D. Pogoryelov, J. Yu, T. Meier, J. Vonck, P. Dimroth, D.J. Muller, The c15 ring of the *Spirulina platensis* F-ATP synthase: F1/F0 symmetry mismatch is not obligatory. *EMBO Rep.* **6**, 1040–1044 (2005).
40. H. Stahlberg, D.J. Muller, K. Suda, D. Fotiadis, A. Engel, T. Meier, U. Matthey, P. Dimroth, Bacterial Na(+)-ATP synthase has an undecameric rotor. *EMBO Rep.* **2**, 229–233 (2001).
41. D. Fotiadis, Y. Liang, S. Filipek, D.A. Saperstein, A. Engel, K. Palczewski, Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **421**, 127–128 (2003).
42. D. Fotiadis, D.J. Muller, G. Tsiotis, L. Hasler, P. Tittmann, T. Mini, P. Jenö, H. Gross, A. Engel, Surface analysis of the photosystem I complex by electron and atomic force microscopy. *J. Mol. Biol.* **283**, 83–94 (1998).
43. D.J. Muller, W. Baumeister, A. Engel, Controlled unzipping of a bacterial surface layer with atomic force microscopy. *Proc. Natl. Acad. Sci. USA* **96**, 13170–13174 (1999).
44. S.H. Yoshimura, K. Hizume, A. Murakami, T. Sutani, K. Takeyasu, M. Yanagida, Condensin Architecture and Interaction with DNA. Regulatory Non-SMC Subunits Bind to the Head of SMC Heterodimer. *Curr. Biol.* **12**, 508–513 (2002).
45. N.P. Barrera, S.J. Ormond, R.M. Henderson, R.D. Murrell-Lagnado, J.M. Edwardson, Atomic force microscopy imaging demonstrates that P2X2 receptors are trimers but that P2X6 receptor subunits do not oligomerize. *J. Biol. Chem.* **280**, 10759–10765 (2005).
46. K.I. Hohmura, Y. Itokazu, S.H. Yoshimura, G. Mizuguchi, Y.S. Masamura, K. Takeyasu, Y. Shiomi, T. Tsurimoto, H. Nishijima, S. Akita, et al. Atomic force microscopy with carbon nanotube probe resolves the subunit organization of protein complexes. *J. Electron Microsc. (Tokyo)* **49**, 415–421 (2000).
47. H. Nishijima, S. Kamo, S. Akita, Y. Nakayama, K.I. Hohmura, S.H. Yoshimura, K. Takeyasu, Carbon-nanotube tips for scanning probe microscopy: Preparation by a controlled process and observation of deoxyribonucleic acid. *Applied Physics Letters* **74**, 4061–4063 (1999).
48. T. Fukuma, K. Kobayashi, K. Matsushige, H. Yamada, True atomic resolution in liquid by frequency-modulation atomic force microscopy. *Applied Physics Letters* **87**, (2005).

49. T. Ando, N. Kodera, E. Takai, D. Maruyama, K. Saito, A. Toda, A high-speed atomic force microscope for studying biological macromolecules. *Proc. Natl. Acad. Sci. USA* **98**, 12468–12472 (2001).
50. M. Yokokawa, C. Wada, T. Ando, N. Sakai, A. Yagi, S.H. Yoshimura, K. Takeyasu, Fast-scanning atomic force microscopy reveals the ATP/ADP-dependent conformational changes of GroEL. *EMBO J.* **25**, 4567–4576 (2006).
51. N. Crampton, M. Yokokawa, D.T. Dryden, J.M. Edwardson, D.N. Rao, K. Takeyasu, S.H. Yoshimura, R.M. Henderson, Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP15I is capable of DNA translocation and looping. *Proc. Natl. Acad. Sci. USA* **104**, 12755–12760 (2007).
52. A. Ikai, R. Afrin, Toward mechanical manipulations of cell membranes and membrane proteins using an atomic force microscope: an invited review. *Cell Biochem. Biophys.* **39**, 257–277 (2003).
53. D. Fotiadis, S. Scheuring, S.A. Muller, A. Engel, D.J. Muller, Imaging and manipulation of biological structures with the AFM. *Micron.* **33**, 385–397 (2002).
54. M. Guthold, X. Zhu, C. Rivetti, G. Yang, N.H. Thomson, S. Kasas, H.G. Hansma, B. Smith, P.K. Hansma, C. Bustamante, Direct observation of one-dimensional diffusion and transcription by *Escherichia coli* RNA polymerase. *Biophys. J.* **77**, 2284–2294 (1999).
55. M.B. Viani, L.I. Pietrasanta, J.B. Thompson, A. Chand, I.C. Gebeshuber, J.H. Kindt, M. Richter, H.G. Hansma, P.K. Hansma, Probing protein-protein interactions in real time. *Nat. Struct. Biol.* **7**, 644–647 (2000).
56. M. Kobayashi, K. Sumitomo, K. Torimitsu, Real-time imaging of DNA-streptavidin complex formation in solution using a high-speed atomic force microscope. *Ultramicroscopy* **107**, 184–190 (2007).
57. K.S. Matthews, DNA looping. *Microbiol. Rev.* **56**, 123–136 (1992).
58. P. Janscak, U. Sandmeier, M.D. Szczelkun, T.A. Bickle, Subunit assembly and mode of DNA cleavage of the type III restriction endonucleases EcoP1I and EcoP15I. *J. Mol. Biol.* **306**, 417–431 (2001).
59. A. Meisel, P. Mackeldanz, T.A. Bickle, D.H. Kruger, C. Schroeder, Type III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATP hydrolysis. *EMBO J.* **14**, 2958–2966 (1995).
60. L.J. Peakman, M.D. Szczelkun, DNA communications by Type III restriction endonucleases—confirmation of 1D translocation over 3D looping. *Nucleic Acids Res.* **32**, 4166–4174 (2004).
61. R. Afrin, T. Yamada, A. Ikai, Analysis of force curves obtained on the live cell membrane using chemically modified AFM probes. *Ultramicroscopy* **100**, 187–195 (2004).
62. H. Haga, S. Sasaki, K. Kawabata, E. Ito, T. Ushiki, T. Sambongi, Elasticity mapping of living fibroblasts by AFM and immunofluorescence observation of the cytoskeleton. *Ultramicroscopy* **82**, 253–258 (2000).
63. A.E. Pelling, F.S. Veraitch, C. Pui-Kei Chu, B.M. Nicholls, A.L. Hemsley, C. Mason, M.A. Horton, Mapping correlated membrane pulsations and fluctuations in human cells. *J. Mol. Recognit.* **20**, 467–475 (2007).
64. Y. Yamane, H. Shiga, H. Haga, K. Kawabata, K. Abe, E. Ito, Quantitative analyses of topography and elasticity of living and fixed astrocytes. *J. Electron Microsc. (Tokyo)* **49**, 463–471 (2000).

65. M. Yokokawa, K. Takeyasu, S.H. Yoshimura, Mechanical properties of plasma membrane and nuclear envelope measured by scanning probe microscope. *J. Microsc.* **232**, 82–90 (2008).
66. I. Sneddon, The relation between load and penetration in the axisymmetric Boussinesq problem for a punch of arbitrary profile. *J. Struct. Biol.* **119**, 84–91 (1965).
67. U.G. Hofmann, C. Rotsch, W.J. Parak, M. Radmacher, Investigating the cytoskeleton of chicken cardiocytes with the atomic force microscope. *J. Struct. Biol.* **119**, 84–91 (1997).
68. M. Nagayama, H. Haga, K. Kawabata, Drastic change of local stiffness distribution correlating to cell migration in living fibroblasts. *Cell Motil Cytoskeleton* **50**, 173–179 (2001).
69. C. Rotsch, M. Radmacher, Drug-induced changes of cytoskeletal structure and mechanics in fibroblasts: an atomic force microscopy study. *Biophys. J.* **78**, 520–535 (2000).
70. M. Nagayama, H. Haga, M. Takahashi, T. Saitoh, K. Kawabata, Contribution of cellular contractility to spatial and temporal variations in cellular stiffness. *Exp. Cell Res.* **300**, 396–405 (2004).
71. D.P. Allison, P. Hinterdorfer, W. Han, Biomolecular force measurements and the atomic force microscope. *Curr. Opin. Biotechnol.* **13**, 47–51 (2002).
72. R. Gamsjaeger, B. Wimmer, H. Kahr, A. Tinazli, S. Picuric, S. Lata, R. Tampe, Y. Maulet, H.J. Gruber, P. Hinterdorfer, et al. Oriented binding of the His6-tagged carboxyl-tail of the L-type Ca²⁺ channel alpha1-subunit to a new NTA-functionalized self-assembled monolayer. *Langmuir* **20**, 5885–5890 (2004).
73. L. Schmitt, M. Ludwig, H.E. Gaub, R. Tampe, A metal-chelating microscopy tip as a new toolbox for single-molecule experiments by atomic force microscopy. *Biophys. J.* **78**, 3275–3285 (2000).
74. S.H. Yoshimura, H. Takahashi, S. Otsuka, K. Takeyasu, Development of glutathione-coupled cantilever for the single-molecule force measurement by scanning force microscopy. *FEBS Lett.* **580**, 3961–3965 (2006).
75. F. Kienberger, G. Kada, H. Gruber, V. Pastushenko, C. Riener, M. Trieb, H. Knaus, H. Schindler, P. Hinterdorfer, Recognition force spectroscopy studies of the NTA-His6 Bond. *Single Mol.* **1**, 59–65 (2000).
76. A. Ebner, L. Wildling, A.S. Kamruzzahan, C. Rankl, J. Wruss, C.D. Hahn, M. Holzl, R. Zhu, F. Kienberger, D. Blaas, et al. A new, simple method for linking of antibodies to atomic force microscopy tips. *Bioconjug. Chem.* **18**, 1176–1184 (2007).
77. P. Hinterdorfer, W. Baumgartner, H.J. Gruber, K. Schilcher, H. Schindler, Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proc. Natl. Acad. Sci. USA* **93**, 3477–3481 (1996).
78. Y. Hirano, H. Takahashi, M. Kumeta, K. Hizume, Y. Hirai, S. Otsuka, S.H. Yoshimura, K. Takeyasu, Nuclear architecture and chromatin dynamics revealed by atomic force microscopy in combination with biochemistry and cell biology. *Pflugers Arch.* **456**, 139–153 (2008).
79. S. Otsuka, S. Iwasaka, Y. Yoneda, K. Takeyasu, S.H. Yoshimura, Individual binding pockets of importin-beta for FG-nucleoporins have different binding properties and different sensitivities to RanGTP. *Proc. Natl. Acad. Sci. USA* **105**, 16101–16106 (2008).
80. V. Dupres, F.D. Menozzi, C. Locht, B.H. Clare, N.L. Abbott, S. Cuenot, C. Bompard, D. Raze, Y.F. Dufrene, Nanoscale mapping and functional

- analysis of individual adhesins on living bacteria. *Nat. Methods* **2**, 515–520 (2005).
81. L.A. Chtcheglova, J. Waschke, L. Wildling, D. Drenckhahn, P. Hinterdorfer, Nano-scale dynamic recognition imaging on vascular endothelial cells. *Biophys. J.* **93**, L11–13 (2007).
 82. A. Ebner, F. Kienberger, G. Kada, C.M. Stroh, M. Geretschlager, A.S. Kamruzzahan, L. Wildling, W.T. Johnson, B. Ashcroft, J. Nelson, et al. Localization of single avidin-biotin interactions using simultaneous topography and molecular recognition imaging. *Chemphyschem.* **6**, 897–900 (2005).
 83. C. Stroh, H. Wang, R. Bash, B. Ashcroft, J. Nelson, H. Gruber, D. Lohr, S.M. Lindsay, P. Hinterdorfer, Single-molecule recognition imaging microscopy. *Proc. Natl. Acad. Sci. USA* **101**, 12503–12507 (2004).
 84. H. Wang, Y. Dalal, S. Henikoff, S. Lindsay, Single-epitope recognition imaging of native chromatin. *Epigenetics Chromatin* **1**, 10 (2008).
 85. S. Araki, T. Nakai, K. Hizume, K. Takeyasu, K. Yoshikawa, Hydrodynamic radius of circular DNA is larger than that of linear DNA. *Chem. Phys. Lett.* **418**, 255–259 (2006).
 86. N. Makita, K. Yoshikawa, ATP/ADP switches the higher-order structure of DNA in the presence of spermidine. *Febs. Letters* **460**, 333–337 (1999).
 87. N. Makita, K. Yoshikawa, Proton concentration (pH) switches the higher-order structure of DNA in the presence of spermine. *Biophys. Chem.* **99**, 43–53 (2002).
 88. K. Yoshikawa, Y. Matsuzawa, Nucleation and growth in single DNA molecules. *J. Am. Chem. Soc.* **118**, 929–930 (1996).
 89. M. Ito, A. Sakakura, N. Miyazawa, S. Murata, K. Yoshikawa, Nonspecificity induces chiral specificity in the folding transition of giant DNA. *J. Am. Chem. Soc.* **125**, 12714–12715 (2003).
 90. T. Iwataki, S. Kidoaki, T. Sakaue, K. Yoshikawa, S.S. Abramchuk, Competition between compaction of single chains and bundling of multiple chains in giant DNA molecules. *J. Chem. Phys.* **120**, 4004–4011 (2004).
 91. N. Yoshinaga, K. Yoshikawa, S. Kidoaki, Multiscaling in a long semiflexible polymer chain in two dimensions. *J. Chem. Phys.* **116**, 9926–9929 (2002).
 92. M.E. Mysiak, M.H. Bleijenberg, C. Wyman, P.E. Holthuizen, P.C. van der Vliet, Bending of adenovirus origin DNA by nuclear factor I as shown by scanning force microscopy is required for optimal DNA replication. *J. Virol.* **78**, 1928–1935 (2004a).
 93. M.E. Mysiak, C. Wyman, P.E. Holthuizen, P.C. van der Vliet, NFI and Oct-1 bend the Ad5 origin in the same direction leading to optimal DNA replication. *Nucleic Acids Res.* **32**, 6218–6225 (2004b).
 94. A. Janicijevic, K. Sugawara, Y. Shimizu, F. Hanaoka, N. Wijgers, M. Djurica, J.H. Hoeijmakers, C. Wyman, DNA bending by the human damage recognition complex XPC-HR23B. *DNA Repair (Amst)* **2**, 325–336 (2003).
 95. I. Ugrinova, E. Mitkova, C. Moskalenko, I. Pashev, E. Pasheva, DNA bending versus DNA end joining activity of HMGB1 protein is modulated in vitro by acetylation. *Biochemistry* **46**, 2111–2117 (2007).
 96. L.I. Pietrasanta, D. Thrower, W. Hsieh, S. Rao, O. Stemann, J. Lechner, J. Carbon, H. Hansma, Probing the *Saccharomyces cerevisiae* centromeric DNA (CEN DNA)-binding factor 3 (CBF3) kinetochore complex by using atomic force microscopy. *Proc. Natl. Acad. Sci. USA* **96**, 3757–3762 (1999).

97. S.H. Yoshimura, H. Maruyama, F. Ishikawa, R. Ohki, K. Takeyasu, Molecular mechanisms of DNA end-loop formation by TRF2. *Genes Cells* **9**, 205–218 (2004).
98. K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
99. M.J. Allen, X.F. Dong, T.E. O'Neill, P. Yau, S.C. Kowalczykowski, J. Gatewood, R. Balhorn, E.M. Bradbury, Atomic force microscope measurements of nucleosome cores assembled along defined DNA sequences. *Biochemistry* **32**, 8390–8396 (1993).
100. C. Zheng, J.J. Hayes, Structures and interactions of the core histone tail domains. *Biopolymers* **68**, 539–546 (2003).
101. A. Kimura, K. Matsubara, M. Horikoshi, A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. *J. Biochem. (Tokyo)* **138**, 647–662 (2005).
102. S.J. Nowak, V.G. Corces, Phosphorylation of histone H3: A balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* **20**, 214–220 (2004).
103. Y. Zhang, D. Reinberg, Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15**, 2343–2360 (2001).
104. J.J. Hayes, D.J. Clark, A.P. Wolffe, Histone contributions to the structure of DNA in the nucleosome. *Proc. Natl. Acad. Sci. USA* **88**, 6829–6833 (1991).
105. K. Hizume, S.H. Yoshimura, H. Maruyama, J. Kim, H. Wada, K. Takeyasu, Chromatin reconstitution: Development of a salt-dialysis method monitored by nano-technology. *Arch. Histol. Cytol.* **65**, 405–413 (2002).
106. K. Hizume, T. Nakai, S. Araki, E. Prieto, K. Yoshikawa, K. Takeyasu, Removal of histone tails from nucleosome dissects the physical mechanisms of salt-induced aggregation, linker histone H1-induced compaction and 30-nm fiber formation of the nucleosome array. *Ultramicroscopy* (2009) *In Press*, doi:10.1016/j.ultramic.2009.1003.1014.
107. J.O. Thomas, C.M. Wilson, Selective radiolabelling and identification of a strong nucleosome binding site on the globular domain of histone H5. *EMBO J.* **5**, 3531–3537 (1986).
108. L.M. Schang, G.J. Hwang, B.D. Dynlacht, D.W. Speicher, A. Bantly, P.A. Schaffer, A. Shilatifard, H. Ge, R. Shiekhattar, Human PC4 is a substrate-specific inhibitor of RNA polymerase II phosphorylation. *J. Biol. Chem.* **275**, 6071–6074 (2000).
109. M.Y. Kim, S. Mauro, N. Gevry, J.T. Lis, W.L. Kraus, NAD⁺-dependent modulation of chromatin structure and transcription by nucleosome binding properties of PARP-1. *Cell* **119**, 803–814 (2004b).
110. B.G. De Groot, C.A. Putman, High-resolution imaging of chromosome-related structures by atomic force microscopy. *J. Microsc.* **168**, 239–247 (1992).
111. C.A. Putman, B.G. De Groot, J. Wiegant, A.K. Raap, K.O. Van der Werf, N.F. Van Hulst, J. Greve, Detection of in situ hybridization to human chromosomes with the atomic force microscope. *Cytometry* **14**, 356–361 (1993).
112. O. Hoshi, T. Ushiki, Three-dimensional structure of G-banded human metaphase chromosomes observed by atomic force microscopy. *Arch. Histol. Cytol.* **64**, 475–482 (2001).

113. G. Holmquist, M. Gray, T. Porter, J. Jordan, Characterization of Giemsa dark- and light-band DNA. *Cell* **31**, 121–129 (1982).
114. T. Kobori, M. Kodama, K. Hizume, S.H. Yoshimura, T. Ohtani, K. Takeyasu, Comparative structural biology of the genome: nano-scale imaging of single nucleus from different kingdoms reveals the common physicochemical property of chromatin with a 40 nm structural unit. *J. Electron Microsc. (Tokyo)* **55**, 31–40 (2006).
115. L.L. Soon, C. Bottema, W.G. Breed, Atomic force microscopy and cytochemistry of chromatin from marsupial spermatozoa with special reference to *Sminthopsis crassicaudata*. *Mol. Reprod. Dev.* **48**, 367–374 (1997).
116. O. Hoshi, M. Shigeno, T. Ushiki, Atomic force microscopy of native human metaphase chromosomes in a liquid. *Arch. Histol. Cytol.* **69**, 73–78 (2006).
117. C.R. Woese, Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. USA* **97**, 8392–8396 (2000).
118. C.R. Woese, G.E. Fox, Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. USA* **74**, 5088–5090 (1977).
119. C.R. Woese, O. Kandler, M.L. Wheelis, Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* **87**, 4576–4579 (1990).
120. C. Robinow, E. Kellenberger, The bacterial nucleoid revisited. *Microbiol Rev.* **58**, 211–232 (1994).
121. S.B. Zimmerman, Cooperative transitions of isolated *Escherichia coli* nucleoids: Implications for the nucleoid as a cellular phase. *J. Struct. Biol.* **153**, 160–175 (2006).
122. T. Ali Azam, A. Iwata, A. Nishimura, S. Ueda, A. Ishihama, Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* **181**, 6361–6370 (1999).
123. J. Kim, S.H. Yoshimura, K. Hizume, R.L. Ohniwa, A. Ishihama, K. Takeyasu, Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. *Nucleic Acids Res.* **32**, 1982–1992 (2004a).
124. T.A. Azam, A. Ishihama, Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.* **274**, 33105–33113 (1999).
125. S. Maurer, J. Fritz, G. Muskhelishvili, A systematic in vitro study of nucleoprotein complexes formed by bacterial nucleoid associated proteins revealing novel types of DNA organization. *J. Mol. Bio.* (2009). doi:10.1016/j.jmb.2009.1002.1050.
126. K. Takeyasu, J. Kim, R.L. Ohniwa, T. Kobori, Y. Inose, K. Morikawa, T. Ohta, A. Ishihama, S.H. Yoshimura, Genome architecture studied by nanoscale imaging: analyses among bacterial phyla and their implication to eukaryotic genome folding. *Cytogenet. Genome Res.* **107**, 38–48 (2004).
127. J. van Noort, S. Verbrugge, N. Goosen, C. Dekker, R.T. Dame, Dual architectural roles of HU: Formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. USA* **101**, 6969–6974 (2004).
128. R.T. Dame, C. Wyman, N. Goosen, H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.* **28**, 3504–3510 (2000).
129. M.S. Luijsterburg, M.F. White, R. van Driel, R.T. Dame, The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* **43**, 393–418 (2008).

130. R.T. Dame, M.C. Noom, G.J. Wuite, Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444**, 387–390 (2006).
131. M.C. Noom, W.W. Navarre, T. Oshima, G.J. Wuite, R.T. Dame, H-NS promotes looped domain formation in the bacterial chromosome. *Curr. Biol.* **17**, R913–914 (2007).
132. D.C. Grainger, D. Hurd, M.D. Goldberg, S.J. Busby, Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res.* **34**, 4642–4652 (2006).
133. D.C. Grainger, D. Hurd, M. Harrison, J. Holdstock, S.J. Busby, Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc. Natl. Acad. Sci. USA* **102**, 17693–17698 (2005).
134. T. Oshima, S. Ishikawa, K. Kurokawa, H. Aiba, N. Ogasawara, *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res.* **13**, 141–153 (2006).
135. D. Marenduzzo, I. Faro-Trindade, P.R. Cook, What are the molecular ties that maintain genomic loops? *Trends. Genet.* **23**, 126–133 (2007).
136. J.M. Lucht, P. Dersch, B. Kempf, E. Bremer, Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of *Escherichia coli*. *J. Biol. Chem.* **269**, 6578–6578 (1994).
137. H. Niki, A. Jaffe, R. Imamura, T. Ogura, S. Hiraga, The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of *E. coli*. *Embo. J.* **10**, 183–193 (1991).
138. A. Oguro, H. Kakeshita, H. Takamatsu, K. Nakamura, K. Yamane, The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit, on *Bacillus subtilis* growth and protein translocation. *Gene* **172**, 17–24 (1996).
139. H. Niki, R. Imamura, M. Kitaoka, K. Yamanaka, T. Ogura, S. Hiraga, *E. coli* MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. *Embo. J.* **11**, 5101–5109 (1992).
140. R.A. Britton, D.C. Lin, A.D. Grossman, Characterization of a prokaryotic SMC protein involved in chromosome partitioning. *Genes Dev.* **12**, 1254–1259 (1998).
141. S. Moriya, E. Tsujikawa, A.K. Hassan, K. Asai, T. Kodama, N. Ogasawara, A *Bacillus subtilis* gene-encoding protein homologous to eukaryotic SMC motor protein is necessary for chromosome partition. *Mol. Microbiol.* **29**, 179–187 (1998).
142. R.L. Ohniwa, K. Morikawa, S.L. Takeshita, J. Kim, T. Ohta, C. Wada, K. Takeyasu, Transcription-coupled nucleoid architecture in bacteria. *Genes Cells* **12**, 1141–1152 (2007).
143. L.D. Murphy, S.B. Zimmerman, Isolation and characterization of spermidine nucleoids from *Escherichia coli*. *J. Struct. Biol.* **119**, 321–335 (1997).
144. R.L. Ohniwa, K. Morikawa, J. Kim, T. Ohta, A. Ishihama, C. Wada, K. Takeyasu, . Dynamic state of DNA topology is essential for genome condensation in bacteria. *Embo. J.* **25**, 5591–5602 (2006).
145. K. Morikawa, R.L. Ohniwa, J. Kim, A. Maruyama, T. Ohta, K. Takeyasu, Bacterial nucleoid dynamics: oxidative stress response in *Staphylococcus aureus*. *Genes Cells* **11**, 409–423 (2006).

146. M. Almiron, A.J. Link, D. Furlong, R. Kolter, A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev.* **6**, 2646–2654 (1992).
147. S. Nair, S.E. Finkel, Dps protects cells against multiple stresses during stationary phase. *J. Bacteriol.* **186**, 4192–4198 (2004).
148. P. Ceci, S. Cellai, E. Falvo, C. Rivetti, G.L. Rossi, E. Chiancone, DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Res.* **32**, 5935–5944 (2004).
149. D.R. Edgell, W.F. Doolittle, Archaea and the origin(s) of DNA replication proteins. *Cell* **89**, 995–998 (1997).
150. A. Hirata, B.J. Klein, K.S. Murakami, The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**, 851–854 (2008).
151. A.G. Kusser, M.G. Bertero, S. Naji, T. Becker, M. Thomm, R. Beckmann, P. Cramer, Structure of an archaeal RNA polymerase. *J. Mol. Biol.* **376**, 303–307 (2008).
152. K. Sandman, J.A. Krzycki, B. Dobrinski, R. Lurz, J.N. Reeve, Hmf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus fervidus*, is most closely related to histones. *Proc. Natl. Acad. Sci. USA* **87**, 5788–5791 (1990).
153. R.J. DeLange, G.R. Green, D.G. Searcy, A histone-like protein (HTa) from *Thermoplasma acidophilum*. I. Purification and properties. *J. Biol. Chem.* **256**, 900–904 (1981).
154. K. Decanniere, A.M. Babu, K. Sandman, J.N. Reeve, U. Heinemann, Crystal structures of recombinant histones HmfA and HmfB from the hyperthermophilic archaeon *Methanothermus fervidus*. *J. Mol. Biol.* **303**, 35–47 (2000).
155. S.L. Pereira, R.A. Grayling, R. Lurz, J.N. Reeve, Archaeal nucleosomes. *Proc. Natl. Acad. Sci. USA* **94**, 12633–12637 (1997).
156. M. Noll, Subunit structure of chromatin. *Nature* **251**, 249–251 (1974).
157. M. Tomschik, M.A. Karymov, J. Zlatanova, S.H. Leuba, The archaeal histone-fold protein Hmf organizes DNA into bona fide chromatin fibers. *Structure* **9**, 1201–1211 (2001).
158. R.A. Grayling, K.A. Bailey, J.N. Reeve, DNA binding and nuclease protection by the Hmf histones from the hyperthermophilic archaeon *Methanothermus fervidus*. *Extremophiles* **1**, 79–88 (1997).
159. S.L. Pereira, J.N. Reeve, Archaeal nucleosome positioning sequence from *Methanothermus fervidus*. *J. Mol. Biol.* **289**, 675–681 (1999).
160. F. Bocard, E. Esnault, M. Valens, Spatial arrangement and macrodomain organization of bacterial chromosomes. *Mol. Microbiol.* **57**, 9–16 (2005).
161. A. Worcel, E. Burgi, On the structure of the folded chromosome of *Escherichia coli*. *J. Mol. Biol.* **71**, 127–147 (1972).
162. A. Rodriguez-Campos, F. Azorin, RNA is an integral component of chromatin that contributes to its structural organization. *PLoS ONE* **2**, e1182 (2007).
163. C. Maison, D. Bailly, A.H. Peters, J.P. Quivy, D. Roche, A. Taddei, M. Lachner, T. Jenuwein, G. Almouzni, Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* **30**, 329–334 (2002).

Near-Field Optical Litography

Eugenio Cefali, Salvatore Patanè, and Maria Allegrini

Summary. This chapter reviews a specific application of scanning near-field optical microscopy (SNOM) to lithography. The working principles of the conventional lithographic techniques and related materials as well as of the SNOM technique are recalled. Detailed results of the aperture and apertureless scanning near-field optical lithography (SNOL) with regard to resolution and effectiveness of the method are dealt with. In particular, attention is focused on SNOL results on azo-polymers. The photo-printing mechanism of those polymers has in return allowed the study of the near-field fundamental features as well as of the probes.

Key words: near-field optics, data storage, polymers, lithography.

Abbreviations

AFM	Atomic Force Microscope
CAMP	Chemical Amplified Photoresist
DOF	Depth of Focus
DUV	Deep Ultraviolet
EM	Electromagnetic
EUV	Extreme Ultraviolet
FE	Field Enhancement
EBL	Electron Beam Lithography
HPC	Hollow Pyramid Cantilevered probe
IBL	Ion Beam Lithography
IOL	Immersion Optics Lithography
MEMS	Micro Electro-Mechanical Systems
PMMA	Polymethyl-methacrylate
RASC	Radiation sensitive compound
SCALPEL	Scattering with Angular Limitation Projection Electron Lithography
SFM	Scanning Force Microscopy
SNOL	Scanning near-field optical lithography

SNOM	Scanning near-field optical microscopy
SPL	Scanning Probe Lithography
STM	Scanning Tunnelling Microscope
TIR	Total Internal Reflection

21.1 Introduction

A noticeable consequence of modern technology is the miniaturization of electronic. The scaling down in semiconductors and in data storage fields leads to enormous benefits in terms of speed and energy consumed for computing function [1–3]. In the past 40 years, following the Moore’s law [4], the dimension of integrated circuits has been shrinking at a rate of 30% every 3 years. In this framework, lithography is an important keyword. The progress in microelectronics is controlled by the rate of advances in microlithographic tools, methods, and materials. Today, the photolithography technologies, that have served the microelectronics community since its beginnings, are approaching their limit of extendibility. If the historical growth rate has to be maintained in the future, new imaging technologies with the capability of forming sub-50 nm features must be devised and refined. *Scanning near-field optical lithography* potentially represents a good candidate. It consists of the specific application of the *Scanning near-field optical microscopy* for patterning photoactive materials. The direct advantage of SNOL compared with other kinds of lithography is to “be an optics-based technique” without diffraction limited resolution. Indeed, in the case of near-field illumination, the spatial confinement of the near-field light–matter interaction is not limited by the light wavelength but rather by both the probe size and the probe-to-matter distance. Although the SNOL is in principle a serial technique, some setups have been advanced with a semi-parallel approach [5]. A full-parallel SNOL however is still in a developing stage. Furthermore, as it will be shown later, the SNOL is a devoted tool not only for lithography but also for optical data storage. In the first part of this chapter, a brief overview on the conventional lithography techniques and related materials is presented. The second part of the chapter is focused on the SNOL working principle and the results attained on polymers films.

21.2 Lithography: Principles and Materials

Lithography is basically defined as the *printing process in which the image to be produced is rendered on a flat surface, and treated to retain ink while the nonimage areas are treated to repel ink* [6]. In its current modern usage, the term is more generally applied to a number of methods for replicating a predetermined master pattern on a substrate. As it will be further described, the results achievable by lithographic process strictly depend on the features

of the writing techniques and on the features of the substrate. Generally, the “*printing*” is achieved by first coating the substrate with a thin film called *resist* (e.g. *polymer film*) and using a localized interaction between this layer and an engraving micro-tool, particles, or radiation beam. Various lithographic techniques can be classified according to the micro-tool or the type of radiation used. Photolithography methods consist in exposing a photosensitive material to electromagnetic radiation (UV, DUV, EUV, or X-ray). The irradiation produces chemical changes in the resist layer, increasing or constricting the solubility of exposed areas. The spatial distribution of the chemical changes reproduces, in a latent image, the mask features. The latent image requires a chemical post-processing to be fixed (develop). An enriched solubility gives a positive image of the mask after, whereas a solubility decrease yields a negative-tone image of the mask. Finally, the obtained latent image is then developed into relief structures by means of etching process [7,8]. Figure 21.1 schematically depicts the lithographic process.

Methods based on writing with particles usually accomplish the same goal. They make use of energetic particles (electrons or ions) rather than photons. Then, in general the expected pattern is usually obtained by interposing a mask between the source of radiation (or particles) and the resist. In this case, the writing method is a parallel process and the whole pattern is made simultaneously by firmly *replicating* (perhaps reduced in size) the mask pattern. The parallel lithography is a fast technique since a whole wafer can be exposed in a single stage. The weakness is that one must first make the mask containing all the patterns that need to be reproduced. The mask serves as a template that can be reused several times. A different method consists in scanning a focused particles beam across the surface of the material. This is a sequential writing method. The patterns are written point by point on the resist surface; therefore these techniques are much slower than parallel ones. On the other hand, there is absolutely no need to produce a mask as template for implementing the lithography stage. The main tasks to be pursued by the lithographic process are: the capability to fabricate pattern structure as small as possible (i.e., resolution assessment); the capability to exposes as big as possible resist area per unit time (i.e., high-writing speed). Unfortunately, these two parameters are difficult to reconcile because they are somehow competing one with each other. In fact, the extremely fast parallel techniques, such as optical lithography often result in poor resolution (> 50 nm), limited by diffraction effects [9]. Conversely, serial methods, like electron beam lithography, with particularly high resolution (< 10 nm) require very slow writing speeds.

Current efforts are focused on set novel technological tools able to solve the problem of reconciling resolution and speed, such as SPL, nanoimprinting lithography, EUV Radiation lithography, soft lithography, and others. These technologies cannot yet be used for large-scale mass production [10–17].

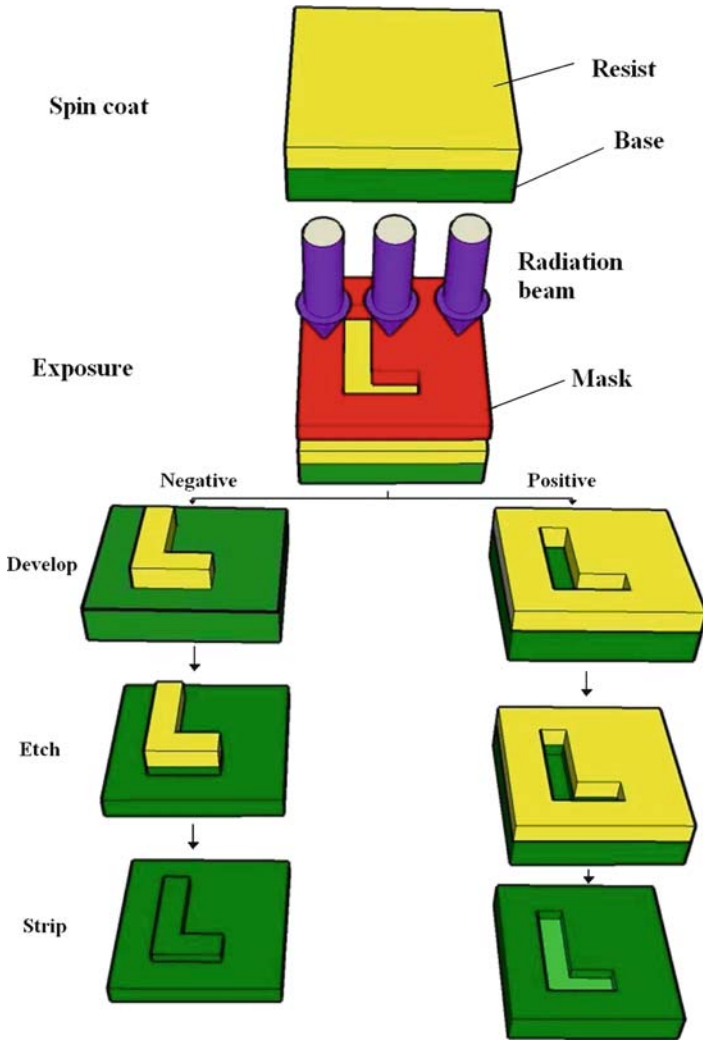


Fig. 21.1. Schematic representation of the lithography process

21.2.1 Photolithography

Photolithography can be performed by employing light with wavelength ranging from UV to X-ray. The choice depends on the demand of structures as small as possible. Photolithography with 193-nm light from ArF excimer laser and synthetic fused silica (or calcium fluoride) lenses yields a reliable and somehow economical solution to patterning features as small as ~ 150 nm [18–23]. The photo-mask lithography is carried out using two different schemes, i.e., contact (or proximity) and printing (or projection printing) [7]. In the first case, an UV light irradiates the photomask which is in physical

contact with, or in close proximity to, the resist film. The mask has opaque and transparent parts that reproduce the relevant pattern. It is generally a quartz plate coated with chromium in those regions where opacity is required. This very simple technique represented the foundation of microfabrication until the mid-1970s. The minimum feature size that can be obtained by this process is primarily determined by the diffraction that occurs as light passes through the gap between the mask and the resist. Even with the use of elaborate vacuum systems to pull the mask and substrate together, it is still difficult to reduce the gap between a conventional rigid mask and a rigid, flat substrate to less than $\sim 1\ \mu\text{m}$ over large areas. As a result, the resolution of contact mode photolithography is typically $0.5\text{--}0.8\ \mu\text{m}$ when UV light of $360\text{--}460\ \text{nm}$ is used. Contact-mode photolithography is not seriously considered as a generally useful technique because the required mechanical contact can damage the fragile structures on the mask or on the sample. The mechanical problems encountered with contact and proximity lithography stimulated the development of projection lithography in the mid-1980s. The mask and wafer are held some distance apart and an optical system is inserted to focus the image of the mask on the sample. In this way, a high-quality image usually demagnified by a factor of 4, is achievable, and alignment or registration can be relatively straightforward. Basically, all integrated circuits are manufactured by projection photolithography, using equipment known as “step-and-repeat machine” (or stepper) [24]. The whole wafer cannot be exposed in a single step, because the field of projection is much smaller than the wafer, therefore the sample or the optics is moved. In the *scanning projection printing*, the sample and mask are moved simultaneously and continuously in front of the optical setup. Reflective optics are used, or a combination of reflective and refractive optics, but without magnification, so as to make simpler the displacement operation. The advantage with this technique is to use only the best zone of the optics providing an excellent definition. The disadvantage is that there is no reduction of the mask.

The minimum size of the writable structure strictly depends on the optical resolution ($R = k_1 * \lambda / \text{NA}$) [25] and on the depth of focus ($\text{DOF} = k_2 * \lambda / \text{NA}^2$) of the optical system. Both the physical quantities depend on the wavelength (λ) of the illuminating EM radiation and on the numerical aperture of the lens system (NA). The resolution depends on the constant k_1 and is limited by *Rayleigh diffraction* [25]. DOF depends on the aberrations of the imaging system represented by the constant k_2 that ranges between 1 and 2. This quantity is also important because it governs the spatial depth over which the image of the mask is faithfully produced. Increasing the DOF to acceptable levels poses a technical challenge for projection photolithography. Practically, there is a notable blurring and distortion of the image for distances beyond the DOF from the image plane. In order to keep the projected image not degraded over a distance of $\pm 0.5\ \mu\text{m}$ from the image plan, a DOF larger than $1\ \mu\text{m}$ is useful to get some insensitivity to slight variations in the thickness of the resist or the position of the surface of the substrate.

The goal is to improve the resolution while maintaining a maximum DOF. Resolution improvements may be realized by reducing the wavelength (λ) of actinic exposure radiation either by increasing the NA or by process and tool refinements that reduce the value of k_1 . There are practical limits to each of these approaches, however. Improved resolution by increasing the NA of the lens system requires that either the diameter of the lens be increased or the focal length be reduced. The result is increased lens size and mass, along with increased complexity of the overall lens system to provide correction for optical aberrations. Due to the DOF definition, a resolution improvement achieved via increased NA is accompanied by a proportionately larger decrease in DOF. Improved resolution achieved reducing the wavelength does not produce a great penalty in DOF, but it requires a suitable photoresist, light source, optical lens. The development of new generations of photoresist for leading edge microlithography applications has been driven in large part by the need of materials optimized for a specific wavelength. Suitable techniques such as Immersion Optics Lithography, EUV Lithography, X-ray lithography and Resolution Enhancement Technique (RET) [26, 27] have been developed for this purpose. A comprehensive description of these approaches is reported elsewhere [28–33].

21.2.2 Electron Beam Lithography

In EBL [34], a flow of electrons with kinetic energies in the range of 1–100 keV is brought into focus on a resist-coated substrate. Collisions of the electrons (primary and secondary) with atoms in the resist film create reactive intermediates whose deactivation chemistry changes the properties of the resist film in the irradiated regions because of local changes in its solubility. In the case of an electron-sensitive resist film of PMMA [35–37], the electrons induce local chain breaking and formation of micropores that causes the material to be soluble in a developer that consists of methylisobutylketone and 2-propanol [7]. High-energy electron beam guarantees an extremely high-writing resolution because of the small *de Broglie* wavelength (for example, $\lambda \sim 0.005$ nm for 50 keV electrons). In practice, it is possible to obtain e-beam with focused spot ranging from 0.5 to 0.25 nm [38]. Patterns with features as small as ~ 50 nm can be routinely generated by e-beam writing [39–41]. Moreover, by using thin (10–100 nm) membranes as substrates to minimize the electron scattering problem, structures as small as ~ 2 nm have been obtained [42–44]. The EBL highest resolution, rather than by electron diffraction, is limited by residual aberrations in the electron optics and by electron–electron repulsion at high-beam currents that blur the beam (space-charge effect). Unfortunately, EBL is a direct-write technology producing pattern by exposing a film of resist in a serial manner. It means a long writing time (approximately 1 h) per wafer (4 in.) and therefore the method is inadequate for large-scale production [35]. However, EBL is fruitfully applied for producing photomasks in optical lithography or to produce small numbers of nanostructures for research purposes.

It is also widely used for fabrication of high-frequency GaAs field-effect transistor (FET) devices that require resolution down to ~ 100 nm [35]. Finally, EBL has been implemented in *projection technique* configuration, so-called SCALPEL. Such a technique gets both the resolution of the approaching direct-write systems and the parallelism of projection lithographic systems. SCALPEL has been intensely developed for mass production of patterned nanostructures [35].

21.2.3 Ion Beam Lithography

The ion beam lithography [45] is also an attractive tool for writing nanostructures. Its working principle is closely related to that of EBL. The impinging electron beam is replaced by a beam of accelerated ions, for example, Ga^+ , H^+ , or He^+ . Likewise in EBL, the ions generate low-energy secondary electrons inside the resist film, setting off the chemical reactions that lead to lithographic imaging. Accordingly, a *resist* material displaying suitable lithographic properties for EBL can also be used for IBL. Such a kind of lithography has been put into practice in several forms: Focused IBL [8, 46, 47], masked IBL [48], projection IBL [49]. FIB lithography is a serial writing process able to produce features as small as ~ 6 nm. Such a result has been reached by using a 50 kV Ga^+ two-lens system [40]. In masked IBL [48], a collimated ion beam is passed through a thin ion-transparent membrane patterned with an absorber material and the transmitted beam exposes a resist-coated substrate in proximity to the mask. This technique is in many ways equivalent to optical projection exposure tools [50]. A divergent ion beam illuminates a mask and is then focused and demagnified by a set of projection ion optics onto a resist-coated substrate. The mask in this case is a stencil, it has physical openings to selectively transmit portions of the ion beam.

Ion beam lithography has at least two advantages over electron-beam lithography: (a) IBL provides a higher (approximately 2 orders of magnitude) resist exposure sensitivity than focused electron beam. It takes place because the energy loss from ions in a resist material is much more efficient than the energy loss from electrons [51]; (b) IBL has a negligible ion scattering in the resist and very low backscattering from the substrate. For these reasons the penetration of the particle beam into the substrate is much less than that with EBL, reducing the image blurring due to scattering, which causes a proximity effect. Second, the more efficient energy deposition effectively yields an increased radiation sensitivity for a given resist. However, while improved sensitivity is desirable in most circumstances, in the IBL case a high-sensitivity resist exposed at the correct dose might require so few ions that statistical fluctuations in the number of ions per pixel leads to poor control. Another problem of IBL is the damage induced to samples by high-energy ions. Nevertheless, IBL and FIB in particular is expected to remain valuable for production of patterned nanostructures in resist films, patterned implantation, repair of defects in photo and X-ray masks, inspection in integrated

circuits processes, failure analysis, and surface characterization (for example, in secondary ion mass spectroscopy, SIMS) [8, 52].

21.2.4 Materials

As already mentioned, the substrate designated for patterning process has to be previously coated with a special material called *resist*, which works as mediator for the functional patterning process. The *resist* (or photoresist) usually is made of polymeric materials that are modified when exposed to radiation (either visible, ultraviolet, X-ray photons) or to energetic electron beams or ions. A photoresist consists of two parts, namely a radiation sensitive compound, which also accelerates or slows down the rate at which the resist dissolves in a solvent, and an alkaline-soluble resin, which is insensitive to the writing radiation and fulfils the mechanical requirements of the resist. The RASC, mixed into the resin, make it insoluble. As a consequence of the coated-substrate exposition to the desired patterned-radiation, the RASC is converted into an acid that makes the resin soluble. In particular, the exposure can increase or decrease the solubility of the resist. In the first case, it deals with positive resist, while in the second case it deals with negative resist. However, after irradiation, the resist can be dissolved and the exposed substrate beneath the resist is chemically etched or metallically coated to match the projected design. It is important to underline that achievements attainable by the lithographic techniques also depends on resist features. In particular, they deal with sill life, availability, cost, purity, etching resistance, resolution sensitivity, and contrast [53]. The last two are the most important parameters. The sensitivity of a resist is usually quoted as the dose ($[D] = \text{mJ}/\text{cm}^2$) of radiation or interaction necessary to induce sufficient modification of the resist to ensure that the desired pattern appears at the development stage. Thus, if we plot the thickness remaining after development versus $\log_{10} D$, we get curves for both negative (exposed areas insoluble in developer) and positive (Fig. 21.2). Usually, for a positive resist, the required dose is slightly

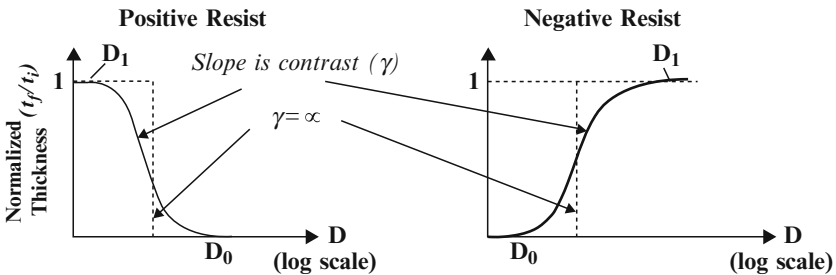


Fig. 21.2. Contrast curve of resist material. Positive resist is softened during exposure and the contrast decreases. Negative resist hardens during exposure and the contrast increases

above the minimum dose required to remove all the resist and similarly for a negative tone. The mean slope of the curve between full thickness and zero thickness is the contrast of the resist ($\gamma = (\log_{10} D_0 - \log_{10} D_1)^{-1}$). Typical values for γ ranges from 2 to 3. Hence, D_1 is $10^{1/3}$ to $10^{1/2}$ times larger than D_0 , e.g., $D_0 \leq 50 \text{ mJ/cm}^2$ and $D_{100} \geq 150 \text{ mJ/cm}^2$. Clearly, for an expected binary image the higher γ the better is. Early resists had a γ of about one, and so the optical image at the wafer (“aerial image”) should have high contrast (i.e., $(I_{\max} - I_{\min})/(I_{\max} + I_{\min}) \cong 1$) to compensate.

As reported in the literature [1], such high contrast in the aerial image is not available for small features (high-spatial frequency), so the resist chemists have come to the rescue of the optical physicists by providing resist with γ -values of ten or more and at the same time by increasing the writing speed. The resist contrast characterises the variation of the solubility rate in its developer as a function of the exposure time (dose). The higher the contrast the better the resist will be able to reveal small variations in the received dose. This is a crucial feature of the resist. Indeed, as we have previously seen, no matter what type of lithography is used, the spatial localization of the exposure on the resist never cuts off abruptly. Because of various physical effects that depend on the type of radiation or interaction used (diffraction for photons, collisions for electrons, etc.), the actually exposed region of the resist extends slightly beyond the intended patterns to include a transition zone that varies in width. These transition zones determine the spatial resolution of the lithography process. It is intuitively clear that, the higher the contrast of the resist the less these edge effects will contribute to patterns spreading. It is therefore a priority to find high-contrast resists.

The most common *resists* are substantially made with one of the following chemical compounds: *diazonaphthoquinone* (DNQ)/*phenol-formaldehyde resin* (Novolak), PMMA, SU-8, *PolyMethyl-GlutarImide* (PMGI). The DNQ/Novolak photoresist represents the bulk resist for the world’s semiconductors manufacturing. It is a positive photoresist used with the *I*-, *G*-, and *H*-lines from a mercury-vapor lamp (UV wavelengths of 365 or 436 nm). The DNQ (the photoactive material) inhibits the dissolution of the novolak resin (matrix). However, upon exposure to light, the dissolution rate increases even beyond that of pure novolak. The mechanism by which unexposed DNQ inhibits novolak dissolution is not well understood, but is believed to be related to hydrogen bonding (or more exactly di-azocoupling in the unexposed region). DNQ-novolak resists are developed by dissolution in a basic solution (usually 0.26N tetra-methyl ammonium hydroxide in water). The performance of these materials has been continually improved by careful optimization of the polymer and DNQ dissolution inhibitor. Today, highly optimized DNQ-novolak resists are used to manufacture integrated circuits with device geometries as small as 350 nm. Unfortunately, these resists are not suitable for use at DUV wavelengths ($\lambda \approx 250 \text{ nm}$). For DUV lithography, the resist are typically polyhydroxystyrene-based polymers with a photoacid generator providing the solubility change. The combined benzene-chromophore

and DNQ-novolac absorption mechanisms lead to stronger absorption in the DUV, and require a much larger exposure. The strong DUV absorption results in diminished photoresist sensitivity. The SU-8, an epoxy-based polymer that can be used to pattern high-aspect ratio (> 20) structures [54], is one of the most commonly used negative photoresist. Its maximum absorption is for ultraviolet light with a wavelength of 365 nm. When exposed, SU-8's long molecular chains cross-link causes the solidification of the material. Moreover, it has a unique property, i.e., it is very difficult to strip. SU-8 is mainly used for fabrication of microfluids and MEMS parts. It is also one of the most bio-compatible materials and is often used in bio-MEMS. Current negative photoresist exhibits better adhesion than positive-tone photoresists to various substrates such as Si, GaAs, InP, glass, and metals including Au, Cu, and Al. Additionally, the current generation of G-, H-, and I-line negative-tone photoresists exhibits higher temperature resistance over positive resists.

For EBL, in principle, it is possible to use photoresists. The exposure of such materials to the electron beam produces the same results as exposure to light. The main difference is that while photons are absorbed, depositing all their energy at once, electrons deposit their energy gradually and scatter within the photoresist during this process. As with high-energy wavelengths, many transitions are excited by electron beams, and heating and out-gassing are still concerns. The dissociation energy for a C–C bond is 3.6 eV. Secondary electrons generated by primary ionizing radiation have energies sufficient to dissociate this bond, causing chain-scission [53]. In addition, the low-energy electrons have a longer photoresist interaction time due to their lower speed. Scission breaks the original polymer into segments of lower molecular weight, which are more readily dissolved in a solvent. An example of polymer suitable for EBL is the PMMA. Electron beams at low-energy (10–50 eV) are able to damage ~ 30 -nm thick PMMA films. The damage manifests as a loss of material. Nonetheless, specific resists has been developed for EBL. One of the most used is ZEP-520. It is a very high-resolution positive tone resist, that like PMMA, it is simple to use and gives reproducible results. Compared to PMMA, it has the advantage of requiring a three times faster radiation exposure and has good dry etch resistance. It has the disadvantage of poor adhesion (requires HMDS prime layer) and normal exposure doses result in reentrant pattern profiles. By using ZEP-520 a pitch resolution limit of 60 nm (30-nm lines and spaces), independent of thickness and beam energy, was found [55]. Others resists suitable for EBL are based on inorganic films as AlF_3 , SiO_2 , and MgO [56]. For Ion beam lithography, resist films based on refractory metal oxides (MoO_3 and WO_3) have been developed. Depending on the preparation condition of the films, these thin oxide films act either as a negative or a positive resist with high-contrast capability [57].

Improvements, in terms of efficiency, have been obtained by developing CAMP. Such resists have been produced for DUV, for electron beam and for shorter wavelengths. This is done in order to combat the larger absorption at shorter wavelengths. In fact, the sensitivity of a photoresist

is measured by its quantum efficiency or by the number of chemical events that occur when a photon is absorbed by the material. In CAMP material (e.g., t-butoxycarbonyloxystyrene-sulfur dioxide) [58], the number of events is dramatically increased by subsequent chemical reactions (hence the amplification), which means that less light is needed to complete the process. Acids released by the exposure radiation diffuse during the post-exposure bake step. These acids render surrounding polymer soluble in developer. A single acid molecule can catalyze many such “de-protection” reactions; hence, fewer photons or electrons are needed [59]. Acid diffusion is important not only to increase photoresist sensitivity and throughput, but also to limit line edge roughness due to shot noise statistics [60]. However, the acid diffusion length is itself a potential resolution limitation. In addition, too much diffusion reduces chemical contrast, leading again to more roughness [60].

21.3 Scanning Near-Field Optical Microscopy and Lithography

Scanning Near-field Optical Microscopy is nowadays an affirmed and tested scanning probe technique able to acquire information on the optical behavior of a surface with sub-diffraction resolution [61, 62]. In principle, it consists of illuminating a sample surface by evanescent components of the electromagnetic field coming up from a subwavelength sized probe [63, 64]. The probe is raster scanned over a sample surface at a few nanometers away from it (near-field region), and concurrently the optical image is then acquired. The reconstructed image is the result of modulation of the EM field emitted by the probe and the local optical properties of the sample. Actually, two main SNOM probe classes exist, i.e., with aperture and apertureless [65, 66]. In the first case, the probe is substantially formed by an optical sensor with a sub-wavelength aperture. To this group, belong the metal coated tapered fiber probes and the hollow-cantilevered probes. In the latter case, the SNOM probe consists of a sharp tip whose apex acts as a nanometre-sized light scatterer. Sensors having this characteristic are represented by STM metallic tips [67], or AFM dielectric and semiconductor tips [68], or nano-emitters, such as single molecules or nanocrystals [69]. Since the high-spectral content (in the k -space) of the near-fields is generated by the probe, this SNOM element is the key feature for super optical resolution. Pohl and coworkers in the 1984 have achieved a 25-nm resolution (corresponding to $\lambda/20$) in the visible range [70], whereas a sub-10 nm resolution has been reached by Wickramasinghe’s group [71]. A complete explanation of the super-resolution is attainable by solving Maxwell’s equations for the electromagnetic field at the SNOM probe apex. Due to the complicated boundary conditions imposed by the sensor geometry, only numerical solutions of the equations are available [72]. Nonetheless, an analytical solution can be achieved by approximating the SNOM aperture probe as a subwavelength circular hole in an infinite perfectly conducting

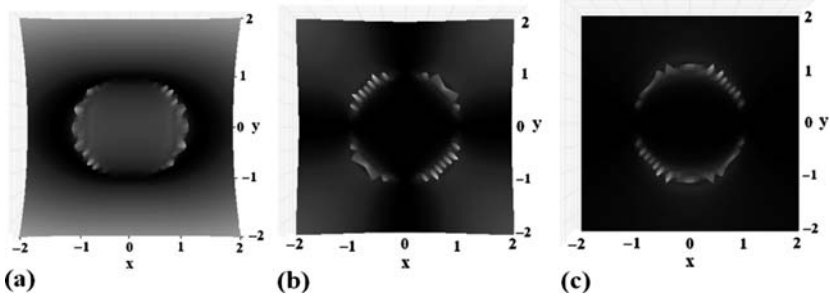


Fig. 21.3. Plot of the squared components of the electric field diffracted by a circular aperture in a screen of infinite perfect conductivity. The plot is performed at $z = 0$ nm into the aperture. (a) and (b) show the $|E_x|^2$ and $|E_y|^2$, respectively, while (c) shows $|H_z|^2$. The axes scales are in units of r

screen [73, 74]. In particular, for a metal screen (in the xy plane) containing an aperture of radius r irradiated by a plane wave of wavelength λ propagating along the z -direction and polarized along the x -direction, the following analytical expressions are obtained:

$$\begin{aligned} E_x &= -\frac{4i\omega E_0}{3\pi c} \cdot \frac{2r^2 - x^2 - 2y^2}{\sqrt{r^2 - x^2 - y^2}} \\ E_y &= -\frac{4i\omega E_0}{3\pi c} \cdot \frac{x \cdot y}{\sqrt{r^2 - x^2 - y^2}} \\ H_z &= -\frac{4H_0}{\pi} \cdot \frac{2r^2 - x^2 - 2y^2}{\sqrt{r^2 - x^2 - y^2}}, \end{aligned} \quad (21.1)$$

where $E_0 = |\mathbf{E}_0|$ and $H_0 = |\mathbf{H}_0|$ are the amplitudes of the electric and magnetic field respectively. Equation (21.1) holds for the electromagnetic field components at the aperture ($r \ll \lambda$, $z = 0$) (near-field components). The squared components $|E_x|^2$ and $|E_y|^2$ of the electric field [75, 76] are reported in Fig. 21.3 for an aperture radius of 50 nm and a wavelength of 514 nm.

The EM field is clearly enhanced near the sharp edge of the aperture, while the fields show divergences at the aperture rims, from which follows the sub-diffraction resolution [75].

Different possible SNOM configurations are available depending on how the near-field is generated and detected:

- Transmission mode: the sample is illuminated through the probe, and the light passing through the sample is collected and detected.
- Reflection mode: the sample is illuminated through the probe, and the light reflected from the sample surface is collected and detected.
- Collection mode: the sample is illuminated with a light source in the far-field and the probe is used to collect the light from the sample surface.
- Illumination/collection mode: the probe is used for both the illumination of the sample and for the collection of the reflected signal.

Contrast mechanisms such as absorption, polarization, and fluorescence contrast, well-known in conventional optical microscopy, can be transferred more

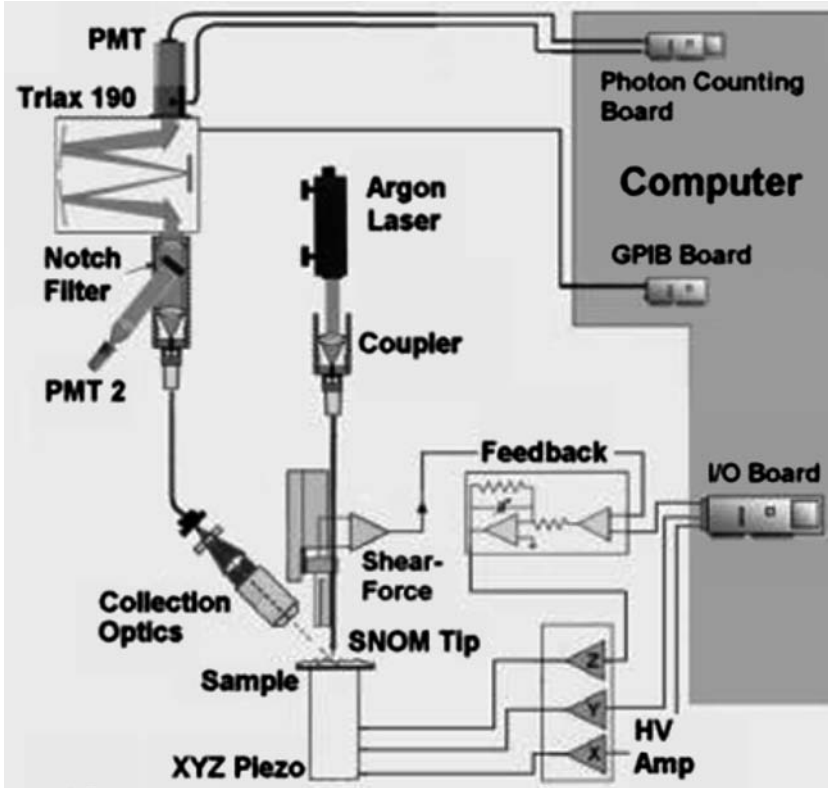


Fig. 21.4. Schematic of a SNOM setup working in illumination-mode

or less directly to SNOM applications. A typical SNOM setup operating in reflection and transmission mode, as well as for near-field spectroscopy is reported in Fig. 21.4. A light beam of suitable wavelength is coupled to the probe which is assembled onto a mechanical holder and roughly approached to the sample surface by a motorized microscrews stage. The specimen is placed on a $x - y - z$ piezo actuator that scans the sample under the tip during the experiment. The control of the probe–surface distance is performed by a tuning fork feedback system [77,78]. The tip attached to a tuning fork quartz is put in oscillation horizontally at the resonance frequency [79, 80]. As the tip approaches the surface, shear forces acting between the probe and the sample dampen the amplitude of the probe vibration [81]. Consequently, a piezoelectric potential between the two arms of the tuning fork is produced and used as a feedback signal for stabilizing the distance and for reconstructing the surface topography. Finally, the light scattered by the interaction of the probe with the sample is collected in the far-field, in reflection and/or in transmission, by means of a microscope lens and imaged on a photodetector (e.g., photomultiplier *PMT*, Avalanche Photodiode *APD*, etc.). As regard as

the near-field optical spectroscopy, the SNOM setup is further provided of a monochromator that analyzes the near-field light collected by the objective lenses.

Among the several applications that can be performed through the SNOM, lithography is the most naturally implemented. In fact, it is simply accomplished by scanning a SNOM optical fibre probe in the illumination mode over a photoresist film. SNOL carries out all of the synthetic flexibility and versatility of photochemical methods at a spatial resolution significantly beyond the limits of conventional photochemical methods. SNOL is also a useful tool to investigate the features of the near-field and to identify photo-induced physical and chemical modification processes at the sub-micrometer scale. For example, it allows to determining the effective aperture size of the tip used in the SNOL experiments [82]. The main drawback of the SNOL is its serial nature (the probe writes only one feature at a time). In principle, a parallel SNOL configuration can be achieved by illuminating a sample surface through a mask provided with a subwavelength transparent pattern placed at nanometer distance from the surface (Fig. 21.5). These conditions enable the sample to be sensitive to the nanometric variation of the field diffracted by the mask. In this case, the main difficulties are the control of the distance over all the mask-sample area and the inevitable presence of nanometric irregular relief on the photosensitive film that can generate defects on the SNOL process. A parallel mode SNOL setup [83] overcoming such troubleshooting needs further development in order to pattern macroscopically extended areas. The system called “SNOMipede” in principle should be made of an array of near-field probes (more specifically hollow-cantilevered probes) [83], similarly to the Millipede technology developed by IBM for parallel AFM lithography [84].

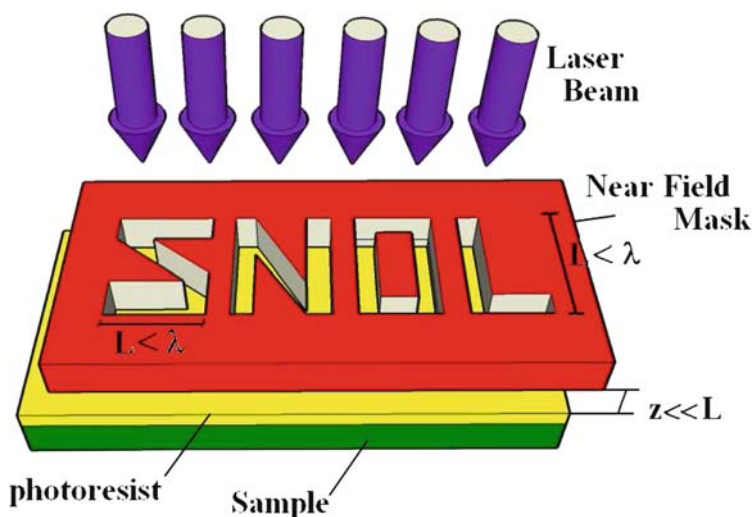


Fig. 21.5. Sketch of the parallel SNOL setup

21.3.1 Aperture and Apertureless SNOM Lithography

By considering the SNOL as a direct application of SNOM to lithography, two main classes of setups can be outlined, namely the aperture SNOL and the apertureless SNOL (Fig. 21.6). The most common setup uses as local illuminator a metal coated optical fibre probe (Figs. 21.6a and 21.7).

This kind of probe usually has a circular shaped aperture that produces a Gaussian-like distribution of electric energy for TE-modes (Fig. 21.8a). The distribution of the TM-modes intensity has the maximum on the aperture edges and strongly decreases toward the centre of the aperture (Fig. 21.8b) [85–90].

This power distribution assures an almost-isotropic illumination of the photosensitive surface and it is suitable to realize relief with a regular topographic profile.

Different SNOL experiments on organic and inorganic photosensitive materials have shown the capability of the technique to realize structure of

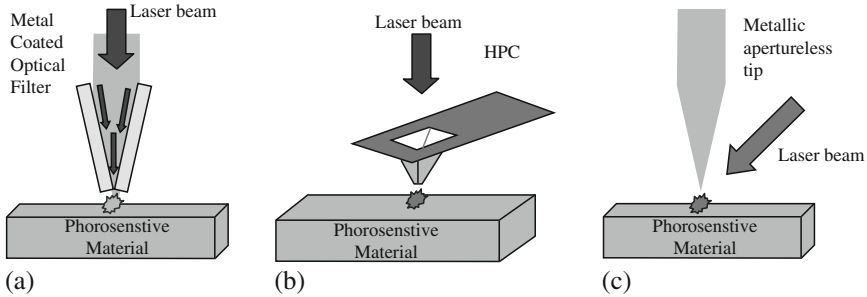


Fig. 21.6. Different SNOL setups using: (a) aperture probe in illumination mode; (b) hollow-cantilever probe in illumination mode; (c) apertureless metallic probe

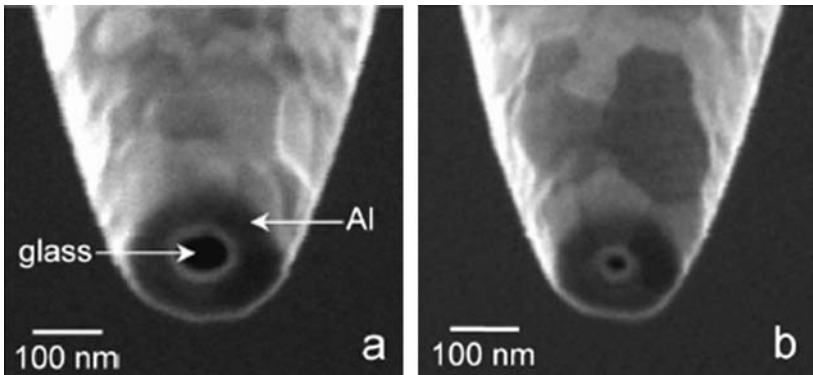


Fig. 21.7. SEM images of two SNOM probes in which the ends is milled flat by FIB. The aperture diameters are (a) 120 nm and (b) 35 nm. Reprinted from [90] copyright 1998, with permission from the American Institute of Physics

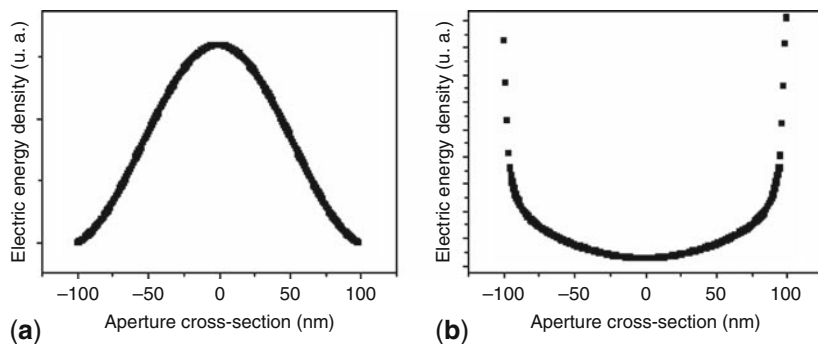


Fig. 21.8. Distribution of the electric energy density on the probe aperture interface for (a) TE and (b) TM polarization mode, respectively. Reprinted from [85] copyright 1998, with permission from the authors

nanometric size as well as a nonlinear dependence of the feature growing by the light power [91–93].

Betzig and co-workers have been pioneers in this field, showing the capability of SNOL as a tool for high-density data storage [94]. They have realized domains of magnetization 60-nm wide into a thin-film with magneto-optic properties (Co/Pt multilayer film). They used a visible light (488 and 514 nm) and heated the small region of the film under the probe aperture near to the Curie temperature (about 300 °C). No external magnetic field has been applied in those experiments. The efficiency of the writing process depends on the optical power that comes out from the probe and is used to heat the sample. Hence, the throughput of the probe is a limit. The power injected into the probe should not be too high in order to avoid the damaging of its aluminium coating.

SNOL experiments were also performed on conventional photoresist [82, 95]. The Krausch's group worked on phenol–formaldehyde resin. By coupling to the optical fiber probe an argon ion laser beam line at 454 nm, they have been able to write structures as small as 80 nm (spatial resolution of $\sim \lambda/5$) [82]. The photoresist reliefs were topographic replicas of the intensity distribution of the optical near-field coming out from the aperture probe and the features could be fitted with a Gaussian intensity distribution with approximately 100-nm width and 15-nm height. Finally, a grating was made with a period of 164 nm and a line width of 82 nm. These results confirmed that SNOM lithography beyond the diffraction limit is possible. However, the lateral dimensions of the produced structure were always greater than the diameter of the optical fibre. Subsequent studies have shown that the main problem was the thickness of the photoresist film, since the electric field associated with the aperture rapidly diverges in the dielectric layer beneath the probe [96]. Madsen's group used metal coated and uncoated fibres coupled to an Ar⁺ laser to write oxide structures into hydrogen-passivated silicon

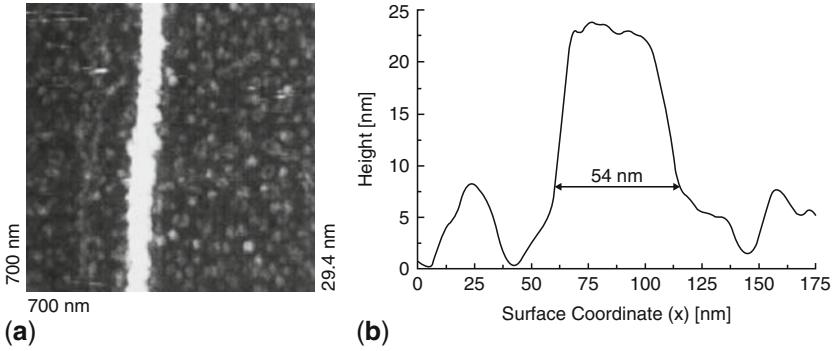


Fig. 21.9. AFM image after KOH etching of a single line written via sample exposure through an Al-coated fibre probe. The diameter of the probe aperture is 100 nm and the illumination wavelength is 457.9 nm. The power density applied to the sample was estimated to be 130 kW/cm^2 . **(b)** Cross-sectional profile of the single line. Reprinted from [98] copyright 1997, with permission from the American Institute of Physics

surfaces, which were then used as resists during etching of the Si substrate with KOH (Fig. 21.9) [97,98]. By using the metal coated fiber probe, they produced structures of 50-nm lateral width, whereas by uncoated fiber probe they realized structure with dimensions ranging from 111 to 126 nm [97].

SNOL processes with pulsed diode laser were performed by Hosaka and co-workers. They build structures of 60-nm lateral size on amorphous GeSbTe films by means of a local heating process performed by pulsed infrared near-field light delivered through the SNOM probe [99].

A good alternative to the optical fiber probe for aperture SNOL are represented by the so-called HPC probe [100,101]. It consists of a noncontact AFM probe whose tip has been nano-drilled to produce an aperture ranging from 40 to 150 nm (Fig. 21.10). If it is illuminated from the back, the aperture works as a confined light source. Additionally, the sensor is employed as a conventional force sensor for SFM. The probe has very good throughput, thanks to the large tapering angle and preserves a well-defined polarization state [101–103], and silicon dioxide as tip material guarantees excellent mechanical stability for long-life performance. The HPC integration in a conventional SNOM setup is straightforward. The stabilization of the tip–sample distance can be accomplished by techniques directly imported by the AFM world such as the contact and the tapping mode. This latter is usually preferred because it preserves the life of the cantilever and the sample integrity. Unfortunately, the shape of the cantilever leads to damping effects produced by the thin air cushion between the cantilever and the surface [104]. This problem has been sometimes overcome by increasing the angle between the sample surface and the cantilever or by more intricate methods, such as drilling a number of micro-holes in the cantilever body to permit air to escape [105]. Shear force operation mode

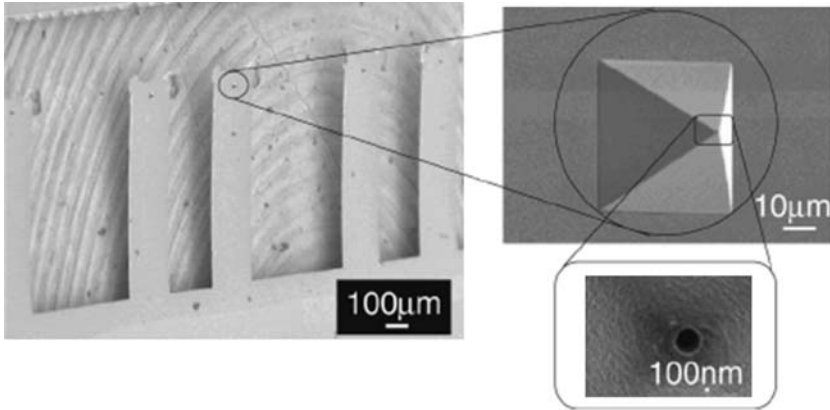


Fig. 21.10. SEM image of an HPC probe. Reprinted from [103] copyright 2004, with permission from the America Institute of Physics

has been demonstrated to be damping free and to guarantee long-term stability. A preliminary SNOL experiment with HPC working in shear force mode has been performed on azopolymeric films (PMA/PMA4) [78]. UV near-field light coming from the HPC aperture produced a local photo-isomerization on a 100-nm thick azopolymeric film with subsequent polymer embossing.

The special features of the HPC and the shear force mode of operation have allowed producing one after the other a set of dots imaging step by step their grown. Structures 10-nm high and < 50 nm (FWHM) large have been obtained by illuminating each dot-area for a few seconds with only 12 nW of optical power coming out from the aperture (Fig. 21.11).

A different approach for SNOL uses the so-called apertureless probes. It can be substantially schematized as a subwavelength photon-scatterer (sharp tip or nanoparticle), that, for the Babinet's reciprocity principle, behaves similarly to a small hole (its reciprocal). The apertureless approach allows confining a virtually unlimited quantity of electromagnetic radiation into regions as small as a few hundreds of squared nanometres. Standard STM metallic tips [67], AFM dielectric or semiconductor tips (Fig. 21.12) [68], and specially designed probes with a single nanostructure at their apex acting as a nanometre-sized light source or antenna [106], may be used.

The apertureless SNOM (ASNOM) configuration exploits the strong localization and enhancement of the fields generated at the tip apex when illuminated by a focused laser beam having the electric field vector of the polarized along the direction of the tip shaft (Fig. 21.13) [107, 108]. It has been shown both theoretically and experimentally that this local Field Effect (FE) actually corresponds to an electromagnetic singularity related to the high-charge density at the tip end. In electrostatics, a similar phenomenon is known as the "rod effect" [109–111]. By analogy, the term "*lightning rod effect*" is often used in near-field optics. The local field in close proximity to a laser irradiated

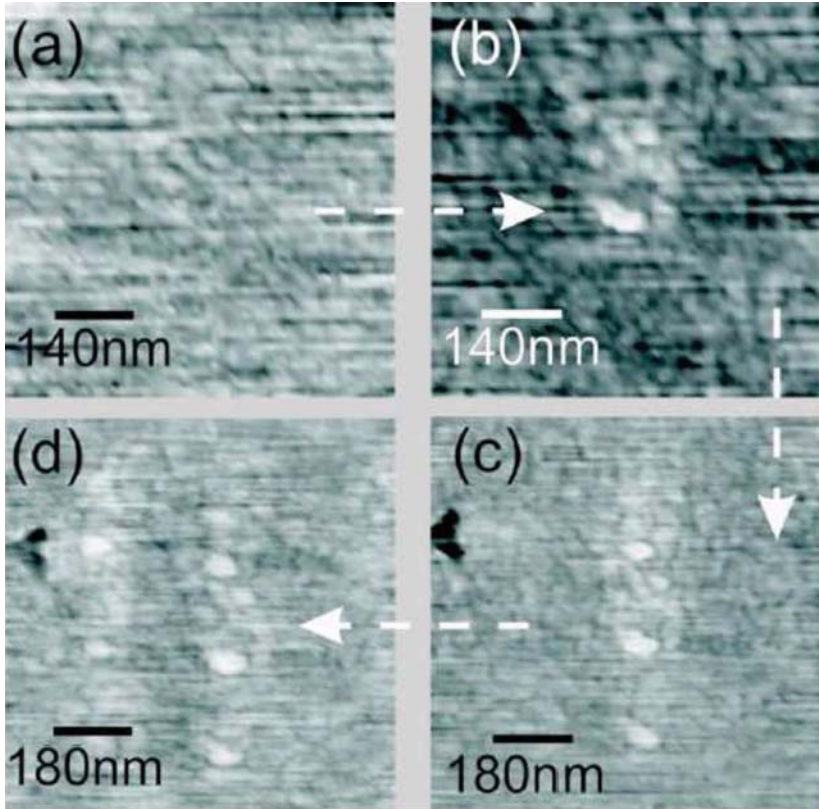


Fig. 21.11. Consecutive writing of five dots on the free surface of a PMA/PMA4 copolymer film. (a) SNOM topographical image of the area before writing. (b) Image of the first dot drawn. (c) Two dots added to the first one. (d) Result of the complete process. Because of the short (15–30 s) exposure time for each dot, the height is small (≤ 14 nm). Reprinted from [78] copyright 2006, with permission from America Institute of Physics

tip can be approximated by the electric field of an effective dipole $\mathbf{p}(\omega)$ located at the centre of the tip apex. The magnitude of the effective dipole can be written as [112]:

$$\mathbf{p}(\omega) = \begin{bmatrix} \alpha_{\perp} & 0 & 0 \\ 0 & \alpha_{\perp} & 0 \\ 0 & 0 & \alpha_{\parallel} \end{bmatrix} \cdot \mathbf{E}_0(\omega) \quad (21.2)$$

where the z -axis coincides with the tip axis and \mathbf{E}_0 is the exciting electric field in the absence of the tip. The transverse polarizability α_{\perp} is identical to the quasistatic polarizability of a small sphere and the longitudinal polarizability α_{\parallel} is given by $\alpha_{\parallel}(\omega) = 8\pi \cdot \epsilon_0 \cdot r_0^3 \cdot f_e(\omega)$; r_0 is the tip radius and f_e the complex field enhancement factor. The enhancement at the tip generates locally a light source that can be accurately represented by a single on-axis oscillating dipole

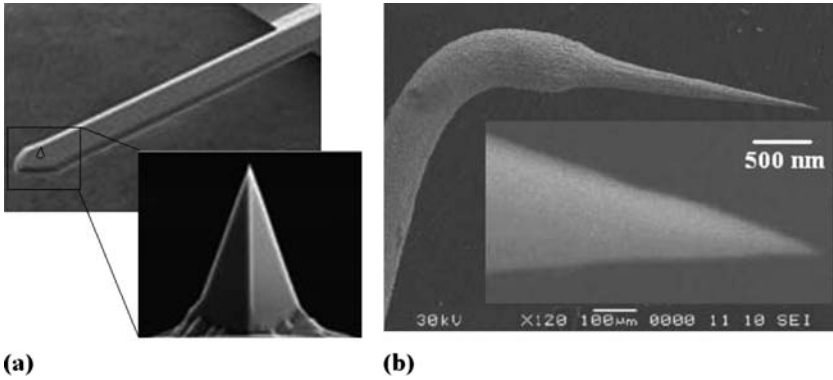


Fig. 21.12. SEM images of (a) an AFM silicon cantilever and (b) a bended STM tungsten tip (figure (a) is reprinted with permission from Nanosensors TM)

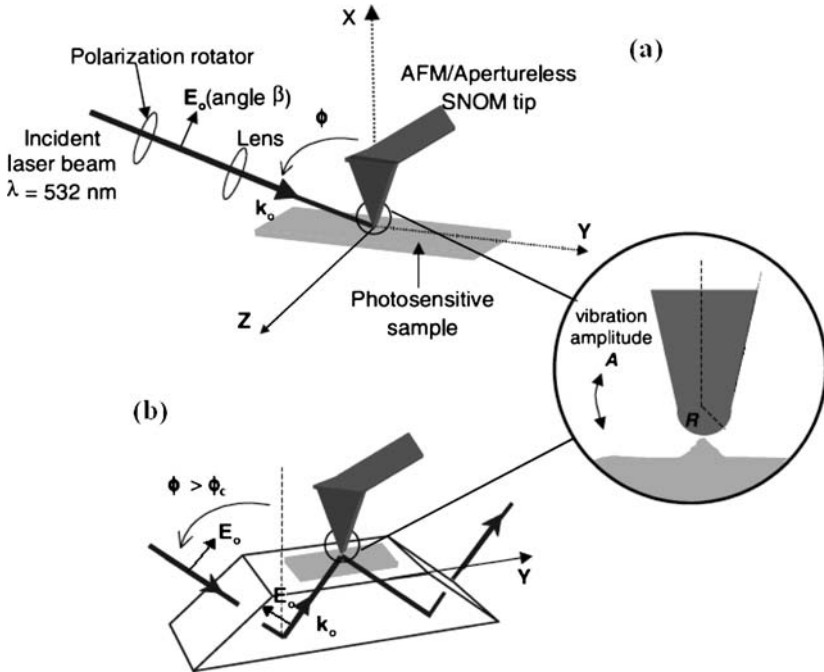


Fig. 21.13. Sketch of ASNOM configurations (a) reflection mode, (b) total internal reflection (TIR). Reprinted from [111] copyright 2003, with permission from Wiley Interscience

[112, 113]. The dipole strength is a direct measure for the field enhancement factor.

This effect has several advantages for writing the desired patterns on photosensitive materials. It does not depend on a particular SNOM-probe

geometry. Moreover, they can support high-energy illumination without suffering any damages at any wavelength also in the mid- and the far-infrared regions. In Fig. 21.13, two typical ASNOM configurations are reported.

SNOL experiments using apertureless probes have been demonstrated on commercial photoresist materials as well as on photosensitive azo-polymer [112, 114–116]. These experiments show that FE induces efficient local polymerization of the resist, while the tip scans the surface according to the pattern to be replicated. Setting the far-field exposure energy below the threshold energy dose (E_{th}) for the photoresist polymerization, the enhanced energy density at the tip apex surpasses the threshold only locally, inhibiting the polymerization process far from the tip. Thus, it is possible to control locally the total exposure energy by modifying the scanning speed in order to avoid an excessive exposure with a consequent increase of the patterns' width. Using platinum-coated AFM silicon tips, Kawata and co-workers [115] have obtained 100-nm FWHM grooves on a commercial-grade positive photoresist, a factor 3 better than in absence of the tip. SNOL experiments have also been carried out by exploiting multi-photons irradiation and the nonlinear optical properties of photosensitive materials [116, 117]. Yin et al. [116] have performed a multi-photon experiment with a red laser on a 60-nm thick spin-coated film of commercial positive photoresist (SU-8 resist precursor diluted with γ -butyrolactone), using AFM tips coated either with Pt/Ir or with Au installed on a reflection ASNOM. Since SU-8 does not absorb at 790 nm, the polymerization of the resist can occur only as a consequence of multi-photons absorption. Being a nonlinear effect, the two-photons-induced polymerization is thus preferred by the field enhancement at the tip apex. For far-field exposure energy below E_{th} , the enhancement provided by the tip allows photo-chemical processes and as consequence enables the polymerization of the resist, yielding to pattern lines with a FWHM of ~ 70 nm.

Apertureless SNOL is also a useful tool for basic studies of the physics underlying the field enhancement effects. For different illumination configurations, it provides snapshots of the electromagnetic field intensity distribution at the sample surface in the presence of apertureless tips made of different materials. Extensive experiments have been carried out by Bachelot et al. [111, 118] using azohybrid sol-gel and polymethylmethacrylate with dispersed Red 1 (PMMA-DR1) as photosensitive materials. The lithography process of azobenzene groups dispersed in polymer matrices involves a *trans-cis-trans* photo-isomerization reaction. During the near-field irradiation, the presence of a nonzero polarization component of the EM field along the intensity gradient, induces mass transport and a consequent deformation of the matrix embedding the molecules [119]. As a result, the field intensity distribution can be directly “read” by acquiring the surface topography (Fig. 21.14).

Moreover, the analysis of the height of the dot-like structures generated at the tip apex, gives information on the field enhancement dependence on the polarization of the excitation field [118]. By studying the dot height as function of the angle of the incident linear polarization of the laser light (Fig. 21.14a–c),

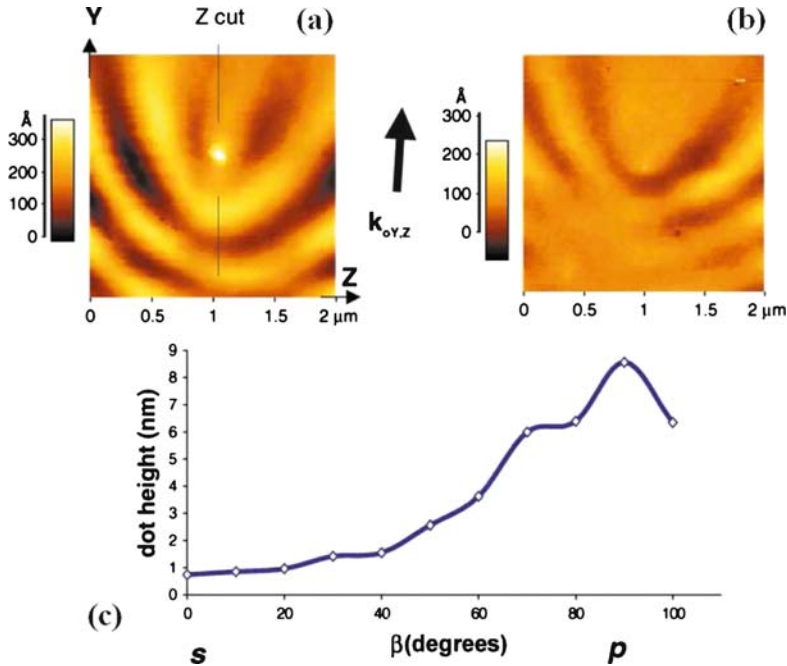


Fig. 21.14. $2 \times 2 \mu\text{m}$ topographic images obtained after reflection-mode illumination of a cobalt tip. (a) *p*-polarization, (b) *s*-polarization. (c) Dot height as a function of the angle of the incident linear polarization. Reprinted from [111] copyright 2003, with permission from Wiley Interscience

it was established that, the more intense the component of the electric field parallel to the tip axis is, the more intense is the field underneath the tip. However, the absence of a threshold energy dose on such materials, make them more similar to the far-field components generated by the tip.

Figure 21.15a,c,e shows how the far-field radiation reflected by the tip interferes with the incident field generating the typical fringe pattern [111].

In order to decrease the fringe pattern intensity, the total internal reflection excitation configuration (Fig. 21.13b) is exploitable. In this case, the theory predicts a much smaller contribution due to the far-field components with respect to the field enhancement due to the tip. This is displayed in the field intensity map and the relative line profile of Figs. 21.15b and d. The TIR mode configuration is thus the best suited for nano-writing experiments on such materials, as shown in [111], where well-separated dots with a FWHM of ~ 40 nm have been obtained (Fig. 21.16).

Recently, SNOL experiments were performed by using hybrid probes made by the integration of an apertureless and an aperture probe [120–122]. These special probes combine the advantages of the light confinement owned by the apertures probe with the field enhancement effect of the apertureless probe

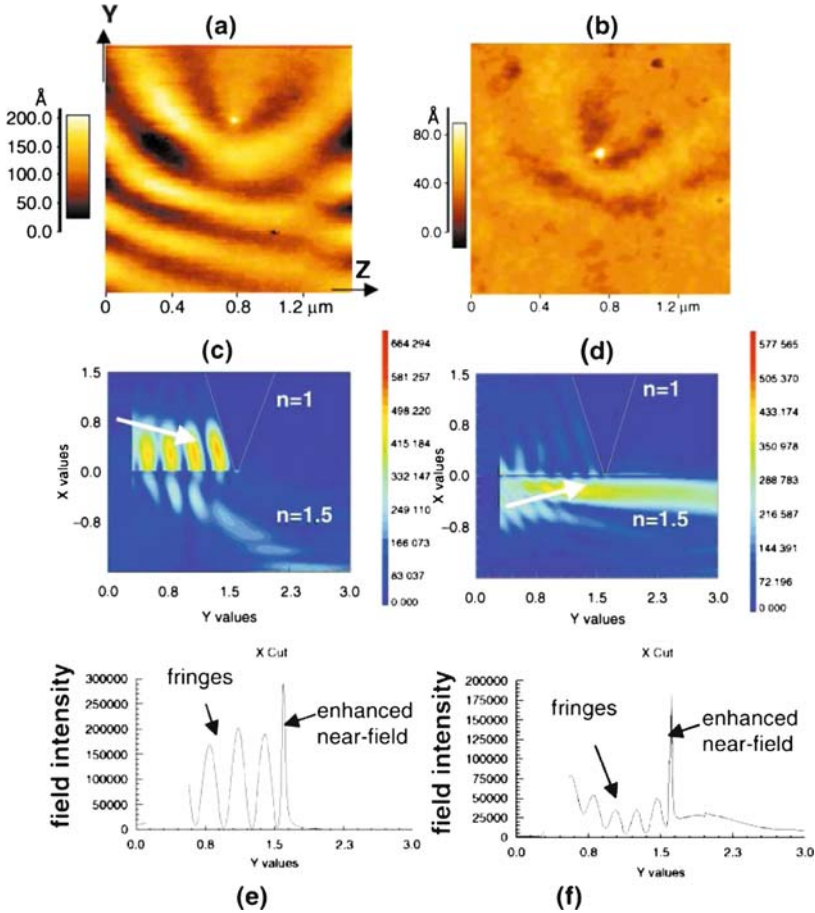


Fig. 21.15. Influence of the illumination geometry. (a, b) Experimental results, (c–f) Theoretical results. (a, b) Topographic images acquired after *p*-polarized illumination of a platinum tip on $1.5 \times 1.5 \mu\text{m}$ scan area. (a) Reflection mode illumination, (b) TIR mode illumination. (c, d) Calculated $3 \times 3 \mu\text{m}$ images of the electric field amplitude in the vicinity of a 20-nm radius perfectly conducting tip by 2D-FDTD (detail of calculation method are reported in [111]). The incident wave vector is represented by the white arrow. (c) Reflection mode illumination, (d) TIR illumination, (e) X cross-section, corresponding to (c), showing intensity distribution beneath the tip end, (f) X cross-section, corresponding to (d), showing intensity distribution beneath the tip end. Reprinted from [111] copyright 2003, with permission from Wiley Interscience

[120,121]. Figure 21.17 shows a hybrid probe made of a ridge antenna aperture in a probe integrated with a nanoscale aperture. The bowtie aperture probe is made by modifying a standard silicon nitride AFM probe. A layer of aluminum 120-nm thick is evaporated onto the tip side of the cantilever. FIB milling is

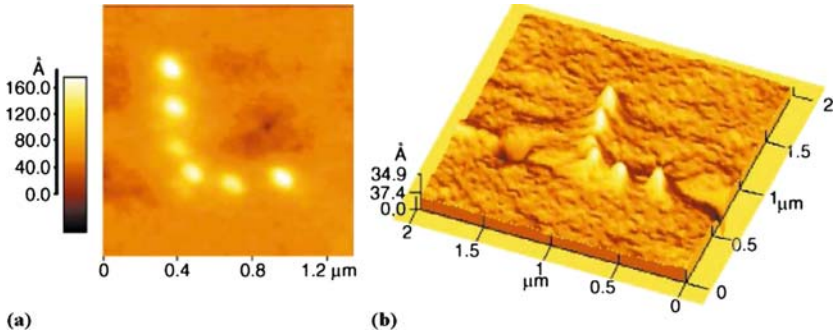


Fig. 21.16. (a) Topographic ASNOM images showing examples of pattern drawn on PMMA-DR1 film by local field enhancement at the end of a platinum tip illuminated in TIR mode in *p*-polarization. The dots have been written by a sequential exposure (duration: 5 s). (b) 3D representation. Reprinted from [111] copyright 2006, with permission from Wiley Interscience

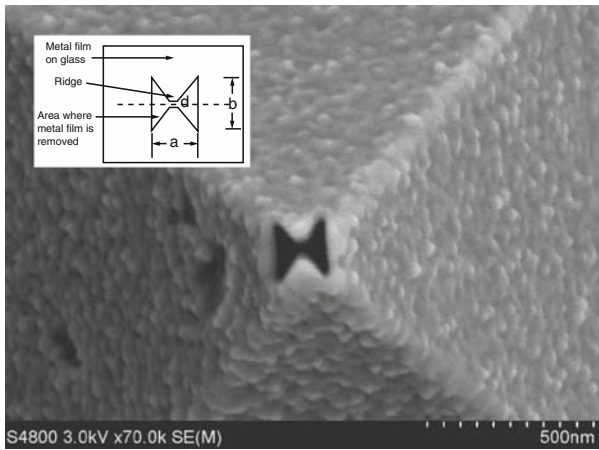


Fig. 21.17. SEM image of a bowtie aperture on AFM probe with inset of idealized geometry. Reprinted from [121] copyright 2008, with permission from the Optical Society of America

used to produce the aperture. The outline dimension of the aperture is about 160 nm.

The SNOL experiment was executed on a positive photoresist film (Shipley S1805) with sensitivity at 436 nm using the bowtie aperture probe for producing line patterns while scanning speeds and incident powers were varied [121]. In Fig. 21.18, the topographic images of the line pattern are reported. Lines with a lateral size smaller than the overall dimensions of the aperture (160 nm) are present, illustrating the light confinement properties of the ridge

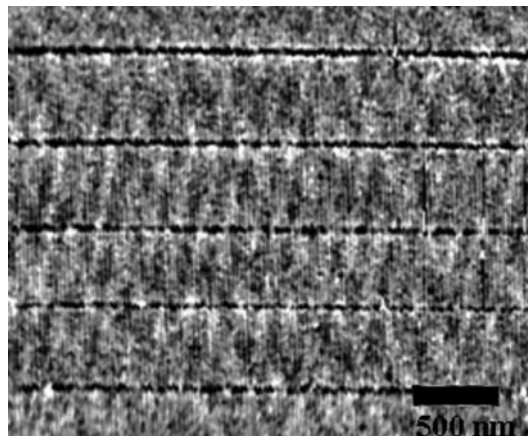


Fig. 21.18. Topographic image of bowtie lines with cross section view. Reprinted from [121] copyright 2008, with permission from the Optical Society of America

aperture. At high scan speed ($5\mu\text{m/s}$), the bowtie probe produced lines as small as 59 nm.

21.3.2 Near-Field Optical Lithography Achievements on Azo – Polymers

In this section, we resume some SNOM applications to nanowriting on azo-polymers [123–126].

Devices for pure optical and reversible writing and reading of sub-wavelength size bits are highly desirable, because of the potential enhanced speed and duration, combined to the increased density [127]. Various aspects embracing the optical writing resolution assessments and subsequent topographical and optical reading have been addressed. Effects of SNOM probe heating and real-time near-field analysis of the writing process were also performed [124, 128]. The experiments were concentrated on Azobenzene-functionalized polymers with a special focus on azo-polymethacrylate polymers. These polymers attract particular interest for their high-optical activity producing erasable birefringence gratings [129, 130] as well as for the intriguing photo-induced surface relief phenomenon, which can be readily induced in a single-step process upon exposure of thin films to an interference pattern of polarized light at an absorbing wavelength [131]. The optical patterning mechanism in azobenzene side-chain polymers is based on the isomerization cycles of the azobenzene side-chains between their *trans* and *cis* forms, with respective high- and low-dipole moments [132, 133]. Polarized light activates *trans-cis-trans* isomerization cycles, leading statistically to a realignment of the azo-group perpendicular to the polarization direction, thus locally inducing birefringence. Photo-isomerization and the interaction with the

optical electric field are essential for the surface relief formation and the concurrent mass migration that shows a marked polarization dependence [134–138].

The early SNOL experiments on azo-polymers have been exploited to produce subwavelength-sized topographic features relying on the photo-induced surface deformation [123, 139, 140]. In particular, the photosensitive chain of azobenzene-groups, inside a polymethacrylate matrix (PMA4), was locally illuminated at 325 nm (HeCd laser line) or at 488 nm (Ar^+ laser line) through the aperture of a metal-coated tapered optical fibre. The 325-nm line excites the $\pi - \pi^*$ *trans-cis* isomer transition and the 488-nm line excites the $n - \pi^*$ *cis-trans* isomer transition.

The 488-nm laser light is responsible for the optical cycling of the *trans-cis-trans* isomerization since the UV peaked *trans* isomer absorption has a long tail in the blue. For laser energy densities above threshold [136, 141], photo-induced isomerization is accompanied by diffusion processes resulting in topological reliefs. Figure 21.19 shows a topographic images of a $3.5 \times 3.5 \mu\text{m}$ area where an array of 21 single dots have been written by SNOL process (each dots is located 500 nm one from the other). The near-field illumination at 325 nm was active for 1 s per dot while moving along a defined path the optical tip driven by the SNOM piezoelectric scanner. The laser power was about 2 mW at the laser-to-fibre coupler. The printing process was repeated until the drawing was completed. A line profile (marked with a black line in Fig. 21.19b and analyzed along the *x*-direction), reveals that the optical printing creates a number of well-separated embossed conical structures with a full width at half height of about 180 nm and a mean height of about 15 nm. The information was stable not only during the measurement time but also in the long term. A similar topographic effect was obtained by illuminating the

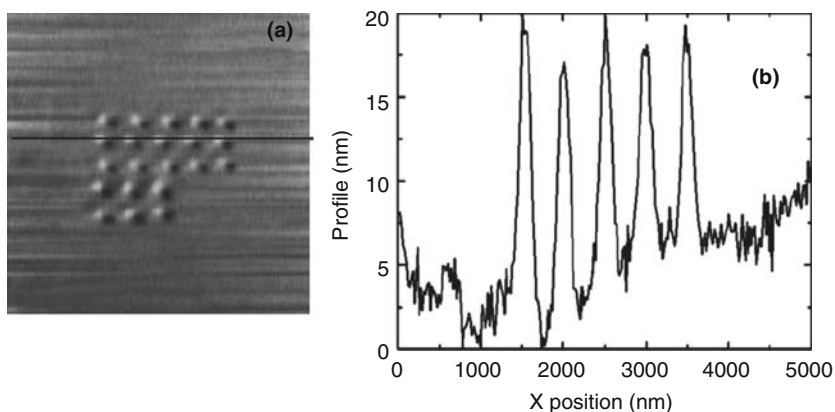


Fig. 21.19. (a) $5 \times 5 \mu\text{m}$ shear force image of the pattern written on a PMA4 film by illuminating the sample for 1 s at 325 nm. (b) Profile analysis along the black line marked in the image (a). Reprinted from [123] copyright 2002, with permission from Elsevier

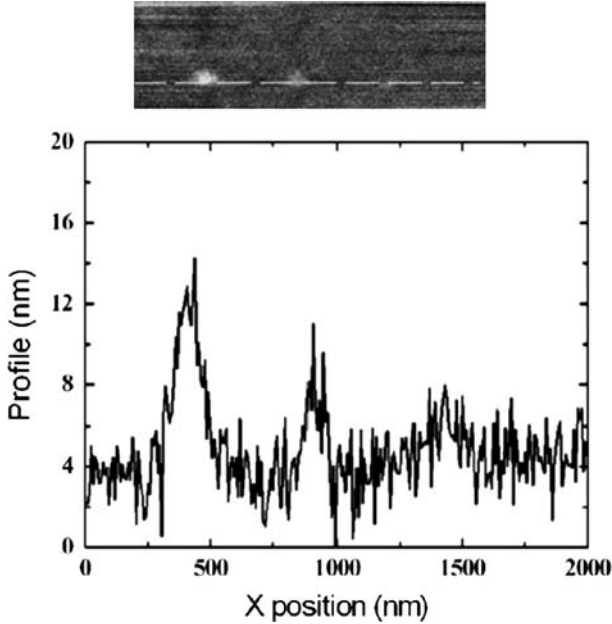


Fig. 21.20. Three dots obtained with different exposure times upon Ar^+ laser illumination at 488 nm. The embossing grows linearly with the exposure time Reprinted from [123] copyright 2002, with permission from Elsevier

sample surface with the Ar^+ laser line at 488 nm. The laser power coupled to the SNOM fibre was 0.5 mW. The printing process was repeated three times, with 15, 30, and 45 s of exposure, to produce three dots placed along a line 500 nm one from the other. Figure 21.20 shows the shear force image taken after this writing process and one topographic profile along the row marked with the dashed line in the left of the picture. Despite the longer recording time, the dimensions of the dots are unchanged and the growth of the structure is linear with respect to the illumination dose, at least in the investigated range.

Figure 21.21 shows a shear force image of a more complex, continuous line draw obtained by exposing a PMA4 film to the unshuttered 325-nm laser light. SNOM probe is driven along the sample surface, following a pre-defined path, at a speed $v = 22 \text{ nm/s}$. The line profile analysis (Fig. 21.21b) shows that the structures have a FWHM of about 75 nm and a mean height of about 10 nm.

Because of the importance the azo-polymers for sub-diffraction optical lithography, further investigation of the formation mechanisms of the photo-induced topographical structures have been performed [124]. Several SNOM experiments have confirmed the existence of two power regimes even for near-field illumination [139, 142]. When the light power density is low (tens of W/cm^2), the structures locally designed by near-field illumination are

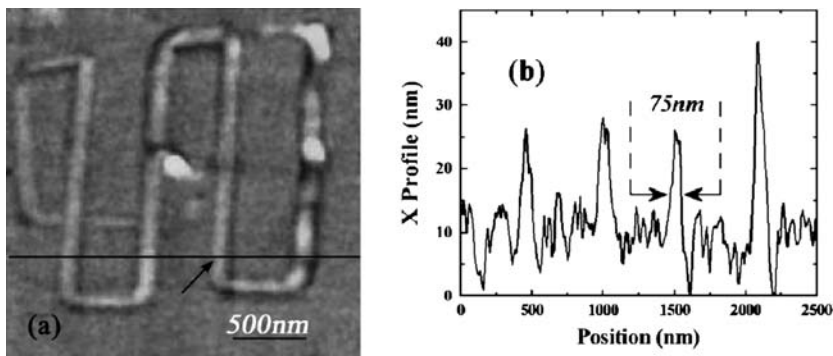


Fig. 21.21. (a) Shear force image of the optically nanostructured PMA4 surface ($\lambda = 325$ nm, $T = 297$ K, $\nu = 22$ nm/s) and (b) line profile extracted along the line marked in (a). Reprinted from [125] copyright 2003, with permission from Taylor & Francis

trenches, and when the power density is high (hundreds of W cm^{-2}), the structures are protrusions. Basing on these results, SNOM was used to monitor in real time the formation of single topographical protrusions (dots) on the free surface of azobenzene-based copolymers [124]. Working into the high-power density regime the dynamics of the mass migration during the embossing of the surface of 50/50 copolymer films of PMA and PMA with 3-methyl-4-pentyloxy-azobenzene units was followed. The experiments consisted in the illumination of the polymer surface by a SNOM optical fiber probe and observing the dynamic of the resulting surface relief formation. These observations have been performed also as a function of the sample temperature. Figure 21.22a shows the 3d-view dot grown on the surface of the PMA/PMA4 copolymer film after a near-field illumination and in Fig. 21.22b the experimental plots of the dots-height versus irradiation time are reported.

The temporal trend of the dot growth was analysed and fitted with a model developed by Barret et al. [143]. It is based on the idea that the mass migration leading to the formation of dots can be modelled as the motion of a Newtonian fluid in the classical laminar flow regime. Starting from the Navier–Stokes equations, the best fit of the experimental data is obtained by using the expression:

$$h(t) = d - \frac{d}{\sqrt{1 + \frac{2}{3}\beta d^2 t}} \quad (21.3)$$

where h is the dot height measured from the sample surface ($z = 0$, Fig. 21.6), β a constant that can be considered dependent on the value of the sample viscosity and the spatial derivative of the force that drives the mass displacement, d the asymptotic height of the grown dot, and t is the exposure time [143]. As evidenced in Fig. 21.22b, an excellent agreement between the experimental data and the fitting formula was found for all the temperature considered.

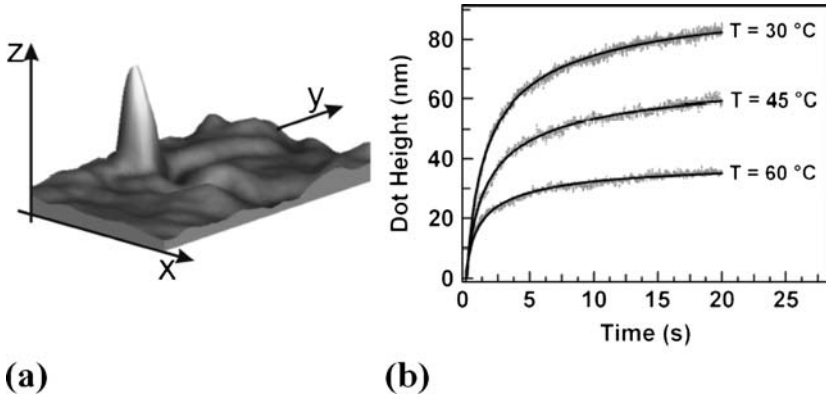


Fig. 21.22. (a) 3D view of a typical dot grown on the surface of the PMA/PMA4 copolymer film. The dot height is about 65 nm and the diameter is about 200 nm. (b) Plots of the experimental data (*grey*) and the derived fitting functions (*black solid lines*) described by (7.3). Reprinted from [124] copyright 2008, with permission from Wiley Interscience

Further experiments of all-optical writing and reading of features on the 100-nm scale on a thin film of an azobenzene side-chain polymer were realized by means of polarization-modulation SNOL. The experiments were performed on a block copolymer of PMA/PMA4 (30/70) deposited as 100-nm thin film on a clean glass substrate by spin coating. The copolymer has been introduced because of the higher photosensitivity with respect to the homopolymer. To perform pure optical writing and reading, a SNOM apparatus appropriate to detect the local optical anisotropy (dichroism or birefringence) of the samples was specifically developed. It has been accomplished by implementing a polarization-modulation setup described in [144, 145]. In practice, a beam of a 690-nm laser diode is prepared in a linear polarization state rotating at an angular frequency ϖ , before being coupled to the NSOM optical fiber (Fig. 21.23). The near-field light coming from the probe-sample interaction is collected in transmission by an aspheric lens and then detected by a miniaturized photomultiplier tube of extended red sensitivity, after crossing a linear polarizer. In order to improve the sensitivity of the optical channel, an intensity modulation of the laser diode beam at relatively high frequency (~ 25 kHz) has been added before the usual polarization modulation detection stage operating at $\varpi/2\pi = 125$ Hz. With synchronous detection, this modification yield imaging of optical anisotropy features produced with aperture diameters of ~ 100 nm. The blue line of an Ar⁺ laser at 488 nm, controlled in power by an acoust-optic modulator, was used for optical writing on the polymer surface at ambient conditions. Figure 21.24 shows birefringence images of optical patterns written by tracing vertical lines with a SNOM tip 150-nm large, delivering about 5 nW. The tracing speed was 1 $\mu\text{m/s}$ corresponding to a writing fluence of about 5 J/cm² [126].

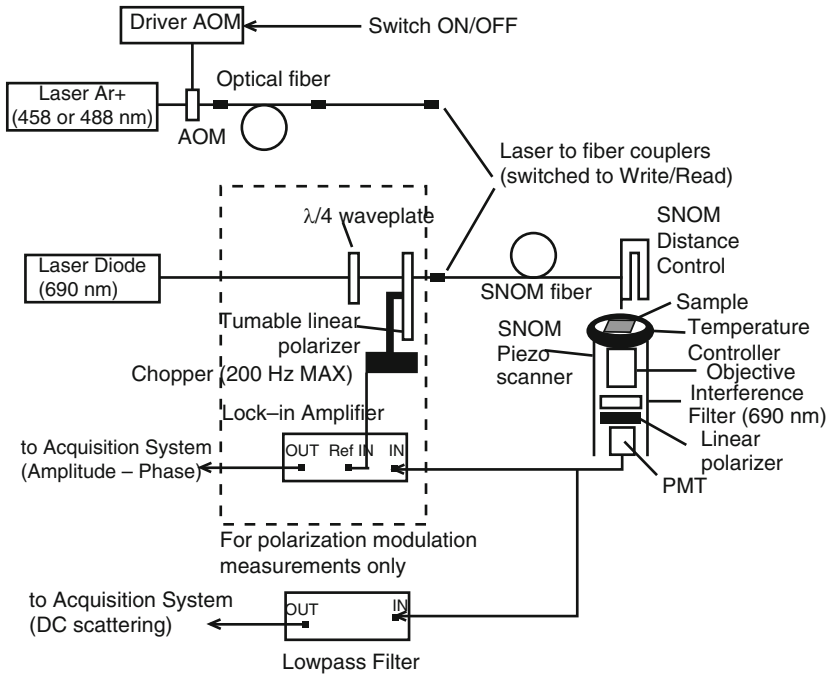


Fig. 21.23. Setup for polarization-modulation SNOM and optical writing experiments. Reprinted from [125] copyright 2003, with permission from Taylor & Francis

Repeated experiments with larger aperture tips and different writing speeds provided direct evidence that the values of light fluence necessary to obtain mass migration are considerably higher than the ones sufficient for changing the local optical properties of the polymer [125].

Figure 21.25 shows a narrow vertical line obtained with a speed of 30 nm/s corresponding to a fluence of about 150 J/cm^2 . The imaging quality obtained in this case is limited by the intrinsic resolution of aperture SNOM due to the tip diameter (estimated to be $\sim 100 \text{ nm}$ in this case), so that the thin line appears with poor contrast and irregular cross section. The full width at half-maximum measured on the profile obtained by averaging all the horizontal scan lines is about 200 nm, while single line profiles show variable thickness from 100 to 200 nm. For example, the one reported in Fig. 21.25b has a width of 120 nm. The higher fluence used to produce the narrow line was not yet sufficient to generate any topographic embossing.

Finally, the optical erasure process was investigated in order to demonstrate the potentiality of the azo-polymethacrylate films as a high-density optical re-writable medium. Since the provided polarization is not fixed the effect is to induce disorder in the orientation of the previously ordered chromophores, and then to erase the optical lines. Thus, a SNOM probe can

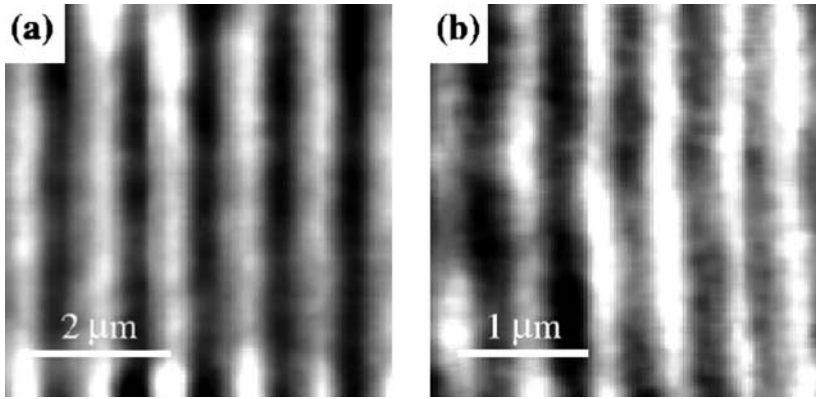


Fig. 21.24. Birefringence SNOM images ($\lambda = 690$ nm) of optical gratings realized by previously tracing parallel vertical lines with the SNOM probe ($\lambda = 488$ nm) on a thin film of PMA4. The lines are spaced by (a) 1,000 nm and (b) 500 nm. Reprinted from [126] copyright 2003, with permission from the American Institute of Physics

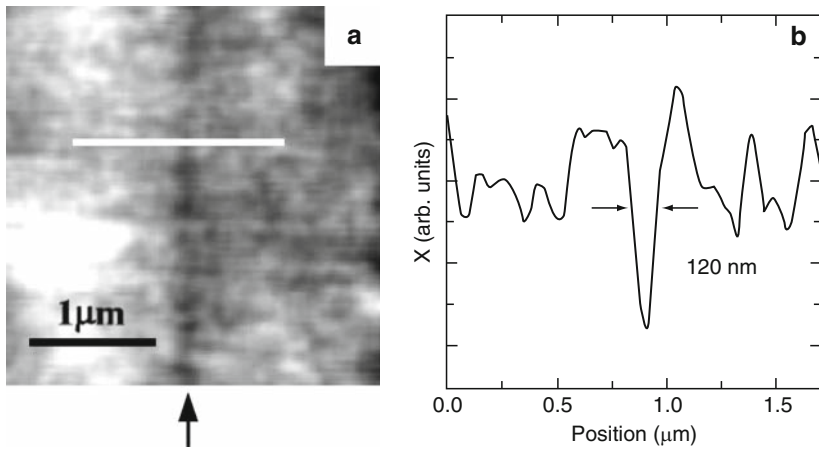


Fig. 21.25. Birefringence SNOM image ($\lambda = 690$ nm) of a line produced by previous writing with the SNOM probe ($\lambda = 488$ nm) on a thin film of PMA4. The arrow indicates the location of the line. (b) Line profile corresponding to the marked line in (a). Reprinted from [126] copyright 2003, with permission from the American Institute of Physics

induce optical disorder of the oriented chromophores. The optical orientation persists only where topographic reliefs are present, indicating that the main chain rearrangement influences the average optical orientation of the chromophores so that it is more difficult to erase such areas. This experiment demonstrated the possibility of local all-optical reversible writing/reading and erasure of optical information [125]. In conclusion, azo-polymethacrylate thin

films have been inscribed and read back optically with a resulting linewidth of the order of 100 nm using polarization-modulation SNOM. Employment of SNOM probes with smaller aperture and higher throughput will help to assess the ultimate limit for all-optical nanowriting in such materials toward possible applications for ultra-high density data storage.

21.4 Conclusions

The results reported in this chapter evidence that SNOL is a good alternative to other lithographic techniques. It reveals high capability in terms of resolution and usefulness at any wavelength and on any photoresist. Structures with lateral dimensions down to about 50 nm have been realized by both aperture and apertureless setups. Further resolution improvements are expected from the development of new probes with higher optical efficiency. Analogously, full development of a parallel setup (SNOMipede) will solve the problem of speed performance. Experiments performed on azo-polymers have shown that SNOL is also useful for data storage. Indeed, extensive near-field optical studies on the photo-isomerization process of such polymers have established the existence of a threshold effect. The dynamic of the relief growing has been followed in analogy to the motion of a Newtonian fluid in the classical laminar flow regime. It has been clearly shown that the illumination of the azo-polymer (optical writing) produces topographical reliefs at high-optical power and pure birefringence effects at low optical power. The first case implies topographical reading of stored data while the second case implies optical data reading.

References

1. F.R. Pease, S.Y. Chou, Proc. IEEE 96(2), 248 (2008)
2. J. Jahns, Q. Cao, S. Sinzinger, Laser Photon Rev., 1 (2008)
3. G.M. Wallraff, W.D. Hinsberg, Chem. Rev. 99, 1801 (1999)
4. G. Moore, Electronics 38(8), 114 (1965)
5. P. Sandoz, R. Giust, G. Tribillon, Opt. Comm. 161(4), 197 (1999)
6. The American Heritage[®] Dictionary of the English Language, Fourth Edition (Houghton Mifflin Company, Boston MA, 2004)
7. W.M. Moreau, Semiconductor Lithography: Principles and Materials (Plenum, New York, 1988)
8. D. Brambley, B. Martin, P.D. Prewett, Adv. Mater. Opt. Electron 4, 55 (1994)
9. P. Choudhury (ed.) Handbook of Microlithography, Micromachining and Microfabrication (SPIE Optical Engineering Press, Bellingham, WA, 1997)
10. E.P. Herzog, Micro-Optics (Taylor & Francis, London, 1997)
11. B.D. Terris, H.J. Mamin, M.E. Best, J.A. Logan, D. Rugar, S.A. Rishton Appl. Phys. Lett. 69, 4262 (1996)
12. M. Li, J. Wang, Zhuang L, Chou, Appl. Phys. Lett. 76, 673 (2000)
13. B. Faircloth, H. Rohrs, R. Tiberio, R. Ruoff, R.R. Krchnavek, J. Vac. Sci. Technol. B 18, 1866 (2000)

14. H. Schulz, H.C. Scheer, T. Hoffmann, C.M. Sotomayor Torres, K. Pfeiffer, G. Bleidiessel, G. Grützner, Ch. Cardinaud, F. Gaboriau, M.C. Peignon, J. Ahopelto, B. Heidari, *J. Vac. Sci. Technol. B* 18, 1861 (2000)
15. Y. Xia, J.A. Rogers, K.E. Paul, G.M. Whitesides, *Chem. Rev.* 99, 1823 (1999)
16. J.A. Rogers, R.G. Nuzzo, *Mater. Today* 8, 50 (2005)
17. J.B. Waldner, *Nanocomputers and Swarm Intelligence* (ISTE Wiley, London, 2008)
18. M. Rothschild, D.J. Ehrlich, *J. Vac. Sci. Technol. B* 5, 389 (1987)
19. W. Hardin *Photonics Spectra*, 94 (1997)
20. M. Hibbs, R. Kunz, M. Rothschild, *Solid State Technol.*, 69 (1995)
21. D.L. White, J.E. Bjorkholm, J. Bokor, L. Eichner, R.R. Freeman, T.E. Jewell, W.M. Mansfield, A.A. MacDowell, L.H. Szeto, D.W. Taylor, D.M. Tennant, W.K. Waskiewicz, D.L. Windt, O.R. Wood, *Solid State Technol.*, 37 (1991)
22. P.N. Dunn, *Solid State Technol.* 57, 49 (1994)
23. H.I. Smith, *J. Vac. Sci. Technol. B* 13, 2323 (1995)
24. Y. Xia, G.M. Whitesides, *Angew Chem. Int. Ed. Engl.* 37, 550 (1998)
25. P. Choudhury, *Handbook of Microlithography, Micromachining and Microfabrication* Ed (SPIE Optical Engineering Press, Bellingham, WA, 1997)
26. S. Owa, H. Nagasaka, K. Nakano, Y. Ohmura, *Proc. SPIE* 6154(61), 5408–5411 (2006)
27. H. Sewell, J. Mulken, D. McCafferty, L. Markoya, B. Streefkerk, P. Graeupner, *Proc. SPIE* 52, 6154 (2006)
28. G. Whitesides, *Chem. Rev.* 99, 1823 (1999)
29. M.D. Levenson, N.S. Viswanathan, R. Simpson, *IEEE Trans. Electron Dev.* 29, 1828 (1982)
30. A. Kwok-Kit Wong, *Proceedings of the SPIE Tutorial Texts in Optical Engineering* Bellingham (SPIE, WA, 2001)
31. P. Kruit, S. Steenbrink, M. Wieland, *J. Vac. Sci. Technol. B* 24, 2931 (2006)
32. J.P.H. Benschop, A.J.J. van Dijsseldonk, W.M. Kaiser, D.C. Ockwell, *J. Vac. Sci. Technol. B* 17, 2978 (1999)
33. Y. Li, F. Zhang, *Jpn. J. Appl. Phys.* 46, 2936 (2007)
34. M. McCord, M. Rooks, in *Handbook of Microlithography, Micromachining and Microfabrication*, vol 1, Chapter 2, ed. by P. Rai-Choudhury (SPIE, Bellingham, WA, 1997)
35. J.M. Gibson, *Phys. Today*, 56 (1997)
36. R.F.W. Pease, *J. Vac. Sci. Technol. B* 10, 278 (1992)
37. R.G. Jones, P.C.M. Tate, *Adv. Mater. Opt. Electron* 4, 139 (1994)
38. A.V. Crewe, J. Wall, *J. Mol. Biol.* 46, 375 (1970)
39. A.N. Broers, A.C.F. Hoole, J.M. Ryan, *Microelectron Eng.* 32 (1996)
40. R.L. Kubena, J.W. Ward, F.P. Stratton, R.J. Joyce, G.M. Atkinson, *J. Vac. Sci. Technol. B* 9, 3079 (1991)
41. P. Choudhury, *Handbook of Microlithography, Micromachining and Microfabrication* (SPIE Optical Engineering Press, Bellingham, WA, 1997)
42. H.I. Smith, H.G. Craighead, *Phys. Today*, 24 (1990)
43. D.A. Muller, Y. Tzou, R. Raj, J. Silcox, *Nature* 366, 725 (1993)
44. P.E. Batson, *Nature* 366, 727 (1993)
45. A. Broers, in *Materials for Microlithography*; ACS Symposium Series 266, ed. by L. Thompson, F.G. Willson (American Chemical Society, Washington, DC, 1984)

46. S. Matsui, Y. Ochiai, *Nanotechnology* 7, 247 (1996)
47. R. Seliger, R. Kubena, R. Olney, J. Ward, V. Wang, *J. Vac. Sci. Technol.* 16, 1610 (1979)
48. J. Bartelt, C. Slayman, J. Wood, J. Chen, C. McKenna, C. Minning, J. Coakley, R. Holman, C. Perrygo, *J. Vac. Sci. Technol.* 19, 1166 (1981)
49. G. Gross, *J. Vac. Sci. Technol. B* 15: 2136 (1997)
50. G. Stengl, H. Loschner, W. Maurer, P. Wolf, *J. Vac. Sci. Technol. B* 4: 194 (1986)
51. W. Brunger, L. Buchmann, M. Torkler, W. Finkelstein, *J. Vac. Sci. Technol. B* 12: 3547 (1994)
52. A. Benninghoven, *Angew. Chem. Int. Ed. Engl.* 33, 1023 (1994)
53. H.J. Levinson, *Principles of Lithography*, vol PM146, 2nd ed. (Bellingham, WA, 2001)
54. J. Liu, B. Cai, J. Zhu, G. Ding, X. Zhao, C. Yang, D. Chen, *Microsyst. Technol.* 10(4), 265 (2004)
55. H. Yang et al., *Proceedings of the 1st IEEE International Conference on Nano/Micro Engineered and Molecular Systems*, 391 (2006)
56. C. Dupas, P. Houdy, M. Lahmani (eds.) *Nanoscience Nanotechnologies and Nanophysics* (Springer-Verlag, Berlin Heidelberg, 2006)
57. M. Hashimoto, S. Watanuki, N. Koshida, M. Komuro, N. Atoda, *Jpn. J. Appl. Phys.* 35, 3665 (1996)
58. O. Nalamasu, M. Cheng, J.M. Kometani, S. Vaidya, E. Reichmanis, L.F. Thompson, *Proc. SPIE* 32, 1262 (1990)
59. US Patent #4,491,628 – J.M.J. Fréchet, H. Ito, C.G. Willson, Positive and Negative Working Resist Compositions with Acid-Generating Photoinitiator and Polymer with Acid-Labile Groups Pendant From Polymer Backbone
60. D. Van Steenwinckel, J.H. Lammers, T. Koehler, R.L. Brainard, P. Trefonas, *J. Vac. Sci. Tech. B* 24: 316 (2006)
61. L. Novotny, in *Progress in Optics* 50, ed. by E. Wolf (Elsevier, Amsterdam, 2007), p. 137
62. G. Kaupp, *Int. J. Photoenergy*, vol 2006, 1, (2006)
63. D.W. Pohl, *Phil. Trans. R Soc. Lond. A* 362, 701 (2004)
64. R.C. Dunn, *Chem. Rev.* 99, 2891 (1999)
65. E. Cefali, S. Patanè, S. Spadaro, R. Gardelli, M. Albani, M. Allegrini, Chap. 3, 77–135 in *Applied Scanning Probe Methods*, vol VIII, ed. by B. Bhushan, H. Fuchs, M. Tomitori (Springer, Berlin Heidelberg New York, 2008)
66. L. Novotny, B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, Cambridge, UK, 2006)
67. G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Phys. Rev. Lett.* 49, 57 (1982)
68. G. Binnig, C.F. Quate, C. Gerber, *Phys. Rev. Lett.* 56, 930 (1986)
69. J. Michaelis, C. Hettich, A. Zayats, B. Eiermann, J. Mlynek, V. Sandoghdar, *Opt. Lett.* 24, 581 (1999)
70. D.W. Pohl, W. Denk, M. Lanz, *Appl. Phys. Lett.* 44, 651 (1984)
71. F. Zenhausern, Y. Martin, H. Wickramasinghe, *Science* 269, 1083 (1995)
72. C. Girard, *Rep. Prog. Phys.* 68, 1883 (2005)
73. H. Bethe, *Phys. Rev.* 66, 163 (1944)
74. C. Bouwkamp, *Philips Res. Rep.* 5, 321 (1950)
75. H.D. Hallen, C.L. Jahncke, *J. Raman Spectrosc.* 34, 655 (2003)
76. H.D. Hallen, E.J. Ayars, C.L. Jahncke, *J. Microsc.* 210, 252 (2003)

77. K. Karrai, R.D. Grober, *Ultramicroscopy* 61, 197 (1995)
78. A. Ambrosio, E. Cefali, S. Spadaro, S. Patanè, M. Allegrini, D. Albert, E. Oesterschulze, *Appl. Phys. Lett.* 89, 163108 (2006)
79. R. Toledo-Crow, P. Yang, Y. Chen, M. Vaez-Iravani, *Appl. Phys. Lett.* 60, 2957 (1992)
80. E. Betzig, P. Finn, S. Weiner, *Appl. Phys. Lett.* 60, 2484 (1992)
81. C. Durkan, I.V. Shvets, *J. Appl. Phys.* 79, 1219 (1996)
82. G. Krausch, S. Wegscheider, A. Kirsch, H. Bielefeldt, J.C. Meiners, J. Mlynek, *Opt. Commun.* 119, 283 and references therein (1995)
83. J. Newey (ed.) *Micro&nano Lett.* 1(1), 4 (2005)
84. P. Vettiger, *Europhys. News* 31, 8 (2000)
85. G. Louarn, S. Taleb, S. Cuenot, Excerpt from the Proceedings of the COMSOL Users Conference Paris (2006)
86. H. Heinzelmann, T.R. Huser, T.D. Lacoste, H.J. Güntherodt, D.W. Pohl, B. Hecht, L. Novotny, O.J. Martin, C.H. Hafner, H. Baggenstos, U.P. Wild, A. Renn, *Opt. Eng.* 34, 2441 (1995)
87. L. Novotny, D.W. Pohl, P. Regli, *J. Opt. Soc. Am. A* 11, 1768 (1994)
88. Y. Leviatan, *J. Appl. Phys.* 60, 1577 (1986)
89. S. Patanè, E. Cefali, S. Spadaro, R. Gardelli, M. Albani, M. Allegrini, *J. Microsc.* 229, 377 (2008)
90. J.A. Veerman, A.M. Otter, L. Kuipers, N.F. van Hulst, *Appl. Phys. Lett.* 72, 3115 (1998)
91. P.K. Wei, J. Hsu, W.S. Fann, B.R. Hseih, *Synth. Met.* 85, 1421 (1997)
92. G.M. Credo, G.M. Lowman, J.A. DeAro, P.J. Carson, D.L. Winn, S.K. Buratto, *J. Chem. Phys.* 112, 7864 (2000)
93. D. Richards, F. Cacialli, *Phil. Trans. R Soc. Lond. A* 362, 771 (2004)
94. E. Betzig, J.K. Trautman, R. Wolfe, E.M. Gyorgy, P.L. Finn, M.H. Kryder, C.H. Chang, *Appl. Phys. Lett.* 61, 142 (1992)
95. I. Smolyaninov, D.L. Mazzoni, C.C. Davis, *Appl. Phys. Lett.* 67, 3859 (1995)
96. R. Riehn, A. Charas, J. Morgado, F. Cacialli, *Appl. Phys. Lett.* 82, 526 (2003)
97. S. Madsen, M. Mullenborn, K. Birkelund, F. Grey, *Appl. Phys. Lett.* 69, 544 (1996)
98. S. Madsen, S.I. Bozhevolnyi, K. Birkelund, M. Mullenborn, J.M. Hvam, F.J. Grey, *J. Appl. Phys.* 82, 49 (1997)
99. S. Hosaka, A. Kikukawa, H. Koyanagi, T. Shintani, M. Miyamoto, K. Nakamura, K. Etoh, *Nanotechnology* 8, A58 (1997)
100. C. Mihalcea, W. Scholz, S. Werner, S. Münster, E. Oesterschulze, R. Kassing, *Appl. Phys. Lett.* 68, 3531 (1996)
101. E. Oesterschulze, O. Rudow, C. Mihalcea, W. Scholz, S. Werner, *Ultramicroscopy* 71, 85 (1998)
102. P. Biagioni, D. Polli, P.A. Labardi, G. Ruggeri, G. Cerullo, M. Finazzi, L. Duò, *Appl. Phys. Lett.* 87, 223112 (2005)
103. T. Masaki, Y. Inouye, S. Kawata, *Rev. Sci. Instrum.* 75, 3284 (2004)
104. P.N. Minh, T. Ono, M. Esashi, *Sens. Actuators A* 80, 163 (2000)
105. P.N. Minh, T. Ono, M. Esashi, *Rev. Sci. Instrum.* 71, 3111 (2000)
106. S. Kühn, U. Hakanson, L. Rogobete, V. Sandoghdar, *Phys. Rev. Lett.* 97, 017402 (2006)
107. L. Novotny, S.J. Stranick, *Ann. Rev. Phys. Chem.* 57, 303 (2006)
108. P. Royer, D. Barchiesi, G. Lerondel, R. Bachelot, *Phil. Trans. R Soc. Lond. A* 362, 821 (2004)

109. L. Novotny, E. Sanchez, X.S. Xie, *Ultramicroscopy* 71, 21 (1998)
110. J.T.I. Krug, E.J. Sanchez, X.S. Xie, *J. Chem. Phys.* 116, 10895 (2002)
111. F. H'Dhili, R. Bachelot, A. Rumanyantseva, G. Lerondel, P. Royer, *J. Microsc.* 209, 214 (2003)
112. A. Bouhelier, M.R. Beversluis, A. Hartschuh, L. Novotny, *Phys. Rev. Lett.* 90, 13903 (2003)
113. A. Bouhelier, M.R. Beversluis, L. Novotny, *Appl. Phys. Lett.* 82, 4596 (2003)
114. F. H'Dhili, R. Bachelot, G. Lerondel, D. Barchiesi, P. Royer, *Appl. Phys. Lett.* 79, 4019 (1999)
115. A. Tarun, M.R.H. Daza, N. Hayazawa, Y. Inouye, S. Kawata, *Appl. Phys. Lett.* 80, 3400 (2002)
116. X. Yin, N. Fang, X. Zhang, I.B. Martini, J. Schwartz, *Appl. Phys. Lett.* 81, 3663 (2002)
117. T. Tanaka, H.B. Sun, S. Kawata, *Appl. Phys. Lett.* 80, 312 (2002)
118. R. Bachelot, F. H'dhili, D. Barchiesi, G. Lerondel, R. Fikri, P. Royer, N. Landraud, J. Peretti, F. Chaput, G. Lampel, J.P. Boilot, K. Lahlil, *J. Appl. Phys.* 94, 2060 (2003)
119. P. Rochon, E. Batalla, A. Natansohn, *Appl. Phys. Lett.* 66, 136 (1994)
120. A. Sundaramurthy, P.J. Schuck, N.R. Conley, D.P. Fromm, G.S. Kino, W.E. Moerner, *Nano Lett.* 6(3), 355 (2006)
121. N. Murphy-DuBay, L. Wang, E.C. Kinzel, S.M.V. Uppuluri, X. Xu, *Opt Express* 16(4), 2584 (2008)
122. N. Murphy-DuBay, L. Wang, X. Xu, *Appl. Phys. A: Mater. Sci. Proc.* 93(4), 881 (2008)
123. S. Patanè, A. Arena, M. Allegrini, L. Andreozzi, M. Faetti, M. Giordano *Opt. Comm.* 210, 37 (2002)
124. A. Ambrosio, A. Camposeo, P. Maddalena, S. Patanè, M. Allegrini *J. Microsc.* 229, 307 (2008)
125. M. Labardi, N. Coppedè, L. Pardi, M. Allegrini, M. Giordano, S. Patanè, A. Arena, E. Cefali, *Mol. Cryst. Liq. Cryst.* 398, 33 (2003)
126. V. Likodimos, M. Labardi, L. Pardi, M. Allegrini, M. Giordano, A. Arena, S. Patanè, *Appl. Phys. Lett.* 82(19), 3313 (2003)
127. F. Gan, L. Hou (eds.) SPIE, vol 4085, in *Fifth International Symposium on Optical Storage, Proceedings* (2001)
128. G. Latini, A. Downes, O. Fenwick, A. Ambrosio, M. Allegrini, C. Daniel, C. Silva, P.G. Gucciardi, S. Patanè, R. Daik, W.J. Feast, F. Cacialli, *Appl. Phys. Lett.* 86, 011102-1 (2005)
129. T. Todorov, L. Nikolova, N. Tomova, *Appl. Opt.* 23, 4309 (1984)
130. R. Wuestneck, J. Stumpe, V. Karageorgieva, L.G. Meier, M. Rutloh, D. Presher, *Colloids Surf A* 198-200, 753 (2002)
131. P. Rochon, E. Batalla, A. Natansohn, *Appl. Phys. Lett.* 66, 136 (1995)
132. L. Cristofolini, S. Arisi, M.P. Fontana, *Phys. Rev. Lett.* 85, 4912 (2000)
133. P. Wu, D.V.G.L.N. Rao, B.R. Kimball, M. Nakashima, B.S. De Cristofano, *Appl. Phys. Lett.* 78, 1189 (2001)
134. N.K. Viswanathan, S. Balasumarian, L. Li, S.K. Tripathy, J. Kumar, *Jpn. J. Appl. Phys.* 38, 5928 (1999)
135. N.C.R. Holme, L. Nikolova, S. Hvilsted, T. Todorov, P.H. Ramussen, R.H. Berg, P.S. Ramanujam, *Appl. Phys. Lett.* 74, 519 (1999)
136. P. Camorani, L. Cristofolini, G. Galli, M.P. Fontana, *Mol. Cryst. Liq. Cryst.* 375, 175 (2002)

137. J. Kumar, L. Li, X.L. Jiang, D.Y. Kim, T.S. Lee, S. Tripathy, *Appl. Phys. Lett.* 72, 2096 (1998)
138. K. Sumaru, T. Fukuda, T. Kimura, H. Matsuda, T. Yamanaka, *J. Appl. Phys.* 91, 3421 and references therein (2002)
139. N. Landraud, J. Perett, F. Chaput, G. Lampel, J.P. Boilot, K. Lahlil, V.I. Safarov, *Appl. Phys. Lett.* 79, 4562 (2001)
140. P.S. Ramanujam, N.C.R. Holme, M. Pedersen, S. Hvilsted, *J. Photochem. Photobiol. A* 145, 49 (2001)
141. S. Arisi, P. Camorani, L. Cristofolini, M.P. Fontana, M. Laus, *Mol. Cryst. Liq. Cryst.* 372, 241 (2001)
142. M. Hasegawa, T. Ikawa, M. Tsuchimori, O. Watanabe, Y. Kawata, *Macromolecules* 34, 7471 (2001)
143. C.J. Barrett, P.L. Rochon, A.L. Natansohn, *J. Chem. Phys.* 109, 1505 (1998)
144. P. Camorani, M. Labardi, M. Allegrini, *Mol. Cryst. Liq. Cryst.* 372, 365 (2001)
145. L. Ramoino, M. Labardi, N. Maghelli, L. Pardi, M. Allegrini, S. Patanè, *Rev. Sci. Instrum.* 73, 2051 (2002)

A New AFM-Based Lithography Method: Thermochemical Nanolithography

Debin Wang, Robert Szoszkiewicz, Vamsi Kodali, Jennifer Curtis,
Seth Marder, and Elisa Riedo

Key words: Atomic force microscopy, Nanolithography, Proteins.

Abbreviations

AFM	Atomic force microscopy or atomic force microscope
DNA	Deoxyribonucleic acid
T_g	glass transition temperature
NSOM	Nearfield scanning optical microscopy
T_d	deprotection temperature
FTIR	Fourier transform infrared spectroscopy
F_N	normal load
SPDP	succinimidyl 3-(2-pyridyldithio)propionate
DTT	dithiothreitol
NHS	N-hydroxysuccinimide
GA	glutaraldehyde
DACM	(N-(7-dimethylamino-4-methylcoumarin-3-yl)maleimide)
DPN	Dip-pen nanolithography
TCNL	Thermochemical nanolithography
SAM	Self-assembled monolayer
p(THP-MA)	Poly(tetrahydro-2H-pyran-2-yl methacrylate);
p(THP-MA) ₈₀ p(PMC-MA) ₂₀	Poly(tetrahydro-2H-pyran-2-yl methacrylate) ₈₀ poly(3-{4-[(E)-3-methoxy-3-oxoprop-1-enyl]phenoxy}propyl 2-methacrylate) ₂₀ copolymer

22.1 Introduction

In the last decade, there has been a tremendous increase in the number of techniques for patterning materials on the nanoscale (10–100 nm), driven by numerous potential applications, for example, in sensing [1], data storage [2], optoelectronic [3], display [4], nanofluidic [5], and biomimetic [6] devices. An ideal nanolithography technique would be able to: (1) write with nm resolution; (2) write with speeds of multiple centimeters per second (while preserving nanometer scale registry) for wafer-scale lithography; (3) impart different chemical functionality and/or physical properties (with or without topographical changes) as desired; (4) function in different laboratory environments (for example, under ambient pressure or in solution); (5) be capable of massive parallelization for both writing and metrology; and (6) write on a variety of materials deposited on a variety of substrates. Specific applications will require one or more of the attributes described earlier, but the most versatile technique would encompass as many as possible. To our knowledge, no technique currently in practice can simultaneously attain all of these features.

Many nanoscale-patterning techniques, including AFM thermomechanical writing [2], AFM-assisted electrostatic lithography [7], scanning-beam lithography (electron- or ion-beam) [8], and block-copolymer templating [9], can provide *topographical* patterning through material deformation, removal, cross-linking, or degradation with resolution down to tens of nanometer in certain cases. *Chemical* patterning has been usually achieved via different strategies including direct assembly of chemically distinct regions by micro-contact printing [10], scanning probe-assisted deposition on the surface [11], and removal or manipulation of functional groups at the surface via photolithography [12], catalytic probe lithography [13,14] or other scanning probe-based methods [15,16]. Self-assembled structures of diblock copolymers can produce features with sub-30 nm linewidth resolution [17]; however, the self-assembly process cannot currently be tailored to afford any arbitrary structure. Scanning probe-based nanografting and replacement lithography techniques using self-assembled monolayers (SAMs) are able to produce chemical features with sizes of the order of 10 nm at speeds of 50 nm/s [18], while electrochemical lithography using a conductive AFM tip can write conducting polymer lines down to 45-nm wide at speeds of the order of micrometers per second [19]. Recently, structures as small as several nanometers have been patterned by immobilizing short, self-organizing DNA fragments onto a mica surface [20].

Dip-pen nanolithography (DPN) is an extremely versatile technique and has contributed a major advance in the field of chemical nanolithography [21]. DPN functions by facilitating the direct transport of molecules to surfaces, much like the transfer of ink from a macroscopic dip-pen to paper. By depositing several different kinds of molecules on the same substrate, DPN can pattern a range of desired chemistries with spatial control without exposing the substrate to harsh solvents, chemical etching, and/or extreme electrical field gradients. The intrinsic speed of DPN depends on molecular transport

between the probe tip and the surface, and is, therefore, limited by mass diffusion. Another challenge facing DPN lies in massive parallelization, which requires independent control of the force applied to, and the height of, the cantilever on or above the surface [21–23].

Herein, we review a simple yet extremely versatile scanning-probe method that is called thermochemical nanolithography. This technique offers some advantages relative to the lithographic methods noted earlier in terms of the combination of speed, resolution, materials flexibility, implementation toward massive parallelization, ability to chemically pattern surfaces and ability to work in a variety of environments. As such, it could have the ability to complement some of these techniques, specifically if high resolution over large surface areas is required.

22.2 Thermochemical Nanolithography

Thermochemical nanolithography (TCNL) employs a resistively heated AFM cantilever (Fig. 22.1A) to induce well-defined chemical reactions to change the surface functionality of thin polymer films (or, potentially, SAMs). Such an

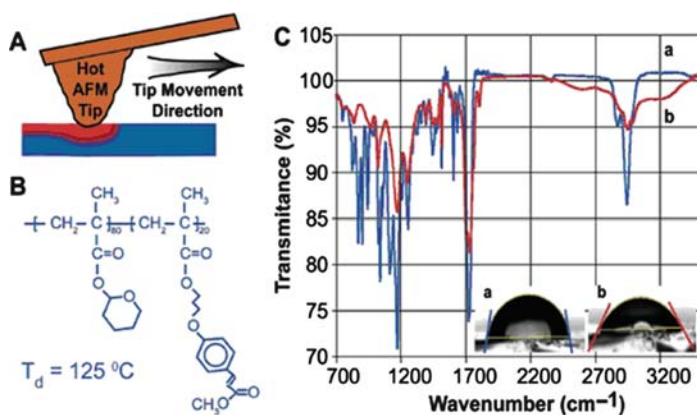


Fig. 22.1. Thermochemical nanolithography setup. (A) Experimental setup showing a resistively heated silicon cantilever scanning across a polymer sample. This heated AFM tip initiates thermal reactions of the polymer film above certain temperatures. (B) Structure of copolymer $p(\text{THP-MA})_{80} p(\text{PMC-MA})_{20}$. (C) Attenuated total-reflection FT-IR spectra of a bulk sample of cross-linked $p(\text{THP-MA})_{80} p(\text{PMC-MA})_{20}$ film (A) before heating and (B) after heating to $150 \pm 5^\circ\text{C}$. These show the growth of an -OH stretching band above $3,000\text{ cm}^{-1}$ and a new carbonyl peak at $1,720\text{ cm}^{-1}$ after heating, consistent with thermal deprotection to give the carboxylic acid. Accordingly, water contact angles show a change from (A) 76° to (B) 66° upon heating. When the film is heated above 180°C new peaks start to develop in the range $1,780$ and $1,820\text{ cm}^{-1}$, characteristic of anhydrides (not shown) [27]

approach is appealing as it is known that the thermal profile in the vicinity of a heated AFM tip can give rise to sharp thermal gradients [24, 25], and that chemical reaction rates increase exponentially with temperature; therefore one can, in principle, achieve a very high degree of spatial resolution. A wealth of thermally activated chemistries can feasibly be employed to change the subsequent reactivity, surface energy, solubility, conductivity, etc., of the material as desired. The distance of the tip from the surface and the temperature of the tip can be modulated independently, and the tip does not need to indent the surface. Therefore, chemical changes can be written very quickly through rapid scanning of the substrate or the tip, as no mass is transferred from the tip to the surface (writing speed is limited only by the heat transfer rate). In addition, judicious choice of the physical properties of a material (e.g., polymer glass transition temperature, T_g) may afford a system wherein chemical changes can be performed either separately from, or accompanied by, topographical modification as required (for example above T_g of a polymer, creation of substantial topographical changes (ripples) is exponentially amplified [26]). Furthermore, the use of a material that can undergo multiple chemical reactions at significantly different temperatures renders the possibility of a multi-state system wherein different functionalities can be addressed at different temperatures.

Szozzkiewicz et al. have recently shown how TCNL on a simple polymer can be used to write chemical features with 12-nm linewidth, while controlling the topography of the surface. It has been shown that TCNL can be performed at rates of millimeters per second [27, 28]. The technique is not limited to “SAM-friendly” surfaces as the polymer films can be deposited on a wide variety of substrates, and TCNL avoids the need for additional chemicals to be present at the surface and/or strong external electrical fields. Furthermore, Szozzkiewicz et al. showed patterning under high-humidity conditions, suggesting that TCNL may be extended to function in liquid environments. More recently, Wang et al. reported the use of TCNL to produce, at speeds of millimeters per second, nanopatterns of different orthogonal chemical functionalities on a polymer surface [28]. These tailored chemical nanopatterns were then used to direct the assembly of different nanoobjects, e.g., proteins, DNA, and C_{60} . In particular, they produced nanoassemblies, as small as 40 nm, of two different species of bioactive proteins coexisting on the surface.

Thermal lithography parallelization with AFM has been already demonstrated with the “millipede” technology, in which 5,000 heated tips can be used to write in parallel. Thus, with existing technology, TCNL could be scaled up to patterning speed of 1 m/s, i.e., 1×10^6 times faster than NSOM and 1×10^7 times faster than with electron beam lithography [2].

22.3 Thermal Unmasking of Chemical Groups on a Polymer

22.3.1 Unmasking Carboxylic Acid Groups

Szoszkiewicz et al. used an AFM tip to locally heat a thin film of poly(tetrahydro-2H-pyran-2-yl methacrylate)₈₀ poly(3-{4-[(E)-3-methoxy-3-oxoprop-1-enyl]phenoxy}propyl 2-methacrylate)₂₀ copolymer, p(THP-MA)₈₀ p(PMC-MA)₂₀ (Fig. 22.1b) which, after spin coating and photocross-linking of the cinnamate groups, exhibited no glass transition at or below the T_d . (Fig. 22.1c).

Szoszkiewicz et al. modified a pristine p(THP-MA)₈₀ p(PMC-MA)₂₀ film on glass by heating it locally with a silicon thermal cantilever to $160 \pm 30^\circ\text{C}$, where they expected a transformation from hydrophobic to hydrophilic. In ambient conditions, the Si tip is covered by a thin layer of native silicon dioxide, which makes the tip somewhat hydrophilic. The magnitude of the friction force between the tip and the sample surface is a sensitive relative measurement of the sample hydrophilicity, e.g., larger the friction force, the more hydrophilic the sample [30–32]. Figure 22.2A' shows a $12\ \mu\text{m} \times 12\ \mu\text{m}$ frictional force image of a copolymer sample, where they have chemically modified the left-hand side by means of local heating. The frictional force is clearly higher in the modified region, supporting the occurrence of a chemical change in this area. The AFM topography image taken over the same $12 \times 12\ \mu\text{m}$ area (Fig. 22.2A) shows no significant induced topographical differences, as desired. The slight difference in height ($\sim 1.6\ \text{nm}$) between the modified and unmodified areas is consistent with removal and vaporization of dihydropyran (bp = 84°C) upon heating. The insets of Fig. 22.2A and 22.2A' show a similar TCNL experiment, this time performed at a scanning velocity of $1.4\ \text{mm/s}$ (i.e., 16 times faster, but with similar results).

If topographical changes are desired, this can be achieved by varying the normal load, tip temperature and scanning velocity used during the TCNL process. Figure 22.2B and 22.2B' show topography and frictional force images, respectively, of a $1 \times 1\ \mu\text{m}$ copolymer area modified on the left side by heating. Due to the higher tip temperature and loads regular ripples appear in the topography. Figure 22.2B' shows that the friction force in the modified area is higher than in the unmodified area, consistent with a chemical change as described earlier (Fig. 22.2A'), and is essentially constant in this area. Thus, both the local topography and chemistry of the polymer film can easily be modified by simultaneously activating a chemical reaction and exploiting the mechanism of ripple formation, as required. Ripples of different sizes (ranging from 2 to 8 nm in height) can be created in a controlled manner on this substrate; however use of a p(THP-MA) homopolymer film enables the formation of even larger ripples, up to 300 nm in height. Szoszkiewicz et al. note that scanning with varying speed provides an easy way to create high-resolution hydrophilicity/hydrophobicity gradients, which could have important impact in the area of nanofluidic devices [33].

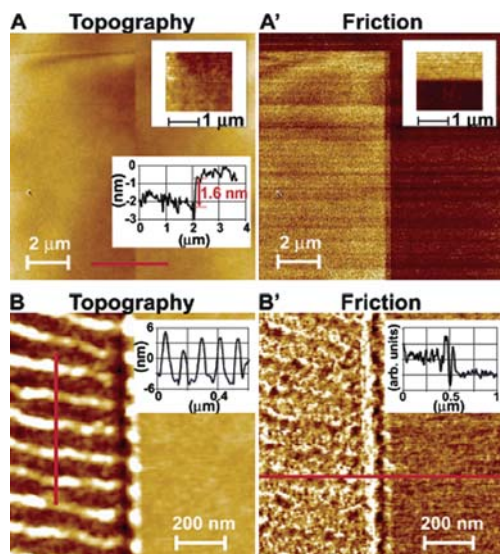


Fig. 22.2. TCNL induced change of surface chemistry. (A) $10 \times 10 \mu\text{m}$ AFM topography image and (A') corresponding friction image of a cross-linked $\text{p}(\text{THP-MA})_{80} \text{p}(\text{PMC-MA})_{20}$ film. The zone showing much higher friction (left side in A') was modified by local heating to $160 \pm 30^\circ\text{C}$ at a scanning speed of $85 \mu\text{m/s}$. Insets show analogous comparison of topography and friction changes between modified (*upper part*) and unmodified (*lower part*) regions of a $\text{p}(\text{THP-MA})_{80} \text{p}(\text{PMC-MA})_{20}$ film patterned at $1,420 \mu\text{m/s}$. (B) $1 \times 1 \mu\text{m}$ AFM topography image and (B') corresponding friction image of a cross-linked $\text{p}(\text{THP-MA})_{80} \text{p}(\text{PMC-MA})_{20}$ film recorded with a sharp contact AFM cantilever. Ripples in (B) ($10 \pm 3 \text{ nm}$) result from thermal modification by local heating at $180 \pm 30^\circ\text{C}$ while scanning at $85 \mu\text{m/s}$. The friction measured in the modified zone is noticeably larger than in the unmodified zone (inset); the changes in friction observable along the edges of ripples are presumably due to artifacts arising from the greater local contact area between the AFM tip and the sample while scanning across a sharp edge [27]

TCNL can be employed to write a controlled chemical pattern on a polymer surface with high density and at high resolution. Fig. 22.3A and 22.3A' show use of TCNL to write a chemical change on the copolymer (via deprotection of the carboxylic acid functionality) as a series of lines with a linear density of about 2×10^7 lines/meter (corresponding to 260 Gbit/in^2), in the absence of significant topographical changes (the differences visible in the topographical image arise from desorption of dihydropyran as before). Figure 22.3B and 22.3B' show topography and friction images of "GIT" written chemically on a copolymer sample. Figure 22.3B" gives the cross-section of a friction line, demonstrating that chemically modified lines can be created easily and reproducibly with width at half-maximum as small as 12 nm . The very

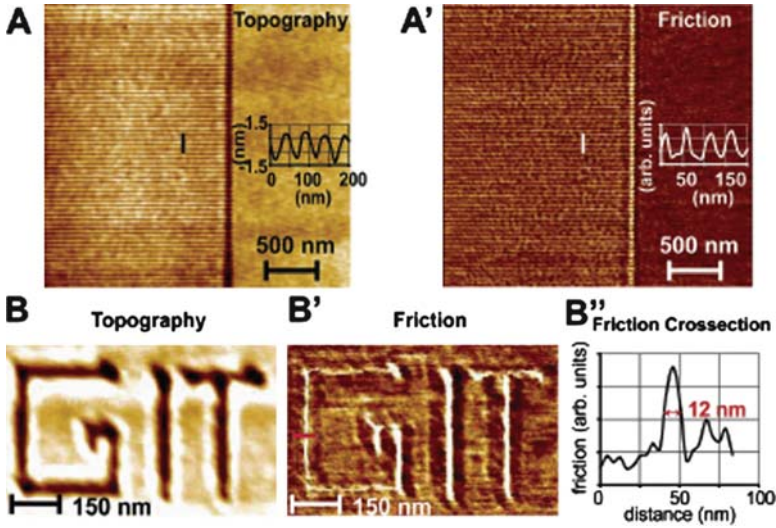


Fig. 22.3. High-resolution patterning. (A) $3 \times 3 \mu\text{m}$ AFM topography image and (A') corresponding friction image of a cross-linked $p(\text{THP-MA})_{80} p(\text{PMC-MA})_{20}$ film showing a high-density line pattern written chemically on the left side. This pattern was produced by modifying alternate lines at a speed of $9.6 \mu\text{m/s}$. (B) AFM topography and corresponding friction image (B') of a modified copolymer film written at a speed of $0.5 \mu\text{m/s}$, with the indentation depth kept within 3 nm. The resulting friction cross-section (B'') shows about 12-nm half-width within the modified zone (in the letter G); topographical changes are minimal, similar to those shown in Fig. 22.2A. Interestingly, the modified zone appears wider in the topography image than in the friction image; this is under further investigation but potentially has its origin in the convolution of the true topography with the tip profile [27]

small feature size achievable is attributed to the large temperature gradients in the polymer in the vicinity of the tip [24].

The fundamental limit to writing speed in TCNL is the thermal diffusivity of the substrate material, as opposed to mass diffusivity which limits deposition-based approaches. For example, the mass diffusivity of small molecules typically used in DPN is $\sim 10^{-10} \text{m}^2/\text{s}$ [34], while the thermal diffusivity of the organic substrates used in the present work is $\sim 10^{-7} \text{m}^2/\text{s}$ [35]. Thus, the speed of TCNL is currently limited by accessible AFM scanning velocity. However, modeling suggests that the maximum patterning speed can be estimated as about 30mm/s and is, therefore, much faster than any comparable chemical nanopatterning technology. In addition, the single-tip actuation time for resistively heated AFM cantilevers approaches $1 \mu\text{s}$, while thermomechanical actuation of similar tips requires closer to 1ms [23].

22.3.2 Unmasking Amines Groups

TCNL employs a resistively heated AFM tip to activate a chemical reaction at the nanometer scale on the surface of a material, for example by locally deprotecting a chemical moiety in the film of a specifically designed polymer. Very recently, Wang et al. showed that it is possible to extend the capabilities of TCNL by introduction of a new polymer to enable the selective patterning of amine groups [36]. Wang et al. synthesized a methacrylate copolymer containing tetrahydropyran (THP) carbamate groups that can be thermally deprotected to unmask primary amines. The mass loss after bulk heating of the polymer above the deprotection temperature, T_d , is consistent with this mechanism. This deprotection can be performed with a hot AFM tip maintained above T_d (between 150 and 220°C), exposing amine groups.

22.4 Two-Step Wettability Modification

In the past decade, the need to introduce local chemical modifications on a variety of substrates, particularly useful for the immobilization of biomolecules, has motivated the development of scanning probe-based nanopatterning techniques [37, 38]. Recently, Wang and collaborators demonstrated that it is possible to use TCNL to write-read-overwrite chemical patterns on thin polymer films in situ and with no need of probe change [36]. A hydrophilic pattern was written on the originally hydrophobic surface by scanning an AFM tip heated to $110 \pm 20^\circ\text{C}$. The wettability change introduced by the first local chemical modification was then reversed by a second chemical modification with further local heat-treatment to $190 \pm 20^\circ\text{C}$. This additional overwriting capability can facilitate the design of complex nanopatterns, with tunable wettability, particularly desirable in the fabrication of chemical/biological sensors and nanofluidic devices.

Chemical modifications in the *macroscopically* heat-treated samples were characterized in ambient conditions by Fourier transform infrared spectroscopy (FTIR) (UMA-600, Digilab) (Fig. 22.4), static water contact angle measurements (Phoenix 150, SEO), and AFM friction measurements (CPII, Veeco) (Fig. 22.5). The FTIR spectrum collected on the untreated, e.g., not heated, sample is characterized by the presence of the carbonyl band corresponding to the stretching vibration of the C = O bond in esters around $1,730\text{ cm}^{-1}$ [39]. The FTIR spectrum of the sample heated to 70°C shows no significant shift of the ester band. The static water contact angle is 77° for temperatures at 70°C , where no chemical modification is expected. A thermogravimetric analysis indicates that the removal of THP groups starts around 110°C (data not shown) and the FTIR spectrum of the sample heated to 130°C shows that the carbonyl band shifts from $1,730$ to $1,700\text{ cm}^{-1}$, consistent with the THP removal and formation of carboxylic acids. The friction force, measured by AFM at a normal load $F_N = 70 \pm 2\text{ nN}$, increases by $48 \pm 7\%$ upon heating to

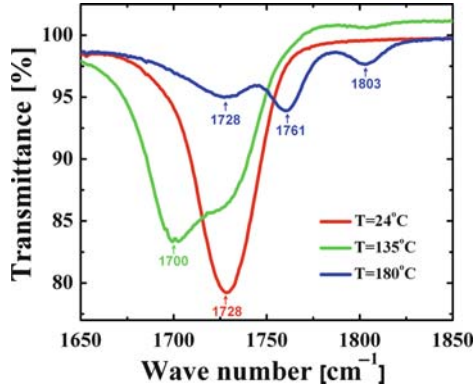


Fig. 22.4. FTIR spectra. FTIR spectra of a p(THP-MA)₈₀ p(PMC-PMCA)₂₀ copolymer upon heating to selected temperatures [28]

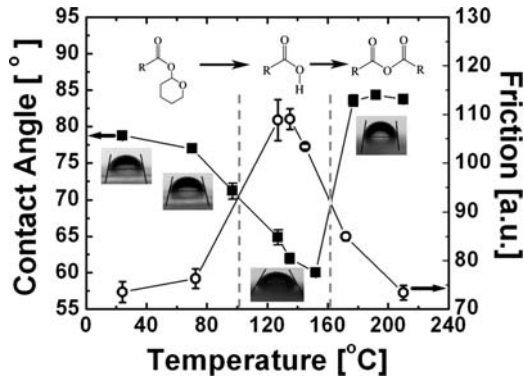


Fig. 22.5. Contact angle measurements. Static water contact angle (*solid squares*) and friction force (*open circles*) as a function of the temperature used to macroscopically heat a p(THP-MA)₈₀ p(PMC-MA)₂₀ copolymer. Contact angle images of the samples without heating treatment, heated to 70, 135, and 180°C, are shown as insets. The error bars represent the standard deviations [28]

135°C. The friction force increase in the chemically modified area is consistent with a local change of the wettability, i.e., the originally hydrophobic surface becomes hydrophilic [40]. This is confirmed by the static water contact angle that decreases from $79 \pm 0.2^\circ$ to $62 \pm 1.0^\circ$ as the temperature increases from 70 to 135°C. After further heating above 180°C, the appearance of two additional peaks at 1,760 and 1,802 cm^{-1} on the FTIR spectrum confirms the conversion of carboxylic acids into anhydrides. This second chemical modification is accompanied by a water contact angle increase to $84 \pm 0.7^\circ$. Consistently, the friction force (at $F_N = 70 \pm 2$ nN) decreases to a value almost equal to that of the untreated sample. The wettability change introduced by the first chemical

modification (THP to carboxylic acid) is fully reversed by the second chemical modification (carboxylic acid to anhydride).

Local chemical modifications were performed by scanning the sample with an AFM tip heated to a selected temperature, in the contact mode with 29 ± 5.8 nN of applied normal force, at a speed of $6 \mu\text{m/s}$, and in a $41 \pm 1.3\%$ humidity environment. Topographic and friction images of the sample were collected before and after the local thermochemical modification under the same conditions, at room temperature.

The $p(\text{THP} - \text{MA})_{80}p(\text{PMC} - \text{MA})_{20}$ copolymer surface was first heated to $70 \pm 20^\circ\text{C}$, below the THP deprotection temperature, by means of a thermal cantilever. No change in topography or friction was detected on the heated area (Fig. 22.6). After heating a $1.5 \times 1.5 \mu\text{m}$ square to $110 \pm 20^\circ\text{C}$ a corresponding pattern in the friction image is observed; the topography shows no depletion inside the square. The friction increase in the written pattern suggests that the THP groups were deprotected, leaving the area covered with hydrophilic carboxylic acid groups. A second smaller $0.65 \times 0.4 \mu\text{m}$ square pattern was overwritten inside the hydrophilic pattern by further heating

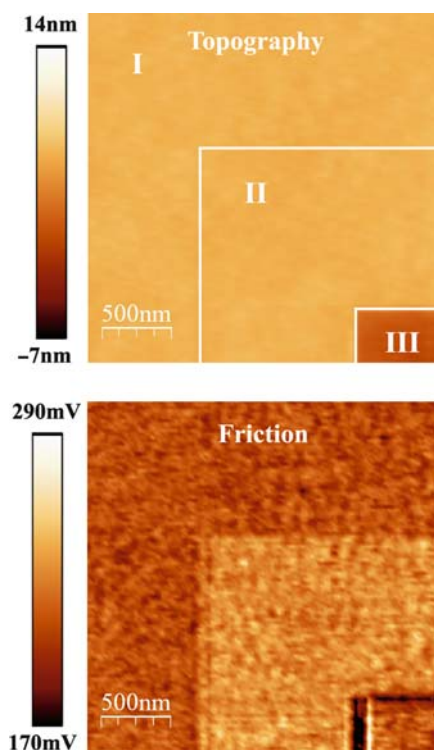


Fig. 22.6. Two-Steps wettability modification. Topography and friction images of a $p(\text{THP-MA})_{80} p(\text{PMC-MA})_{20}$ copolymer surface before (I) and after a first (II) and second (III) TCNL modification, and second (III) TCNL modification [28]

to $190 \pm 20^\circ\text{C}$. The friction image shows that after a second local chemical modification at $190 \pm 20^\circ\text{C}$ the surface becomes again hydrophobic (lower friction). The corresponding topography depletion is 6-nm deep. This change is consistent with the anhydride formation observed for macroscopic heating.

22.5 Covalent Functionalization and Molecular Recognition

Wang et al. demonstrated that by using covalent functionalization and molecular recognition it is possible to create multi-functional nanopatterns coexisting on the same surface, that later can be selectively functionalized with the desired species of nanoobjects [36]. The approach is conceptually simple, as outlined in Fig. 22.7. After using TCNL to write the first of the desired amine

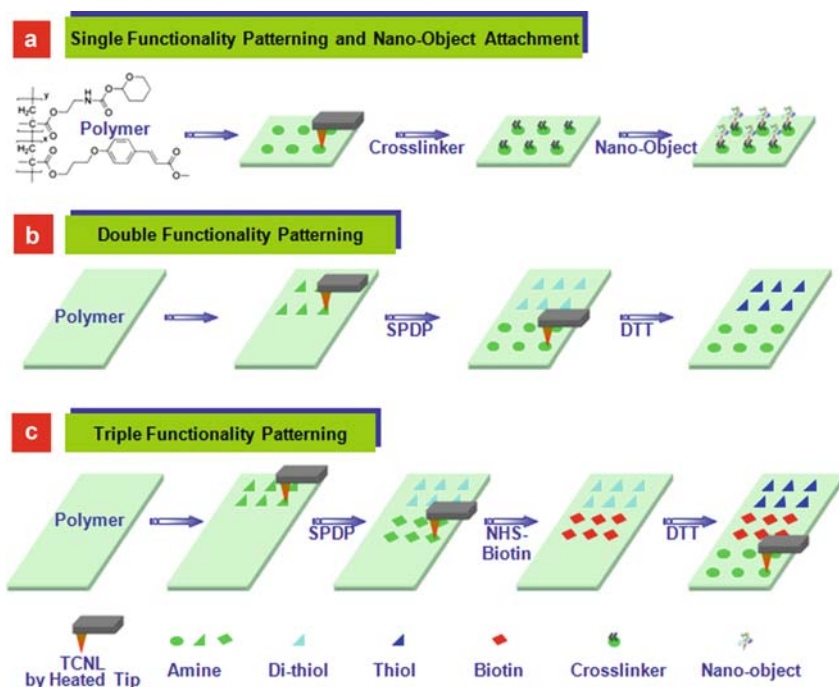


Fig. 22.7. Multiple functionality patterning flow chart. (a) A single nano-object pattern is created through three steps: TCNL, crosslinker incubation and nano-object incubation. (b) A double functionality pattern of thiols (*blue triangles*) and amines (*green disks*) is created through two rounds of TCNL and incubation processes. (c) A triple functionality pattern of thiols (*blue triangles*), biotins (*red diamonds*), and amines (*green disks*) is created through three rounds of TCNL and incubation processes [28]

patterns, the deprotected amines can be thiolated via a reaction with *N*-succinimidyl 3-(2-pyridyldithio) propionate (SPDP) (triangles, Fig. 22.7b and c). A second TCNL pattern can then be created on a different area of the same polymer surface (diamonds, Fig. 22.7b and c), exposing additional amines (Fig. 22.7b) in a new region. The di-thiols obtained after the first patterning step are then reduced to thiols using dithiothreitol (DTT) [41], thereby producing a surface with tailored patterns of amine and thiol groups.

Alternately, Wang et al. reacted the amine pattern from the second application of TCNL, with *N*-hydroxysuccinimide (NHS)-biotin, transforming the amine functionality to biotin. At this stage, reduction with DTT can create a co-patterned surface of thiols and biotins. In order to nanopattern surfaces with three different functionalities (Fig. 22.7c), a third application of TCNL produces a triple-patterned surface consisting of thiols, biotins, and amines.

Once the surfaces have been patterned with the desired chemical groups, they can be stored in a desiccator for later functionalization with proteins or other nanoobjects. So far, Wang et al. have stored the nanopatterned surfaces for a maximum duration of 3 weeks before successful two-protein functionalization. Future studies will investigate longer storage times.

Figure 22.8a shows the binding of different fluorescently labeled proteins to the TCNL amine patterns via two different crosslinking mechanisms, NHS-biotin and glutaraldehyde (GA). The fluorescent patterns of single-protein species in Fig. 22.8a show bound streptavidin, anti-CD3, and fibronectin. Images produced using epi-fluorescence microscopy verified that fluorescent patterns are consistent with the patterns written by TCNL and with the post-patterning functionalization utilized.

The viability of multi-functional patterns was then verified using the approach described in Fig. 22.8c. A co-pattern of thiols and biotins was fluorescently labeled by incubation of the surface first with the blue fluorescent dye DACM (*N*-(7-dimethylamino-4-methylcoumarin-3-yl) maleimide) and then the red-emitting Cy5-streptavidin, in order to determine if they would selectively bind to the thiol and biotin patterns, respectively [41]. Figure 22.8c shows the resultant fluorescent patterns of thiols (blue triangles) and biotins (red diamonds). The two fluorophores were imaged separately on a monochrome camera using a DAPI filter set for the DACM and a Cy5 filter set for the Cy5-streptavidin. The third combined image represents the overlay of the two independently acquired images. Little crosstalk was observed between the two channels, demonstrating both the effective localization of the dyes onto the TCNL pattern and the specificity of their chemical crosslinking to the surface.

TCNL is capable of extremely high-resolution chemical patterning of surfaces due to the high-temperature gradients in the vicinity of the heated tip [27]. Therefore, Wang et al. investigated their ability to organize small amounts of proteins with high definition. An array of 500 nm features decorated with fluorescently labeled fibronectin is shown in Fig. 22.9. The AFM phase and epi-fluorescence images are in agreement with what would be

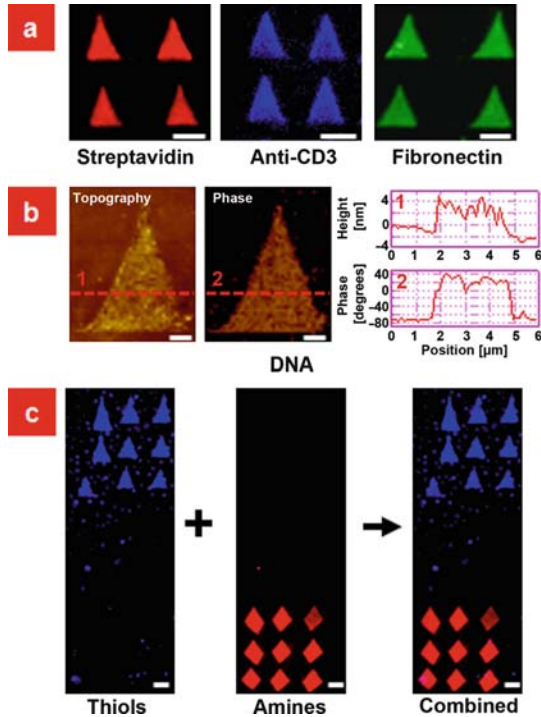


Fig. 22.8. Multifunctional nanopatterning. (a) Epi-fluorescence images of micro triangle patterns of Cy5-streptavidin (*red*) (crosslinked to amine via NHS-biotin), biotinylated Alexa350-antiCD3 (*blue*) (crosslinked to NHS-biotinylated amine with streptavidin), and atto488 fibronectin (*green*) (crosslinked to amine with GA). (b) AFM tapping mode topography and phase images of a triangular pattern of thiol-terminated DNA single strands crosslinked to amines through PMPI. (c) Epi-fluorescence images of DACM and Cy5-streptavidin orthogonal functionality patterns on the same surface. Blue DACM was crosslinked to amine-terminated triangles by means of SPDP-DTT thiolation, while red Cy5-streptavidin was crosslinked to amine-terminated diamonds by means of NHS-biotin. Scale bars: (a) 5 μm , (b) 1 μm , (c) 10 μm [28]

expected from the TCNL patterning. The AFM phase image provides information on the local viscoelasticity of the sample. A larger phase change is indicative of a “softer” surface, which is what Wang et al. expected in the patterns decorated with proteins.

To investigate length scales below the resolution limit of optical microscopy, Wang et al. employed topographical and phase AFM imaging in ambient conditions. Using a single touchdown approach at each feature, TCNL was used to de-protect amines in extremely small areas. Three such features are shown in Fig. 22.10 as topographical and phase images. The topography indicates a shallow indentation of approximately 10 nm. This depth can be varied by

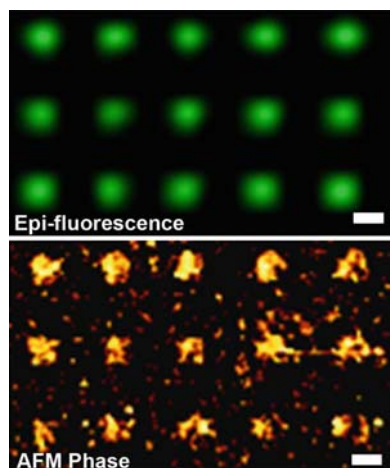


Fig. 22.9. Five hundred nanometers fibronectin nanoarray. Epi-fluorescence and AFM phase images of a regular array of Atto488 fibronectin patches. Fibronectin was crosslinked to amine groups by means of GA. Scale bar: 500 nm [28]

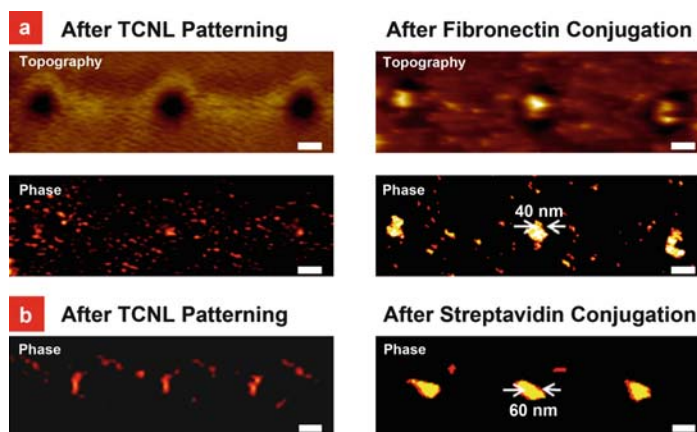


Fig. 22.10. High-resolution protein patterning. Fibronectin and streptavidin nanoarrays down to 40 nm. AFM topography and phase images of a TCNL nanoarray before and after fibronectin (a) and streptavidin (b) attachment. Scale bar: 100 nm [28]

changing the temperature of the tip, higher temperatures producing deeper indentations [27, 42]. The surface was then functionalized with fibronectin (Fig. 22.10a) or with streptavidin (Fig. 22.10b). The topographical data revealed that the TCNL “holes” were filled with proteins. The phase images are also consistent with the deposition of proteins in the holes. Fibronectin phase features as small as 40 nm and streptavidin as small as 60 nm have been measured. The former value compares well with previous measurements of

single dry fibronectin molecules [43]. Considering the AFM tip convolution during imaging and a tip radius of about 15 nm, it is reasonable to assume that there are as few as 1 or 2 fibronectin molecules covalently attached to each location.

References

1. Y.H. Wang et al., Controlling the shape, orientation, and linkage of carbon nanotube features with nano affinity templates. *Proc. Natl. Acad. Sci. USA* 103, 2026–2031 (2006)
2. P. Vettiger et al., The “Millipede” - More than one thousand tips for future AFM data storage. *Ibm. J. Res. Dev.* 44, 323–340 (2000)
3. J. Wang,, X.Y. Sun, L. Chen, L. Zhuang, S.Y. Chou, Molecular alignment in sub-micron patterned polymer matrix using nanoimprint lithography. *Appl. Phys. Lett.* 77, 166–168 (2000)
4. J.A. Rogers et al., Paper-like electronic displays: Large-area rubber-stamped plastic sheets of electronics and microencapsulated electrophoretic inks. *Proc. Natl. Acad. Sci. USA* 98, 4835–4840 (2001)
5. D. Mijatovic, J.C.T. Eijkel, A. van den Berg, Nanofluidic methods review. *Lab Chip* 5, 492–500 (2005)
6. K.H. Jeong, J. Kim, L.P. Lee, Biologically inspired artificial compound eyes. *Science* 312, 557–561 (2006)
7. S.F. Lyuksyutov et al., Electrostatic nanolithography in polymers using atomic force microscopy. *Nat. Mater.* 2, 468–472 (2003)
8. S. Matsui, Nanostructure fabrication using electron beam and its application to nanometer devices. *Proc. IEEE.* 85, 629–643 (1997)
9. M. Park, C. Harrison, P.M. Chaikin, R.A. Register, D.H. Adamson, Block copolymer lithography: Periodic arrays of similar to 10(11) holes in 1 square centimeter. *Science* 276, 1401–1404 (1997)
10. Y.N. Xia, G.M. Whitesides, Soft lithography. *Annu. Rev. Mater. Sci.* 28, 153–184 (1998)
11. A.A. Tseng, A. Notargiacomo, T.P. Chen, Nanofabrication by scanning probe microscope lithography: A review. *J. Vacuum Sci. Technol. B* 23, 877–894 (2005)
12. S.Q. Sun, K.S.L. Chong, G.J. Leggett, Photopatterning of self-assembled. monolayers at 244 nm and applications to the fabrication of functional microstructures and nanostructures. *Nanotechnology* 16, 1798–1808 (2005)
13. W.T. Muller, et al., A strategy for the chemical synthesis of nanostructures. *Science* 268, 272–273 (1995)
14. M. Peter, X.M. Li, J. Huskens, D.N. Reinhoudt, Catalytic probe lithography: Catalyst-functionalized scanning probes as nanopens for nanofabrication on self-assembled monolayers. *J. Amer. Chem. Soc.* 126, 11684–11690 (2004)
15. R.M. Nyffenegger, R.M. Penner, Nanometer-scale surface modification using the scanning probe microscope: Progress since 1991. *Chem. Rev.* 97, 1195–1230 (1997)
16. P. Samori, *Scanning Probe Microscopies Beyond Imaging* (Wiley-VCH, Weinheim, 2006)
17. M.P. Stoykovich et al., Directed assembly of block copolymer blends into nonregular device-oriented structures. *Science* 308, 1442–1446 (2005)

18. S. Kramer, R.R. Fuieler, C.B. Gorman, Scanning probe lithography using self-assembled monolayers. *Chem. Rev.* 103, 4367–4418 (2003)
19. S.Y. Jang, M. Marquez, G.A. Sotzing, Rapid direct nanowriting of conductive polymer via electrochemical oxidative nanolithography. *J. Amer. Chem. Soc.* 126, 9476–9477 (2004)
20. U. Feldkamp, C.M. Niemeyer, Rational design of DNA nanoarchitectures. *Angew. Chem. Int. Edit.* 45, 1856–1876 (2006)
21. D.S. Ginger, H. Zhang, C.A. Mirkin, The evolution of dip-pen nanolithography. *Angew. Chem. Int. Edit.* 43, 30–45 (2004)
22. M. Su, Z.X. Pan, V.P. Dravid, T. Thundat, Locally enhanced relative humidity for scanning probe nanolithography. *Langmuir* 21, 10902–10906 (2005)
23. D. Bullen et al., Design, fabrication, and characterization of thermally actuated probe Arrays for dip pen nanolithography. *J. Microelectromech. Syst.* 13, 594–602 (2004)
24. D.S. Fryer, P.F. Nealey, J.J. de Pablo, Thermal probe measurements of the glass transition temperature for ultrathin polymer films as a function of thickness. *Macromolecules* 33, 6439–6447 (2000)
25. W.P. King et al., Atomic force microscope cantilevers for combined thermomechanical data writing and reading. *Appl. Phys. Lett.* 78, 1300–1302 (2001)
26. B. Gotsmann, U. Durig, Thermally activated nanowear modes of a polymer surface induced by a heated tip. *Langmuir* 20, 1495–1500 (2004)
27. R. Szoszkiewicz et al., High-speed, sub-15 nm feature size thermochemical nanolithography. *Nano Lett.* 7, 1064–1069 (2007)
28. D. Wang et al., Local wettability modification by thermochemical nanolithography with write-read-overwrite capability. *Appl. Phys. Lett.* 91, 3 (2007)
29. D. Wang et al., Reversible nanoscale local wettability modifications by thermochemical nanolithography, *Mater. Res. Soc. Symp. Proc.* 1059, KK10-36 (2008)
30. R.W. Carpick, M. Salmeron, Scratching the surface: Fundamental investigations of tribology with atomic force microscopy. *Chem. Rev.* 97, 1163–1194 (1997)
31. R. Szoszkiewicz, E. Riedo, Nanoscopic friction as a probe of local phase transitions. *Appl. Phys. Lett.* 87, 033105 (2005)
32. R. Szoszkiewicz, A.J. Kulik, G. Gremaud, M. Lekka, Probing local water contents of in vitro protein films by ultrasonic force microscopy. *Appl. Phys. Lett.* 86, 123901 (2005)
33. S.M. Morgenthaler, S. Lee, N.D. Spencer, Submicrometer structure of surface-chemical gradients prepared by a two-step immersion method. *Langmuir* 22, 2706–2711 (2006)
34. P.E. Sheehan, L.J. Whitman, Thiol diffusion and the role of humidity in “dip pen nanolithography”. *Phys. Rev. Lett.* 88, 156104 (2002)
35. E.A. Grulke, A. Abe, D.R. Bloch, *Polymer Handbook*. (Wiley, New York, 2003)
36. D. Wang et al., Thermochemical nanolithography of multifunctional nanotemplates for assembling nano-objects, *Adv. Funct. Mater.* in press (2009)
37. R. Garcia, R.V. Martinez, J. Martinez, Nano-chemistry and scanning probe nanolithographies. *Chem. Soc. Rev.* 35, 29–38 (2006)
38. X.N. Xie, H.J. Chung, C.H. Sow, A.T.S. Wee, Nanoscale materials patterning and engineering by atomic force microscopy nanolithography. *Mater. Sci. Eng. R-Reports* 54, 1–48 (2006)
39. G.W. Gokel, *Dean’s Handbook of Organic Chemistry* (McGraw-Hill, New York, 2004)

40. R. Szoszkiewicz, E. Riedo, Nanoscopic friction as a probe of local phase transitions. *Appl. Phys. Lett.* 87, 033105 (2005)
41. G.T. Hermanson, *Bioconjugate Techniques*, 1st ed. (Academic, London, 1996)
42. W.P. King, K.E. Goodson, Thermomechanical formation of nanoscale polymer indents with a heated silicon tip. *J. Heat Transfer-Trans. ASME* 129, 1600–1604 (2007)
43. Y. Hu, A. Das, M.H. Hecht, G. Scoles, Nanografting de novo proteins onto gold surfaces. *Langmuir* 21, 9103–9109 (2005)

Scanning Probe Alloying Nanolithography

Luohan Peng, Hyungoo Lee, and Hong Liang

Summary. The probe-based nanolithography method is attracting increasing attentions due to its unique and simple process characteristics. This chapter, first, provides a brief review of nanolithography in order to see where the probe-based nanolithography current stands. Subsequently, it discusses the roles of probe-based nanolithography in the nanoscience and technology. Followed are some examples of the nanofabrication processes using the probe-based lithography method. At the later part of the chapter, new applications in nanofabrication and nanoalloying are discussed.

Key words: Scanning probe based nanolithography, Nanofabrication, Nanoalloying, Nanofabrication, Nanostructures, Scanning probe alloying nanolithography (SPAN), Scanning probe-based nanolithography.

23.1 Brief Review of Nanolithography

23.1.1 Introduction

Lithographic techniques at nanometer scale include optical lithography, electron beam lithography (EBL), focused ion beam lithography, and scanning probe microscope (SPM)-based nanolithography. Lithography literally means “writing on stone.” The basic principle of lithography was first investigated by an Austrian author Alois Senefelder [1], and the technique experienced an explosive boost in the mid-twentieth century. The history of developed lithographic techniques [2] can be found in Table 23.1.

In modern nanoscience and engineering, a critical issue is to create metallic nanostructures with few defects. These nanostructures have been created through a number of methods in either top-down or bottom-up approaches. Planar technology is a typical top-down approach where nanostructures are fabricated layer-by-layer with precise control [3]. In this approach, high-resolution lithographic techniques, such as optical lithography, EBL, and x-ray

Table 23.1. List of lithography techniques developed in history

Methods	Author	Year
Lithography	A. Senefelder	1798
Optical lithography	G. Moore	1965
Electron beam lithography	D.A. Buck et al.	1957
X-ray lithography	E. Spears et al.	1972
Extreme ultraviolet lithography	A.M. Hawryluk et al.	1988
Imprint lithography	G. Whitesides et al.	1994
Scanning probe microscope-based lithography	M. Ringger et al.	1985

lithography, are used to transfer the nanoscale pattern on the substrate. Metal thinfilm are then deposited via vacuum coating processes, followed by either etching or lift-off to finalize the structure. Alternatively, focused ion beam can be used to directly carve metal thinfilm into nanostructures [4]. The bottom-up approach relies on the molecules' self-assembly process where very fine nanostructures are built up atom by atom, molecule by molecule, or cluster by cluster. This approach is less conventional than the top-down approach, but it is synthesized from fundamental building blocks based on different physical or chemical processes [5–9].

The resolution of regular optical lithography strongly depends on the numerical apertures (NAs) of the optics and the wavelength of the light source. A higher NA and lower wavelength will provide better resolution. Immersion lithography has an advantage in constructing the projection optics in an immersion fluid with numeric aperture greater than one. With immersion lithography, patterns less than 90 nm have been demonstrated [10]. On the contrary, in extreme ultraviolet lithography, an extremely short wavelength of 13.4 nm is applied. Patterns with feature sizes of less than 50 nm are expected to be realized [11, 12]. In addition, there are other unconventional lithographic techniques that offer high spatial resolutions of features. Nanoimprint lithography has proven to be a promising patterning technique at nanoscale [13]. In 1999, Colbrun et al. reported step- and flash imprint lithography [14], which has led recent research and development forward. Two-photon 3D lithography [15], a state-of-the-art technique, is able to fabricate 3D micro/nanostructures required in more complicated systems, such as photonic crystals [16] and microfluidic devices [17].

There is a critical problem with the photolithography, i.e., resolution. The resolution is limited by optical diffraction. As show in the following equations, the minimum printed-feature size (R , resolution) depends on the wavelength (λ) and the Numerical Aperture (NA) of exposure system [18–26].

$$R = k_1 \frac{\lambda}{\text{NA}}; \quad \text{DOF} = k_2 \frac{\lambda}{\text{NA}^2}; \quad \text{NA} = n \sin(\theta) \quad (23.1)$$

where R is resolution, k_1 and k_2 are constants (normally 0.5–0.8 for k_1 and 0.1–0.5 for k_2) depending on the resist technology, λ is the wavelength of the transmitted light, NA is the numerical aperture (the sine of the convergence angle of the lens), and DOF is the depth of focus.

For higher resolution, the wavelength λ should be decreased and the NA should increase. Thus, the wavelength has been reduced from 435 nm (g-line [27–32]) to 405 nm (h-line [33]) to 365 nm (i-line [34]) to 248 nm (deep ultraviolet [35]). The critical resolutions are 0.5, 0.4, 0.35, and 0.25 μm , respectively. The depth of focus will be decreased by increasing the resolution. The decreased depth of focus will reduce the process latitude of the optical system [26].

One of other problems involved with lithography is the alignment [36–44]. For fabricating an integrated circuit, a number of layers should be printed. Typically 15–25 layers are demanded. Those layers should be precisely aligned for fabricating a nondefected chip. It is difficult to make a precise alignment for each layer. This leads to a trade-off between alignment accuracy and alignment time. In addition, as the patterns on a chip get smaller, the alignment becomes increasingly difficult. There are alternative nanolithography methods to solve the problems of above-mentioned lithography. One of the promising technologies is the AFM lithography [45–70].

23.1.2 Probe-Based Lithography

The AFM Lithography has significant advantages over the regular lithography. With AFM lithography, the patterns can be fabricated on various materials such as metals, semiconductors, and polymers. AFM lithography can be processed in air, liquid, or vacuum environments. The fabricated patterns by AFM lithography can be characterized by using AFM immediately. There are many AFM-based lithography methods [71–78]. In general, the methods can be categorized into two groups; force-based lithography or bias-based ones. For force-based lithography, larger forces are loaded on the substrate surface, which fabricate a structure on the surface. The SPM has been widely used in many disciplines for surface science research and applications since its invention in the early 1980s [79]. In recent years, researchers have realized that atomic interaction between the scanning probe and interrogated sample surface not only is useful for nanoimaging but also can be utilized for revolutionizing conventional bottom-up nanofabrication methods. These methods include nanomanipulation, nanodeposition, and nanopatterning. The SPM-based nanofabrication processes offer unique advantages of high resolution and registration accuracy, in-vivo inspection capability readily available, process automation with computer control, and low-cost process under ambient conditions. Application for mass production is also possible with large scanning probe arrays [80–82]. For material deposition and patterning, a popular SPM-based technique called the dip-pen nanolithography (DPN) [83–85] has

been developed. With the sharp scanning probe tip, DPN can directly dispense a wide variety of solution-based chemicals onto a substrate surface to create single or arrayed nanoscale pattern or structures [86–88]. The effectiveness of DPN relies on the material transfer from the probe tip to the substrate surface via the water condensation existing in between [89, 90].

23.1.3 Probe Materials and Properties

The spring constant is the ratio of the applied force to the deflection of the cantilever. For AFM oxidation lithography, probes are needed to have low spring constant. In the processes of oxidation lithography, the probes can remove the oxidized structures from the substrate surface. There are two main types of AFM probes; rectangular shape and V-shaped probes, as shown in Fig. 23.1.

For those probe types, the spring constant (k) can be calculated from the following equations:

$$\text{For rectangular probe, } k = \frac{Et^3w}{4L} \quad (23.2)$$

$$\text{For V-shaped probe, } k = \frac{Et^3w}{2b(L_1^3 - L_2^3) + 6wL_2^3} \quad (23.3)$$

where E is the Young's modulus, t is the thickness, w is the width, and L is the length of the cantilever [91]. According to the equations, one can decrease the spring constant with decreasing the thickness and increasing the length (L or L_1). The young's modulus (E) is dependent on the materials of the probes. Table 23.2 shows two main materials of probes having different dimensions (l , w , and t); SiO_2 and Si_3N_4 . The k values in the last column of Table 23.2 were measured. From the first three rows in the table, as the length (l) increases, the spring constant (k) decreases for the rectangular shape probe.

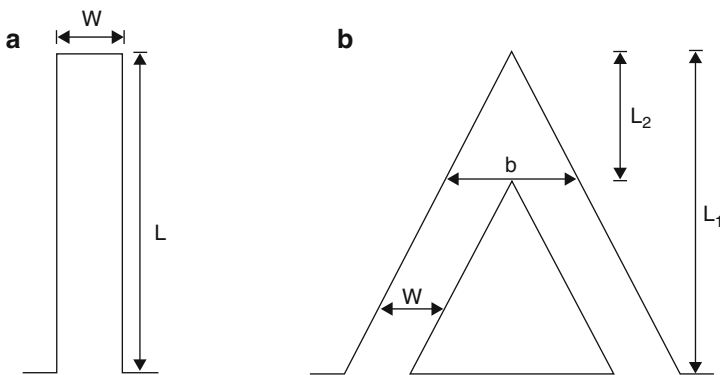


Fig. 23.1. The main types of AFM; (a) rectangular probe and (b) V-shaped probe

Table 23.2. Examples of mechanical properties of cantilevers [92]

Probe number	Shape	Material	$l(\mu\text{m})$	$w(\mu\text{m})$	$t(\mu\text{m})$	$k(\mu\text{m})$
1	Rectangular	SiO ₂	100	20	1.5	1.0
2	Rectangular	SiO ₂	180	20	1.5	0.16
3	Rectangular	SiO ₂	240	20	1.5	0.07
4	Rectangular	SiO ₂	500	5	1.0	0.0006
5	Rectangular	SiO ₂	500	10	1.0	0.001
6	Rectangular	SiO ₂	500	20	1.0	0.002
7	Rectangular	Si ₃ N ₄	100	20	0.4	0.05
8	Rectangular	Si ₃ N ₄	100	20	0.7	0.25
9	V	SiO ₂	100	13	1.5	1.3
10	V	SiO ₂	100	22	1.5	2.2
11	V	Si ₃ N ₄	100	13	0.4	0.06
12	V	Si ₃ N ₄	100	13	0.7	0.34
13	V	Si ₃ N ₄	100	22	0.4	0.11
14	V	Si ₃ N ₄	100	22	0.7	0.59
15	V	Si ₃ N ₄	200	18	0.4	0.01
16	V	Si ₃ N ₄	200	18	0.7	0.06
17	V	Si ₃ N ₄	200	36	0.4	0.02
18	V	Si ₃ N ₄	200	36	0.7	0.12

When increasing only the width (w), k is also increased as expected from the Eq. (23.2). Although the thickness increased from 0.4 to 0.7 μm (probe 7 and 8), the value of the spring constant of the probe 8 increased five times more than the probe 7. Comparing the probe 9 with 11, the calculated k value of the probe 11 should be about 0.0246. However, the measured k value was 0.06 because the materials of the probes were different. The Young's modulus of SiO₂ (70 GPa) is smaller than that of Si₃N₄ (380 GPa). The material, dimension, and shape should be considered in lithography.

23.1.4 Probe Wear

Research effort has been made to study wear of tips during contact mode AFM sliding [93–99]. The tip wear should be eliminated for AFM imaging or fabrication of nanostructures. Understanding the wear of tips has yet be obtained. It was founded that the end of the AFM tips became amorphous during the sliding process [99]. It was generally accepted that Archard's wear equation is utilized for quantification of wear of the amorphous layers. According to the Archard's wear equation [100], one can calculate the wear volume (V) with the normal load (L), the sliding distance (x), and the hardness (H) of the AFM tip.

$$V = k \frac{Lx}{H}$$

where k is the wear coefficient that implies the probability of debris particle formation. The wear of the AFM tips increases with the normal load and the sliding distance. The harder the AFM tip, the less it wears.

The removal of surface materials is a rather complicated process in nanomanufacturing. Shiari et al. developed a multiscale numerical simulation approach on single crystal aluminum to further advance the understanding of material removal mechanisms at the nanometer length scale [101]. Nanoabrasive wear is expected to be a dominant factor in a contact-sliding lithographic process, such as the scanning probe lithography. The interactions between the sharp tip of a SPM and the sample surface are responsible for nanoscale patterning in the scanning probe lithography. Sung et al. have investigated tip–surface interactions in the nanolithography process through both experiments and molecular dynamics simulation [102]. Nanoprecision sensing in microelectromechanical system is critical. Narayanan et al. have established a model to achieve this goal in manufacturing low-friction and low-wear parts that have nanoscale resolution of the position control [103].

Wear plays important roles in our daily life and has many major industrial applications [104–106]. Since its invention, the atomic force microscopy (AFM) has been used for nanoscale studies of material's surface morphology and tribological properties. Nanotribology has then generated great interests in the broad fields of research and applications [107–112]. In polysilicon microelectromechanical systems, contact within small areas requires highly scratch-wear-resistant ultrathin films. This study was made possible by using an AFM [113, 114]. Using an AFM, nanowear experiments on polymer films were conducted successfully [115, 116]. Reports have been found that a specifically processed probe played a crucial role in contemporary AFM-based recording techniques. Among those, Bhushan et al. investigated the wear mechanism between platinum-coated probes and recording medium [117]; and Chung et al. studied tribological properties of a metal-doped silicon probe on recording materials [118]. Nanoscale study of tribological properties of AFM probes remains to be important.

Nanoabrasive wear was observed on stainless steel induced by a diamond AFM probe [119]. Such wear is simply downsizing of the conventional wear found at the macroscale. In situ techniques have been applied to nanoscale wear, and it was found that the abrasive wear was initiated by a series of localized onset activities [120]. On the contrary, materials can be transferred by adhesive wear [121]. Friction, in addition, can induce nanoscale crystallization [122]. To date, the roles of materials and fundamental mechanisms of wear need to be further understood, which is possible with emerging techniques and methodologies.

23.2 Nanoalloying and Nanocrystallization

23.2.1 Background

In most nanomanufacturing processes, nanoalloying and nanocrystallization are commonly involved when materials are in contact, that is where an interface is formed. Special attention has been paid in previous sections to SPM-based nanolithography, in which an interface between the sharp probe and the sample surface is formed during writing. In such type of applications, two materials are involved. Here we focus on binary materials systems. In the current section, we firstly define nanoalloying and nanocrystallizations, and then review the methods generating binary nanoalloys. Finally, recent work on mechanical stimulation-induced nanoalloying and nanocrystallization is discussed.

According to the Merriam-Webster Dictionary, alloy is defined as “a substance composed of two or more metals or of a metal and a nonmetal intimately united usually by being fused together and dissolving in each other when molten” [123]. Nanoalloying literally means the process through which two elements are mixed in order to form an alloy at the atomic or nanometer scale. During the nanoalloying process, solid crystals can be generated from a uniform solution, melt or a gas phase, under the control of certain external conditions. We define this solid crystal formation as nanocrystallization.

When two different types of elements (metal or nonmetal) are mixed, under external mechanical and thermal stimulations, nanoalloys can be synthesized atom by atom, molecule by molecule, and cluster by cluster [124]. These nanoalloys have shown enhanced mechanical, catalytical, optical, electronic, and magnetic properties [2–125, 125, 125–161]. Chen et al. discovered a structure of $\text{Ag}_3\text{Au}_{10}$ with different optical properties from pure Au_{13} and clusters [129]. The same group has reported other gold–silver nanoalloy structures such as $\text{Au}_{44}\text{Ag}_{44}$ [130], which shows high structural, thermal, and electronic stability, and a core-shell structure Ag–Au alloy, which has martensitic transformations [131].

23.2.2 Synthesis of Nanoalloys

Nanoalloys can be generated through various compositions of basic elements. Nanocrystallizations will occur under certain treatment such as annealing. There are mainly six different means to synthesize nanoalloys. Recent reports on bimetallic nanoalloys synthesized by these approaches are presented here:

- Thermolysis. Nazir et al. synthesized iron (Fe)–palladium (Pd) bimetallic nanoalloys by thermal decomposition of Fe and Pd salts at low temperature. The nanoalloys appear to be superparamagnetic in nature. [132]

- Energetic beams methods. Abdelsayed et al. have reported the laser vaporization controlled condensation method to obtain several gold (Au) nanoalloys such as gold–silver, gold–platinum (Pt), and gold–palladium (Pd). During formation, each alloyed nanoparticle corresponded to a selective laser wavelength. The bimetallic alloys synthesized by this approach were found to be superparamagnetic [133].
- Chemical reduction. Kim et al. prepared a tin (Sn)–cobalt (Co) nanoalloy ($\text{Sn}_{87}\text{Co}_{13}$) by two different chemical reduction agents. The products showed improved electrochemical properties such as good capacity retention. The average nanoalloy particle size can be easily controlled at less than 15 nm [134].
- Electrochemical approach. Kim et al. also synthesized a tin–cobalt nanoalloy ($\text{Sn}_{92}\text{Co}_8$) through the electrochemical approach in different solutions, followed by annealing. The particle size could be controlled as small as 10 nm [135].
- Radiolysis. Chen et al. generated a tin–bismuth (Bi) nanoalloy through a novel sonochemical approach. This type of nanoalloy powders showed great antiwear property [136].
- Biosynthesis. Senapati et al. obtained a gold–silver nanoalloy through a biosynthesis approach. The advantage of biosynthesis was that the nanoparticles were entrapped within the cell matrix [137].

23.3 Probe-Based Nanoalloying and Nanocrystalizations

23.3.1 Background

Mechanical stimulation has been found to generate nanostructures. As early as in the late 1960s, scientists were able to use electron microscopes to observe smoothed metal surfaces of packed and reorganized atoms due to mechanical polishing [138]. Fundamental study has shown that mechanical stimulation (e.g., friction) triggers chemical reactions that are different from equilibrium chemical reactions. For example, when materials are under friction, they can be oxidized and dissolved by the interaction of frictional stresses and chemical attack in water [139]. The feasibility of friction-stimulated nanostructure growth has been demonstrated on brushite with probe scanning on AFM [140]. Indium has been found nucleated, and then grew into triangular-shaped nanostructures on a silicon substrate due to friction. The interfacial junction between the nanostructures and the substrate is strongly due to the eutectic reactions. The metal–silicon interfaces are of great salience to materials scientists and engineers. The interfacial phenomena involve a nondiffusive interface of silver and silicon (Ag/Si) [141,142], and a diffusive one of gold and silicon (Au/Si), without a stable compound [143].

Mechanical forces present in wide industrial applications. During polishing, surface atoms repack and reorganize under stress leading to a smooth surface.

[144] Friction force reportedly triggers chemical reactions that are different from equilibrium processes. [145] With a mechanical force, materials can be oxidized and dissolved in water [146–149] and form a boundary layer in oil as lubricants. [69, 150, 151] In addition, mechanical stress introduces deformation, fracture, and wear, resulting in changes in a large number of nucleus positions and therefore the physical motion of an atom. [152] Being able to control surface forces precisely allows the mentioned damages to be significantly reduced or eliminated.

It has been established that silicon alloyed with gold forms a eutectic alloy that can be used as a low melting point solder for silicon devices [153, 154]. During soldering, it is desirable for silicon to have a high diffusion rate inside the gold [143, 155]. According to the equilibrium phase diagram (shown in the supporting materials), the gold (Au) and silicon (Si) system forms a eutectic with a small quantity of about 5%wt Si at 363°C, existing as a perfectly mixed liquid over a fairly broad temperature and composition range [156]. Equilibrium eutectic systems have been used to fabricate nanorods and nanowires at a relatively low-temperature through vapor–liquid–solid synthesis method [157–160]. Among those, metastable phases of AuSi and Au₇Si appeared with different structural and electronic properties.

23.3.2 Scanning Probe-Based Alloying Nanolithography

This technique utilizes an AFM as a writing tool, an AFM probe coated by a thin metal layer as a “nanopen,” to perform raster sliding on a single crystal silicon substrate. The process performed here is a directly nanoscale writing. The precisely control of AFM scanning is the critical part of the scanning probe alloying nanolithography.

AFM Functionality as a Processing Tool

The main components of this microscope are: a 10× Mitutoyo objective, a scanner, a tripod stage and a sample holder (Fig. 23.2). Each of them plays an important role in the writing process development.

The 10× objective plus a CCD camera are capable of producing a resolution as small as 1 μm and display the digitized image on a television screen.

The scanner is the most critical part of the microscope. In Nano-R2’s design, the AFM probe physically scans over a fixed sample. The built-in X and Y piezoelectric sensors control tip movement in the horizontal plane. The maximum displacement along one direction is 85 μm.

The motorized tripod stage contains a separate Z-piezoelectric sensor, which offers fast and precise manipulation of the scanner’s vertical movement. The maximum vertical displacement is 22 mm, and the smallest step of vertical movement is 160 nm.

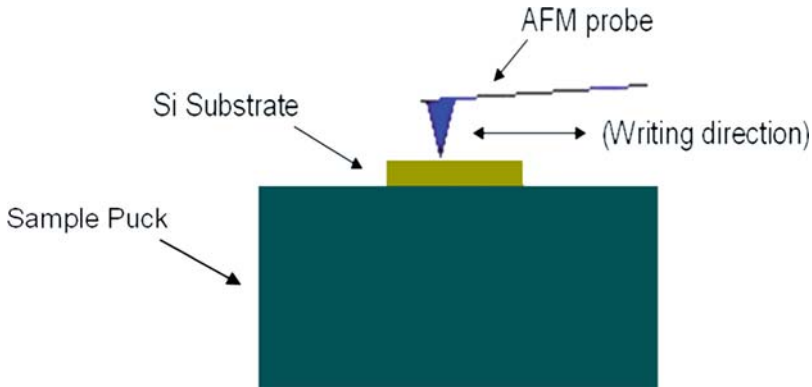


Fig. 23.2. AFM setup for room temperature writing

The sample holder is made of five cylindrical metal pieces. Each piece can be easily inserted or pulled out, to achieve the appropriate height for different specimens. There is an automated X-Y translation stage beneath the sample holder that moves the sample holder below the AFM probe at certain steps (from 1 to 25,000 μm).

With the design of Nano-R2, one can quickly locate the point of interest on the sample and perform a robust scan.

Basic AFM Setup for Nanoprocessing

The basic setup can be found in Fig. 23.2. We started the writing experiment with a standalone AFM at room temperature. This is a process called scanning probe alloying nanolithography.

Interfacial Interactions Between Tip and Substrate

In this section, we use a nanoscale approach to study the wear behavior of eutectic material systems. Simple sliding experiments were conducted using Au-coated AFM probes to slide on a single crystal silicon (Si) wafer. The extreme dimensions of the AFM probes were observed using high-resolution techniques. We found that nanowear resulted from mixed modes of abrasion and adhesion. We propose that with the addition of frictional heating, it is possible to observe morphological change due to eutectic reaction. The present research contributes to fundamental understanding of nanoscale wear mechanisms.

Mechanical Sliding

A silicon nitride (Si_3N_4) AFM probe (Contact Mode probes from Veeco) was used in this study. The probe was coated with an adhesion layer of 5-nm-thick chromium (Cr) and subsequently, a 145-nm-thick gold (Au) layer. The film

was deposited through vacuum evaporation. A single crystalline silicon (100) wafer was used as the substrate, and an AFM (AFM, Nano-R2, PNI) was used for sliding experiments as well as characterization.

The sliding experiment was processed in a contact mode. During the probe's approach, there were two sequential steps. The probe was first brought into physical contact with the sample surface in the approaching mode (controlled by the software), with no stress applied on the cantilever. The probe was then forced down with the cantilever bent. The amount of bending can be controlled through the software with the number of down steps. A total of three steps were applied and each step was $5\ \mu\text{m}$ down. The sliding speed was set at $480\ \mu\text{m/s}$, with the test duration at 0, 5, 10, 20, 40, and 100 strokes. At each point, the probe was characterized using a field emission scanning electron microscope (Zeiss1350). Chemical composition at the probe tip was analyzed using energy dispersion spectroscopy. Micro- and nanohardness tests were conducted using a Vickers hardness tester and a nanoindenter (Hysitron). The morphological analysis of the wear track on the silicon substrate was carried out using the same AFM with a different probe.

Morphology of AFM Tips

The Au-coated probe slid against the silicon. The applied load (N) was 504 nN and the sliding speed was $480\ \mu\text{m/s}$ [84]. An FE-SEM was used to observe the wear scar at fixed sliding distances (0, 5, 10, 20, 40, and 100 strokes). Results are shown in Fig. 23.3, which is the top view of the Au-coated probes before and after sliding; in this figure, the center plateau is the wear scar of coated probes indicating abrasive wear.

To quantitatively study the abrasive wear behavior, we defined the diameter of the wear scar (D) as the average diagonal length of the observed rectangle. The diameter of the wear scar versus the number of strokes was plotted (Fig. 23.4). The error bars reflect the systematic uncertainty. As seen in the figure, the final scar diameter reaches the size of 400 nm, which is much greater than the original tip diameter (less than 100 nm). This indicates the material's removal due to mechanical sliding.

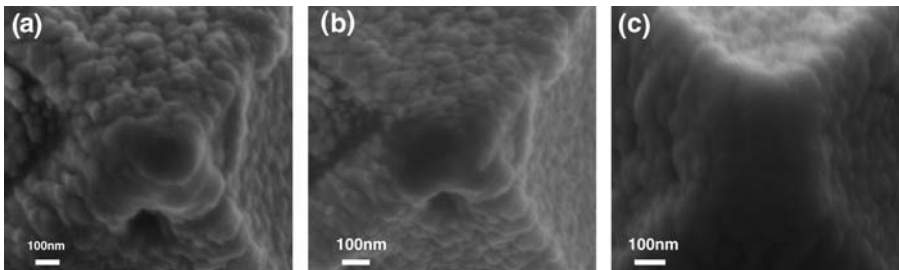


Fig. 23.3. SEM images of AFM probes. (a) Au-coated probe before sliding, (b) same probe after sliding 5 strokes, and (c) same probe after 100 strokes

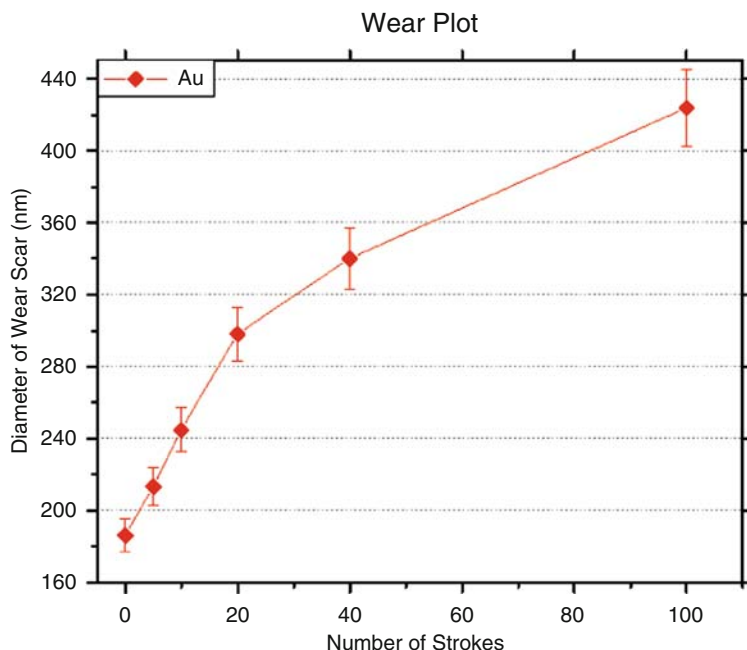


Fig. 23.4. Plot of diameter of wear scar versus # of strokes for Au-coated probe. The probe is showing abrasive wear

Table 23.3. Au weight (%) and atomic (%) on probe before and after sliding

	Element	Wt (%)		Element	Atomic (%)
Before sliding	Au	89.32	Before sliding	Au	54.39
	Si	10.68		Si	45.61
After sliding	Au	66.59	After sliding	Au	22.13
	Si	33.41		Si	77.87

Chemical Analysis

Energy dispersive x-ray spectroscopy analysis was conducted on both unused and worn probes. The areas of detection were exactly the same on each tip. The accelerating voltage was 20 kV. The data is shown in Table 23.3. Here the silicon weight percentage within the interested area significantly changed, both before and after sliding, by 22.73%: silicon was transferred from the substrate surface onto the AFM tip due to sliding.

Morphological Analysis of “Wear” Tracks

The morphology of the silicon substrates was characterized using the same AFM with a different probe. Results are shown in Fig. 23.5. A well-defined line is seen here, instead of a wear groove. This structure was seen in both the Au- and Ag-wear tracks as grown lines. The width of the line is 240 and 170 nm, respectively, which is correlated with the wear scar diameter of the AFM probe.

To investigate whether the hardness of the three materials is responsible for the nanowear, we conducted microhardness test on bare silicon wafers and both Au and Ag thin films. The results (data now shown) indicated that silicon has the highest hardness whereas gold has the lowest ($H_{\text{Si}} > H_{\text{Ag}} > H_{\text{Au}}$). It is possible that Au deforms more than the other two materials. To see the temperature effects, we continued the experiments at an elevated temperature, 50°C. Figure 23.6 shows the results. A well-defined line was formed due to sliding that has a width of 600 nm and height of 8 nm. How were the line

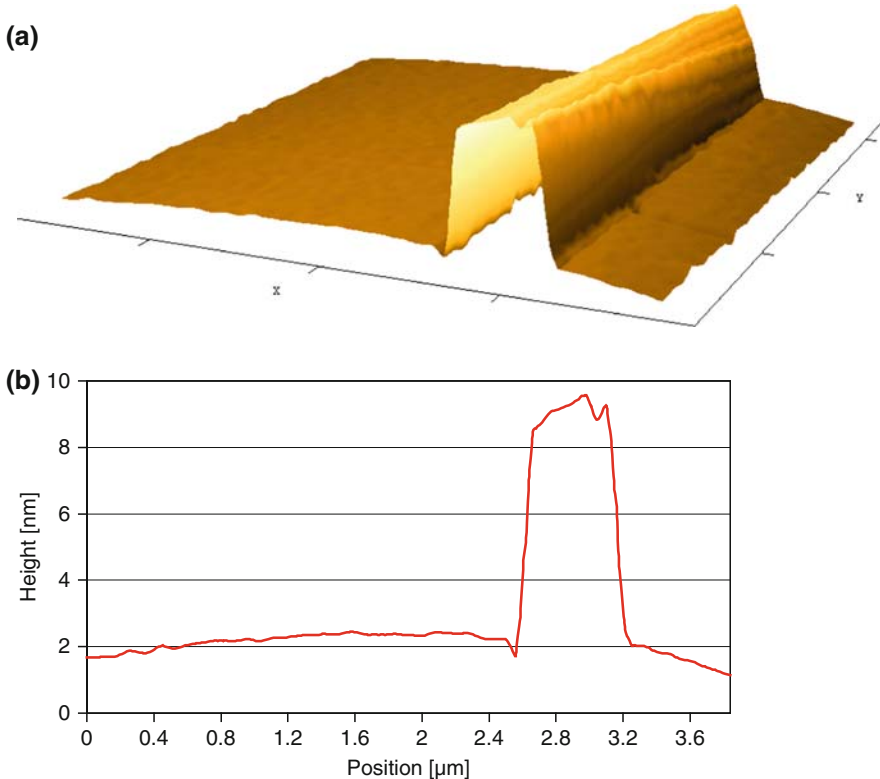


Fig. 23.5. A structure formed by sliding the Au-coated probe on Si substrate. The height and width are 8 and 600 nm, respectively. (a) three-dimensional AFM image and (b) the profile of the structure

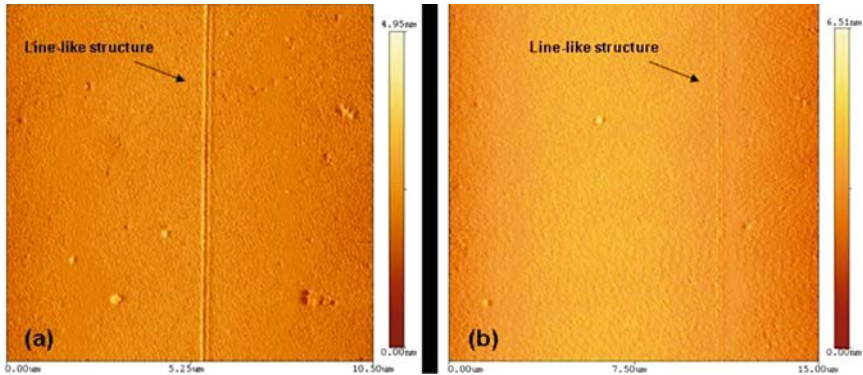


Fig. 23.6. AFM height images of silicon surface after tests. A line-like structure was found along the track of sliding in both cases. (a) Processed by Au-coated tip. The line is $60\ \mu\text{m}$ long, average height and width are 0.9 and $240\ \text{nm}$, respectively. (b) Processed by Ag-coated tip. The whole line is $60\ \mu\text{m}$ long, average height and width are 0.4 and $170\ \text{nm}$, respectively

Table 23.4. Volume comparison between wear volume on the probe V_w and volume of the line structure V_l (one can see $V_l > V_w$)

	$V_w\ (\text{nm}^3)$	$V_l\ (\text{nm}^3)$
Au-Si	1.2×10^7	1.3×10^7

structure and the Si substrate bonded together? Under an in situ experiment inside a TEM, we used an Au-coated diamond indenter to indent the Si substrate. Details about the experiments have been reported elsewhere [161, 162]. After a brief contact, the Au was seen impregnated into the Si (not shown). It's noticed that this was done in a vacuum chamber and the surface was etched before the experiment. This means that mechanical locking could be responsible to bond the Au and Si together. In presence of oxygen, the bonding would be stronger.

We define the tip wear volume as V_w , the volume of the line structure as V_l . Here V_w can be estimated from the diameter of the wear scar (D) from the AFM tip, and V_l can be estimated from the dimensions of the line structure. It is shown in Table 23.4 that V_l is greater than V_w . The volume comparison indicates that the line is indeed formed from mixed Au-Si.

In the EDS analysis, silicon was found to adhere onto the Au tip during sliding. Well-defined line structures were found on the substrate in addition to the groove. All the facts suggest that in the highly localized area, the stress induced nanowelding, such that the tip material was transferred onto the Si substrate. The wear mechanism of an Au-Si system at the nanometer scale was found to be a combination of abrasive wear and localized adhesive (transfer) wear.

Acknowledgements

This work was sponsored by the National Science Foundation (grant number 0506082).

References

1. P.B. Meggs, *A History of Graphic Design* (Wiley, New York, 1998).
2. H.J. Levinson, *Principles of Lithography*, 2nd ed. (SPIE, Bellingham, 2005).
3. J.R. Maldonado, *J. Electron. Mater.* **19**, 699 (1990).
4. L.A. Giannuzzi, F.A. Stevie, *Introduction to Focused Ion Beams: Instrumentation, Theory, Techniques and Practice* (Springer, New York, 2005).
5. J.I. Pascual, J. Mendez, J. Gomezherrero, A.M. Baro, N. Garcia, V.T. Binh, *Phys. Rev. Lett.* **71**, 1852 (1993).
6. N.M. Miskovsky, T.T. Tsong, *Phys. Rev. B* **46**, 2640 (1992).
7. G.S. Hsiao, R.M. Penner, J. Kingsley, *Appl. Phys. Lett.* **64**, 1350 (1994).
8. D.H. Huang, T. Nakayama, M. Aono, *Appl. Phys. Lett.* **73**, 3360 (1998).
9. A. Laracuent, M.J. Bronikowski, A. Gallagher, *Appl. Surf. Sci.* **107**, 11 (1996).
10. R.D. Allen, P.J. Brock, L. Sundberg, C.E. Larson, G.M. Wallraff, W.D. Hinsberg, J. Meute, T. Shimokawa, T. Chiba, M. Slezak, *J. Photopolymer Sci. Technol.* **18**, 615 (2005).
11. R.L. Brainard, G.G. Barclay, E.H. Anderson, L.E. Ocola, *Microelectron. Engineer.* **61–2**, 707 (2002).
12. B. Fay, *Microelectron. Engineer.* **61–2**, 11 (2002).
13. Y. Hirai, Y. Kanemaki, K. Murata, Y. Tanaka, *Jpn. J. Appl. Phys.* **138**, 7272 (1999).
14. P. Ruchhoeft, M. Colburn, B. Choi, H. Nounu, S. Johnson, T. Bailey, S. Damle, M. Stewart, J. Ekerdt, S.V. Sreenivasan, J.C. Wolfe, C.G. Willson, *J. Vac. Sci. Technol. B* **17**, 2965 (1999).
15. M. Miwa, S. Juodkazis, T. Kawakami, S. Matsuo, H. Misawa, *Appl. Phys. A-Mater. Sci. Process.* **73**, 561 (2001).
16. Y.A. Vlasov, X.Z. Bo, J.C. Sturm, D.J. Norris, *Nature* **414**, 289 (2001).
17. H.K. Wu, T.W. Odom, D.T. Chiu, G.M. Whitesides, *J. Am. Chem. Soc.* **125**, 554 (2003).
18. Mack C.A., *Microolitho. World* **7**, 23 (1998).
19. Fukuda H., Yamanaka R., Terasawa T., Hama K., Tawa T., Okazaki S. (IEEE, San Francisco, CA, USA, 1992), pp. 49.
20. Broers A.N. (Academic, College Station, TX, USA, 1989), pp. 421.
21. C. Vieu, F. Carcenac, A. Pepin, Y. Chen, M. Mejias, A. Lebib, Manin-L. Ferlazzo, L. Couraud, Launois H. (Elsevier, Ile de Porquerolles, France, 2000), **164**, 111.
22. A.N. Broers, *IBM J. Res. Develop.* **32**, 502 (1988).
23. A.N. Broers, *Japanese J. Appl. Phys, Kobe, Japan* **43** (1989).
24. A.N. Broers, *IEEE, Washington, DC, USA* **2** (1980).
25. A.N. Broers, *IEEE Transac. Electron Devices ED-28*, 1268 (1981).
26. C.W. Gwyn, R. Stulen, D. Sweeney, D. Attwood, *in Papers from the 42nd international conference on electron, ion, and photon beam technology and nanofabrication* (AVS, Chicago, Illinois, USA, 1998), **16**, 3142.

27. R. Hirose, San Jose, CA, USA, **1088**, 178 (1989).
28. K. Eguchi, S. Miyazaki, C. Takai, T. Suganuma, Santa Clara, CA, USA, **922**, 335 (1988).
29. J. Przybyla, T. Emery, H. Mukaled, San Jose, CA, USA, **1927**, 794 (1993).
30. M. Ohta, T. Kojima, C. Sato, T. Ogawa, M. Noguchi, Santa Clara, CA, USA, **922**, 291 (1988).
31. W.H. Ostrout, Mark W. Hiatt, A.E. Kozlowski, *Microelectron. Manufact. Technol.* **14**, 16 (1991).
32. J.R. Sheats, *Polymer Engineer. Sci.* **29**, 965 (1989).
33. J.D. Buckley, *Solid State Technol.* **30**, 87 (1987).
34. G. Degiorgis, P. Pateri, A. Pilenga, R.J. Hurditch, B.T. Beauchemin, Jr., E.A. Fitzgerald, San Jose, CA, USA, **1262**, 368 (1990).
35. S.J. Holmes, P.H. Mitchell, M.C. Hakey, *IBM J. Res. Develop.* **41**, 7 (1997).
36. W. Zhang, S.Y. Chou, *Appl. Phys. Lett.* **79**, 845 (2001).
37. H. Sakou, T. Miyatake, S. Kashioka, M. Ejiri, *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing]*, *IEEE Transac.* **37**, 2148 (1989).
38. O. Wilhelm, S.S. Peredkov, A.L. Bogdanov, *Microelectron. Engineer.* **61–62**, 1107 (2002).
39. Cheol- K. Kyun, H. Cheol, Young- K. Sik, Ki- B. Ho, *SPIE-Int. Soc. Opt. Eng.*, Santa Clara, CA, USA, **3676**, 528 (1999).
40. N. Bobroff, A.E. Rosenbluth, *Opt. Soc. America*, Monterey, CA, USA **42** (1991).
41. T. Nakasugi, A. Ando, K. Sugihara, Y. Yamazaki, M. Miyoshi, K. Okumura, *AIP for American Vacuum Soc*, Washington, DC, USA, **19**, 2869 (2001).
42. H.Y. Wang, Z.H. Wu, *Semicond. Technol.* **31**, 576 (2006).
43. L. Wang, Y. Ding, B. Lu, Z. Qiu, H. Liu, *Int. Soc. Opt. Engineer. Bellingham WA, WA 98227–0010, United States*, Xian, China, **6149**, 61491 (2006).
44. D.P. Stumbo, G.A. Damm, D.W. Engler, F.O. Fong, S. Sen, J.C. Wolfe, J.N. Randall, P. Mauger, A. Shimkunas, H. Loschner, San Jose, CA, USA, **1263**, 35 (1990).
45. L.B. Zhang, J.X. Shi, J.L. Yuan, S.M. Ji, M. Chang, *Adv. Mater. Manufact. Sci Technol.* **471–472**, 353 (2004).
46. Y. Takemura, J.I. Shirakashi, *J. Magnet. Mag. Mat.* **304**, 19 (2006).
47. H. Sugimura, N. Nakagiri, *Nanotechnology* **8**, A15 (1997).
48. Z.J. Davis, G. Abadal, O. Hansen, X. Borise, N. Barniol, Perez- F. Murano, A. Boisen, *Ultramicroscopy* **97**, 467 (2003).
49. S. Rozhok, P. Sun, R. Piner, M. Lieberman, C.A. Mirkin, *J. Phys. Chem. B* **108**, 7814 (2004).
50. R.L. Lo, W.C. Lee, J. Kwo, *Japanese J. Appl. Phys. Part 1-Regular Papers Brief Communications & Review Papers* **45**, 2067 (2006).
51. A. Notargiacomo, V. Foglietti, E. Cianci, G. Capellini, M. Adami, P. Faraci, F. Evangelisti, C. Nicolini, *Nanotechnology* **10**, 458 (1999).
52. Fernandez- I. Cuesta, X. Borrise, Perez- F. Murano, *Nanotechnology* **16**, 2731 (2005).
53. L. Santinacci, T. Djenizian, P. Schmuki, *J. Electrochem. Soc.* **148**, C640 (2001).
54. B.W. Maynor, Y. Li, J. Liu, *Langmuir* **17**, 2575 (2001).
55. J.W. Park, D.W. Lee, N. Takano, N. Morita, *Progress Adv. Manufact. Micro/Nano Technol.* 2005, Pt 1 and 2, **505–507**, 79 (2006).

56. G. Agarwal, L.A. Sowards, R.R. Naik, M.O. Stone, *J. Am. Chem. Soc.* **125**, 580 (2003).
57. R.D. Piner, J. Zhu, F. Xu, S.H. Hong, C.A. Mirkin, *Science* **283**, 661 (1999).
58. A. Ivanisevic, C.A. Mirkin, *J. Am. Chem. Soc.* **123**, 7887 (2001).
59. S. Kuwahara, S. Akita, M. Shirakihara, T. Sugai, Y. Nakayama, H. Shinohara, *Chem. Phys. Lett.* **429**, 581 (2006).
60. Van C. Haesendonck, L. Stockman, G. Neuttiens, C. Strunk, Y. Bruynseraede, Denver, CO, USA, **13**, 1290 (1995).
61. Van C. Haesendonck, L. Stockman, Y. Bruynseraede, L. Langer, V. Bayot, J.P. Issi, J.P. Heremans, C.H. Olk, *Physica Scripta T* **T55**, 86 (1994).
62. D. Wouters, U.S. Schubert, *Angewandte Chemie-International Edition* **43**, 2480 (2004).
63. R. Held, T. Heinzel, P. Studerus, K. Ensslin, *Physica E* **2**, 748 (1998).
64. B. Irmer, M. Kehrlé, H. Lorenz, J.P. Kotthaus, *Semicond. Sci. Technol.* **13**, A79–A82 (1998).
65. R.D. Ramsier, R.M. Ralich, S.F. Lyuksyutov, *Appl. Phys. Lett.* **79**, 2820 (2001).
66. Rubio- F.J. Sierra, W.M. Heckl, R.W. Stark, *Adv. Engineer. Mater.* **7**, 193 (2005).
67. W.C. Moon, T. Yoshinobu, H. Iwasaki, *Japanese J. Appl. Phys. Part 1-Regular Papers Short Notes & Review Papers* **38**, 6952 (1999).
68. X.J. Duan, J. Zhang, X. Ling, Z.F. Liu, *J. Am. Chem. Soc.* **127**, 8268 (2005).
69. T. Zhenhua, B. Bharat, *Rev. Scientific Instrum.* **77**, 103705 (2006).
70. K. Wilder, C.F. Quate, D. Adderton, R. Bernstein, V. Elings, *Appl. Phys. Lett.* **73**, 2527 (1998).
71. H. Sugimura, N. Nakagiri, IOP Publishing, Tsukuba, Japan, **8**, 15 (1997).
72. S.C. Wimbush, M. Tachiki, Takayama- E. Muromachi, H. Itozaki, *Japanese J. Appl. Phys., Part 1: Regular Papers and Short Notes and Review Papers* **45**, 5742 (2006).
73. X. Tian, N. Jiao, L. Liu, Y. Wang, Z. Dong, N. Xi, W. Li, Institute of Electrical and Electronics Engineers Inc., New York, NY 10016–5997, United States, Chengdu, China, pp. 18–22. (2004).
74. T. Qian, San- S. Qiang, Z. Limin *J. Nanosci. Nanotechnol.* **4**, 948 (2004).
75. H. Sugihara, A. Takahara, T. Kajiyama, *J. Vac. Sci. Technol. B (Microelect. Nanometer. Struct.)* **19**, 593 (2001).
76. M. Kato, M. Ishibashi, S. Heike, T. Hashizume, *Japan Soc. Appl. Phys, Atagawa Haitsu, Japan*, **40**, 4317 (2001).
77. R.E. Ricker, A.E. Miller, D.F. Yue, G. Banerjee, S. Bandyopadhyay, TMS, Anaheim, CA, USA, **25**, 1585 (1996).
78. S. Sasa, T. Ikeda, C. Dohno, M. Inoue, *Japanese J. Appl. Phys, Kanazawa, Japan*, **36**, 4065 (1997).
79. G. Binnig, C.F. Quate, C. Gerber, *Physical Rev. Lett.* **56**, 930 (1986).
80. K. Salaita, S.W. Lee, X.F. Wang, L. Huang, T.M. Dellinger, C. Liu, C.A. Mirkin, *Small* **1**, 940 (2005).
81. D. Bullen, S.W. Chung, X.F. Wang, J. Zou, C.A. Mirkin, C. Liu, *Appl. Phys. Lett.* **84**, 789 (2004).
82. J. Haaheim, R. Eby, M. Nelson, J. Fragala, B. Rosner, H. Zhang, G. Athas, *Ultramicroscopy* **103**, 117 (2005).
83. R.D. Piner, J. Zhu, F. Xu, S.H. Hong, C.A. Mirkin, *Science* **283**, 661 (1999).

84. S.H. Hong, J. Zhu, C.A. Mirkin, *Science* **286**, 523 (1999).
85. S.H. Hong, C.A. Mirkin, *Science* **288**, 1808 (2000).
86. S.W. Lee, R.G. Sanedrin, B.K. Oh, C.A. Mirkin, *Adv. Mater.* **17**, 2749 (2005).
87. L. Fu, X.G. Liu, Y. Zhang, B.P. Dravid, C.A. Mirkin, *Nano Letters* **3**, 757 (2003).
88. J.M. Nam, S.W. Han, K.B. Lee, X.G. Liu, M.A. Ratner, C.A. Mirkin, *Angewandte Chemie-International Edition* **43**, 1246 (2004).
89. J. Jang, G.C. Schatz, M.A. Ratner, *Phys. Rev. Lett.* **90**, 156104 (2003).
90. P.E. Sheehan, L.J. Whitman, *Phys. Rev. Lett.* **88**, 156104 (2002).
91. M. Tortonese, *Engineer. Med. Biol. Magaz., IEEE* **16**, 28 (1997).
92. T.R. Albrecht, S. Akamine, T.E. Carver, C.F. Quate, *J. Vac. Sci. Technol. A: Vac., Surfaces, Films* **8**, 3386 (1990).
93. B. Bhushan, J.N. Israelachvili, U. Landman, *Nature* **374**, 607 (1995).
94. B. Bhushan, V.N. Koinkar, 38th Annual Conference on Magnetism and Magnetic Materials (AIP, Minneapolis, Minnesota, USA, 1994) **75**, 5741.
95. B. Bhushan, J.A. Ruan, J. Tribol., *Transac. ASME* **116**, 389 (1994).
96. T. Miyamoto, R. Kaneko, S. Miyake, Fifth international conference on scanning tunneling microscopy spectroscopy (AVS, Boston, Massachusetts, USA, 1991) **9**, 1336.
97. J.A. Ruan, B. Bhushan, *Transactions of the ASME. J. Tribol.* **116**, 378 (1994).
98. S. Venkatesan, B. Bhushan, *Wear* **171**, 25 (1994).
99. K.H. Chung, Y.H. Lee, D.E. Kim, *Ultramicroscopy* **102**, 161 (2005).
100. J.F. Archard, *J. Appl. Phys.* **24**, 981 (1953).
101. B. Shiari, R.E. Miller, D.D. Klug, *J. Mechan. Phys. Sol.* **55**, 2384 (2007).
102. I.H. Sung, D.U. Kim, *Ultramicroscopy* **107**, 1 (2007).
103. N. Ramakrishnan, P.B. Chu, E.C. Johns, Y. Zhao, X. Yan, *Sensors and Actuators A: Physical* **145–146**, 214 (2008).
104. K. Hokkirigawa, K. Kato, *Tribol. Int.* **21**, 51 (1988).
105. J.F. Archard, *J. Appl. Phys.* **24**, 981 (1953).
106. J.B. Adams, L.G. Hector, D.J. Siegel, H.L. Yu, J. Zhong, *Surface and Interface Anal.* **31**, 619 (2001).
107. F. Iwata, Y. Suzuki, Y. Moriki, S. Koike, A. Sasaki, *J. Vac. Sci. Technol. B* **19**, 666 (2001).
108. W.S. Kim, J.K. Kim, P. Hwang, *J. Electron. Mater.* **30**, 503 (2001).
109. W. Lu, K. Komvopoulos, *J. Tribol.-Transac. ASME* **123**, 717 (2001).
110. T. Tokai, S. Umemura, S. Hirono, A. Imoto, R. Kaneko, *Surface & Coatings Technol.* **169**, 475 (2003).
111. S. Umemura, S. Hirono, Y. Andoh, R. Kaneko, *J. Japan Inst. Metals* **67**, 286 (2003).
112. B. Bhushan, *Wear* **250**, 1105 (2001).
113. M.W. Bai, K. Kato, N. Umehara, Y. Miyake, J.G. Xu, H. Tokisue, *Surface & Coatings Technol.* **126**, 181 (2000).
114. M.A. Baker, J. Li, *Surface and Interface Anal.* **38**, 863 (2006).
115. R. Berger, Y. Cheng, R. Forch, B. Gotsmann, J.S. Gutmann, T. Pakula, U. Rietzler, W. Scharl, M. Schmidt, A. Strack, J. Windeln, H.J. Butt, *Langmuir* **23**, 3150 (2007).
116. A. Ghorbal, S. Bistac, M. Schmitt, *J. Polymer Sci. Part B-Polymer Phys.* **44**, 2449 (2006).
117. B. Bhushan, K.J. Kwak, *Nanotechnology* **18**, 345504 (2007).

118. K.H. Chung, Y.H. Lee, D.E. Kim, J. Yoo, S. Hong, *IEEE Transact. Magnet.* **41**, 849 (2005).
119. K. Degiampietro, R. Colaco, *Wear* **263**, 1579 (2007).
120. R. Ribeiro, Z. Shan, A.M. Minor, H. Liang, *Wear* **263**, 1556 (2007).
121. T. Kayaba, K. Kato, *Asle Transactions* **24**, 164 (1981).
122. S. Guruzu, M. Kulkarni, S. Ingole, G. Xu, C. Chen, H. Liang, *Wear* **259**, 524 (2005).
123. Merriam-Webster Dictionary (2002).
124. R. Ferrando, J. Jellinek, R.L. Johnston, *Chem. Rev.* **108**, 845 (2008).
125. F. Congiu, G. Concas, G. Ennas, A. Falqui, D. Fiorani, G. Marongiu, S. Marras, G. Spano, A.M. Testa, *J. Magnet. Magnetic Mater.* **272–276**, 1561 (2004).
126. Rodriguez- B. Gonzalez, A. Burrows, M. Watanabe, C.J. Kiely, L.M.L. Marzan, *J. Mater. Chem.* **15**, 1755 (2005).
127. P. Mulvaney, *Langmuir* **12**, 788 (1996).
128. S.C. Tjong, H. Chen, *Mater. Sci. Engineer. R-Reports* **45**, 1 (2004).
129. F.Y. Chen, R.L. Johnston, *Appl. Phys. Lett.* **90** (2007).
130. F.Y. Chen, R.L. Johnston, *ACS Nano* **2**, 165 (2008).
131. F.Y. Chen, R.L. Johnston, *Appl. Phys. Lett.* **92**, 023112 (2008).
132. R. Nazir, M. Mazhar, M.J. Akhtar, M.R. Shah, N.A. Khan, M. Nadeem, M. Siddique, M. Mehmood, N.M. Butt, *Nanotechnology* **19**, 185608 (2008).
133. V. Abdelsayed, El- M.S. Shall, *J. Chem. Phys.* **126**, 024706 (2007).
134. H. Kim, J.P. Cho, *Electrochimica Acta* **52**, 4197 (2007).
135. H. Kim, J.P. Cho, *J. Electrochem. Soc.* **154**, A462 (2007).
136. H.J. Chen, Z.W. Li, Z.S. Wu, Z.J. Zhang, *J. Alloys Comp.* **394**, 282 (2005).
137. S. Senapati, A. Ahmad, M.I. Khan, M. Sastry, R. Kumar, *Small* **1**, 517 (2005).
138. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids* (Clarendon, Oxford, 1958).
139. J.M. Martin, *Tribol. Lett.* **6**, 1 (1999).
140. J.A. Stroschio, D.M. Eigler, *Science* **254**, 1319 (1991).
141. A. Hiraki, A. Shimizu, M. Iwami, T. Narusawa, S. Komiya, *Appl. Phys. Lett.* **26**, 57 (1975).
142. K. Nakashima, M. Iwami, A. Hiraki, *Thin Sol. Films* **25**, 423 (1975).
143. A. Cros, P. Muret, *Mater. Sci. Rep.* **8**, 271 (1992).
144. E. Rabinowi, *Sci. Am.* **218**, 91 (1968).
145. T.E. Fischer, W.M. Mullins, *J. Phys. Chem.-Us* **96**, 5690 (1992).
146. T.E. Fischer, H. Tomizawa, *Wear* **105**, 29 (1985).
147. H. Liang, T.E. Fischer, M. Nauer, C. Carry, *J. Am. Ceram. Soc.* **76**, 325 (1993).
148. T.E. Fischer, *Annu. Rev. Mater. Sci.* **18**, 303 (1988).
149. G. Heinicke, *Tribochemistry* (Carl Hanser, Munchen, 1984).
150. J.M. Martin, T. LeMogne, *Analisis* **25**, M28 (1997).
151. D.A. Rigney, J.E. Hammerberg, *Mrs. Bull.* **23**, 32 (1998).
152. R.G. Pearson, *Chemical Hardness* (Wiley, Weinheim, 1997).
153. F.G. Yost, *J. Electron. Mater.* **3**, 353 (1974).
154. H. Kato, *J. Electrochem. Soc.* **134**, 1750 (1987).
155. A. Hiraki, *Jpn. J. Appl. Phys.* **1** **22**, 549 (1983).
156. C.D. Thurmond, M. Kowalchik, *AT&T Tech. J.* **39**, 169 (1960).
157. M.K. Sunkara, S. Sharma, R. Miranda, G. Lian, E.C. Dickey, *Appl. Phys. Lett.* **79**, 1546 (2001).
158. J. Westwater, D.P. Gosain, S. Tomiya, S. Usui, H. Ruda, *J. Vac. Sci. Technol. B* **15**, 554 (1997).

159. X.F. Duan, C.M. Lieber, *Adv. Mater.* **12**, 298 (2000).
160. S.Q. Feng, D.P. Yu, H.Z. Zhang, Z.G. Bai, Y. Ding, *J. Cryst. Growth* **209**, 513 (2000).
161. G. Xu, H. Liang, S. Guruzu, *Friction-Induced Nucleation of Nanocrystals* (Springer, AA Dordrecht, 2006).
162. L.H. Peng, H.G. Lee, W. Teizer, H. Liang, *Wear* (2009), in press.

Structuring the Surface of Crystallizable Polymers with an AFM Tip

Cvetlin Vasilev, Günter Reiter, Khalil Jradi, Sophie Bistac, and Marjorie Schmitt

Summary. In this chapter, we demonstrate that atomic force microscopy (AFM) is a well-suited technique for manipulating polymers, in particular for inducing crystallization and for controlling crystal melting. As examples, we present results on the crystallization and melting behavior of highly asymmetric semi-crystalline block copolymers where the crystallizable block is confined in nanometer size mesophase domains and on lowering the crystal nucleation barrier by deforming polymer chain conformations via friction (rubbing) with an AFM tip. In the here discussed studies, the main focus is on questions concerning primary nucleation, morphological reorganization, crystal growth under confinement, non-equilibrium annealing processes, etc., which still remain unsolved despite numerous theoretical and experimental works. As will become clear from these examples, AFM offers unique possibilities for manipulations of polymers at a molecular scale, which affect their behavior on macroscopic scales up to sizes visible with the naked eye.

Key words: AFM, Crystallization, Friction, Melting, Polymer.

24.1 Introduction

The key advantage of high-resolution real space observations lies in the chance to observe and to follow in real time changes on fundamental length scales of polymer systems, such as the radius of gyration of these molecules or the thickness of lamellar crystals, even during the processes of crystallization and melting. Naturally, the main advantages of non-destructive atomic force microscopy (AFM) [1] in comparison to other types of microscopy can be described as follows: AFM bypasses the diffraction limit of the optical microscopy, thus allowing for very high-resolution images to be obtained. AFM also provides a means to image non-conducting and even relatively soft materials, without damaging the sample. Intermittent AFM (e.g., Tapping Mode AFM) allows mapping elastic properties of a sample surface. Moreover, it is possible to distinguish between crystalline and amorphous parts within

a sample. The advantages of AFM measurements have been often applied for real-time observation of polymer samples [2, 3], also for measurements at elevated temperatures [4–12]. AFM allowed studying central phenomena in polymer physics like ordering [4], crystallization [5–10], and melting [5, 11, 12] behavior of semi-crystalline block copolymers confined in very thin films. The processes of nucleation and crystal growth can get rather complex if lateral confinement of the crystallizable polymer in nanometer-sized phase separated domains is introduced [13–20]. Recent studies of such phase separated systems where the crystallizable block was confined in spherical nano-domains [21] or systems where the crystallizable block is confined in nano-droplets [22] showed that nucleation in such systems is only possible at large undercoolings and the resulting crystals were each restricted to a single mesophase domain. All crystals were formed independently without any visible coupling between neighboring domains. Crystallization at large undercoolings (i.e., at high crystallization rate) and within the restricted geometry of the mesophase compartments results in a variety of metastable crystals, which contained many defects and thus, they are quite imperfect. As a consequence of the metastability of the crystals, relaxation processes were occurring in these systems.

Polymer crystallization has been the subject of intensive studies for more than 50 years. Interest has centered on the fundamental understanding of the complex problem of the transformation from a disordered melt into a highly ordered crystalline structure. In particular, questions concerning nucleation, the transition from a randomly coiled state to an ordered crystalline state, crystal growth kinetics, etc. were intensively studied. Various microscopic, scattering and calorimetric techniques, such as optical and electron microscopy, small- and wide angle X-ray scattering (SAXS and WAXS), differential scanning calorimetry (DSC), etc., made available essential information about the polymer crystallization.

Optical microscopy, although fundamentally limited in resolution by diffraction, provided information on the growth of spherulitic super structures (growth rate measurements) [23, 24] and on local crystal orientation [25–27]. Electron microscopy, with its improved resolution (compared to the resolution of the optical microscopy) due to the short wavelength of the electrons used, has the capability of probing many of the important length scales of the polymer crystallization shedding light on the morphology of crystalline polymers. However, the high energy electrons can considerably damage the polymer samples even at very low doses of irradiation, causing for example cross-linking of the polymers or changes in the morphology of the crystalline and amorphous regions. SAXS and WAXS techniques allow measurement of the crystallinity, lamellar thickness, and lamellar separation, but they cannot provide a real space data (direct visualization) of the crystallization process. The DSC technique allows for measurements of the thermodynamic properties of the crystallization and melting behavior of the polymer crystals.

Our knowledge of polymer crystallization is based on a number of key observations: the folded chain crystals [28], the existence of a characteristic thickness of the crystalline lamellae depending on the undercooling at which growth occurred [29, 30] and the relationship between the crystallization temperature on the one hand and the growth rate and the apparent melting temperature of the polymer crystals. Recently, based on detailed interpretation of SAXS data, a new model was suggested [31], where the crystallization does not occur spontaneously at a clear crystal–melt interface, but rather through semi-ordered mesophase, which then transforms into the lamellae. However, many questions concerning primary nucleation, morphological reorganization, crystal growth under confinement, non-equilibrium annealing processes, etc. are still remaining unclear and the models that have been developed frequently fail to explain new observations. One of the most significant gaps in our knowledge is due to an inability to obtain a real space information with sufficiently high resolution on the fundamental (nanometer) length scales of the system, such as the size (radius of gyration) of the molecules and the thickness of the crystals in real time, during the process of crystallization, without significantly damaging the sample.

AFM allows following the processes of morphological changes, crystallization, and melting of the polymers, in situ, in real time, with nanometer resolution.

Using ordered polymer nanostructures such as block copolymers [32–39] allows demonstrating very impressively some of the unique possibilities of AFM for the study of crystallization and melting of semi-crystalline polymers. Molecular relaxation processes occurring in such systems upon heating above the crystallization temperature can be followed in situ. However, it is definitely less common to induce processes like local melting or characteristic deformations on a molecular scale via an AFM tip. The take-home message of this chapter is that it is possible to melt individual nanometer-sized polymer crystals without perturbing significantly neighboring crystals. Besides such surface structuring possibilities of AFM, the mechanical deformations, which an AFM tip can induce in a polymer surface, provide a unique way for directing crystal growth via a selective reduction of the crystal nucleation barrier.

Here, we present experiments where such ordered polymeric nanostructures, which can crystallize, are combined with local surface manipulation. In particular, we demonstrate how nanometer-sized polymer crystals can be molten individually when they are confined in a spherical compartment surrounded by a non-crystallizable matrix of a block copolymer mesostructure. One of the ways to control size, shape, and order of the mesophase patterns is to vary the properties of the blocks (composition, affinity, length, etc.). However, periodic patterns together with arbitrarily non-periodic features chosen by design cannot be obtained by such a self-assembly process. Using techniques with high spatial resolution, allowing for local manipulation on surfaces, one

can overcome this problem and generate a predefined, non-periodic, lateral pattern with a hierarchy of length scales.

We will demonstrate that melting of nanometer size polymer crystals via a hot AFM tip is feasible. A possibility for local modification of materials properties with heated AFM probe provides an extra dimension in obtaining custom-made nano-structured non-periodic materials. Moreover, our results indicate that by providing limited thermal energy, which is only increasing polymer mobility within the crystallites but insufficient to melt them, we also can improve the degree of order (crystallinity) of individual crystals in a controlled fashion.

Besides melting a polymer crystal via a hot AFM tip, one also can introduce mechanical deformations of the polymer conformations at the molecular scale. As a clear indicator for such deformations, we can analyze the probability of nucleating polymer crystals. A clear correlation between rubbing the surface with an AFM tip and the orientation of lamellar polymer crystals can be established.

24.2 Experimental Part

24.2.1 Characteristics of the Polymers Used

In block copolymers [33, 34], combining incompatible polymer chains in a single macromolecule causes to intra-molecular phase separation leading to the formation of highly ordered self-assembled phase-separated structures with a periodicity proportional to the radius of gyration (\mathbf{R}_g) of the constituent blocks of the molecules. The nature and size of these structures can be controlled by various means, including interfacial and architectural modifications [34].

Here, a block copolymer system consisting of two polymer chains of different length and chemical composition – a crystallizable block [polyethylene oxide (PEO)] and an amorphous block [polybutadiene (PB_h)] – resulting in hydrogenated poly(butadiene-block-ethylene oxide) (PB_h -b-PEO), with a molecular weight of the blocks of 21,100 and 4,300 g mol^{-1} , respectively, corresponding to a PEO weight fraction of 17%. The PB_h block contains statistically distributed 1–4 and 1–2 units, with a majority of 1–2 units ($\approx 85\%$). This block is therefore amorphous and exhibits a glass transition at $T_g \approx -18^\circ\text{C}$. This highly asymmetric diblock copolymer forms microphase structures consisting of spherical domains of the minority block (PEO) embedded in a matrix formed by the majority block (PB). Melting temperature of PEO in these domains is in the range from 42 to 47°C depending on the crystallization temperature and the thermal history of the sample after the crystallization. Upon annealing in the molten state, the mesophase pattern can be highly ordered, yielding a hexagonal packing of the PEO spheres over large areas (up to several square μm).

In a second series of experiments, we used a slowly crystallizing homopolymer, isotactic polystyrene (i-PS) with a tacticity of 90%, having a weight-averaged molecular weight of $M_w = 400,000 \text{ g mol}^{-1}$ and a polydispersity index of $I = M_w/M_n \approx 2.8$ (purchased from Scientific Polymer Products).

24.2.2 Sample Preparation

All samples used in this work were prepared by spin coating from diluted toluene solution onto pre-cleaned substrates, either silicon wafers or mica sheets. After the spin coating, the films were not in the equilibrium state (mostly due to the fast evaporation of the solvent) and thus it was necessary to anneal them. During the annealing of the sample, the polymer chains were allowed to rearrange and so changed from a less equilibrated state immediately after spin coating to a more equilibrated state after annealing. Annealing temperatures were in the range from 120 to 260°C and annealing times were in the range from 10 to 120 min. Heating and cooling rates were in the range from 1 to 100°C min⁻¹. The hot-stage used for annealing was purged with dry nitrogen.

Crystallization of the block copolymer samples was performed in a DSC device without measuring the heat flows during the heating and cooling runs. Our goal was just to perform the crystallization process at well-controlled conditions. During crystallization, the DSC chamber was purged with a flow of dry nitrogen.

Thin films of i-PS on a smooth silicon wafer were also prepared by spin coating. i-PS adsorbed on the oxide surface of silicon wafers, which were cleaned in water saturated UV-ozone atmosphere. The amorphous thin films with uniform thickness of about 11 nm were obtained with use of a 0.4 wt% cyclohexanone solution. Before thermal and frictional treatment, amorphous thin films were melted at 260°C for 3 min, quenched to room temperature (RT), which is much lower than the glass transition temperature, T_g (about 95°C). Subsequently, rubbing/scratching of these films at RT was realized with the tip of an AFM.

24.2.3 The Employed AFM Working Mode

The AFM apparatus used in these studies was a Multimode from Veeco with a NanoScope IV controller. Three-dimensional topographical maps of the surface were constructed by plotting the local sample height versus horizontal sample position. Tapping Mode AFM [40, 41] allows high-resolution topographic imaging of sample surfaces that are easily damaged, like polymeric materials. Tapping Mode (TM-AFM) overcomes problems associated with friction, adhesion, electrostatic forces, and other difficulties that obstruct conventional AFM scanning methods by periodically placing the tip in contact

with the surface to provide high resolution by measuring tip–sample interaction forces and then lifting the tip off the surface to avoid dragging the tip across the surface.

In the here presented experiments, TM-AFM probes with a nominal tip radius of the order of 10 nm (manufacturer’s data) and cantilevers with lengths of about 125 μm and about 225 μm were used. This corresponds to resonant frequencies of about 330 kHz and 190 kHz, and nominal spring constants (manufacturer’s data) of 42 and 48 N m^{-1} , respectively.

During scanning, the vertically oscillating tip periodically contacts the surface and lifts off. As the oscillating cantilever begins to intermittently contact the surface, the cantilever oscillation amplitude is necessarily reduced due to energy loss caused by the tip contacting the surface, but also the amplitude of oscillation scales directly proportional to the average distance of the probe and the sample. The reduction in oscillation amplitude is used to identify and measure surface features. The root mean square amplitude of the sinusoidal voltage signal from the photodiode varies with the amplitude of the oscillating cantilever. This signal is used to control the piezoelectric scanner in Z direction and to adjust the tip–sample separation in a way to maintain a constant force on the sample and the tip oscillation amplitude at a constant level, called the set-point amplitude. When an oscillating tip contacts the surface at high frequency, the contact time t_c between the tip and the sample surface is very short. At high enough frequencies, t_c can be shorter than the relaxation time of the sample t_{relax} (depending on the viscoelastic properties of the sample) and the sample surface appears stiff [42]. Thus, the influence of tip–sample adhesion forces can be greatly reduced. Because of the short contact time, Tapping Mode helps to prevent the tip from sticking to the surface and causing damage during scanning.

Changes in the phase angle of the oscillating cantilever allow producing an image which for polymers often provides significantly more contrast than the topographic image. The AFM phase signal is sensitive to material surface properties, such as stiffness, viscoelasticity, and chemical composition. During an oscillation near the sample surface (a few tens of nanometers above it) a combination of long-range attractive forces (mostly van der Waals forces) and short-range repulsive forces are acting on the AFM tip. AFM operating in Tapping Mode has two regimes of operation – the so-called “attractive” and “repulsive.” In the attractive regime, a net long-range attractive force dominates the reduction of the amplitude of the oscillating tip while in the repulsive regime a net short-range repulsive force controls the cantilever dynamics. The free oscillation amplitude \mathbf{A}_0 and more precisely the ratio $r_{\text{SP}} = \mathbf{A}_0/\mathbf{A}_{\text{SP}}$, where \mathbf{A}_{SP} is the amplitude set point, the drive frequency ω_0 and the viscoelastic properties of the sample control the phase image [43].

In general, for relatively soft samples like most polymers, AFM operates in attractive regime at relatively low \mathbf{A}_0 (in the range from 10 to 30 nm) and high r_{SP} while using higher \mathbf{A}_0 (50 to 100 nm) and lower r_{SP} describes the repulsive regime [43].

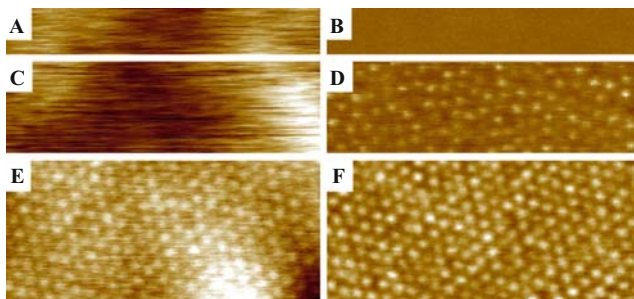


Fig. 24.1. The influence of the tapping force: (A), (C), (E) – AFM topographic and (B), (D), (F) – corresponding phase images taken at r_{SP} of 0.975, 0.957, 0.915, respectively. The width (horizontal axis) of the images is 500 nm. The Z-range is 3 nm and 3° for all topographic and phase images, respectively

In the systems used, due to the lower surface tension of the majority block (PB), the film surface is covered with a thin PB-layer of a few nanometers covering the spherical PEO domains. Thus, in the attractive regime or at low “tapping force” (high r_{SP} close to 1) only featureless topography and phase images were observed corresponding most likely to the real surface of the sample (see Fig. 24.1a, b). Upon lowering of r_{SP} (the system shifts to repulsive regime), the topographic image did not change much but on the phase image the stiffer crystalline domains could be distinguished (see Fig. 24.1c, d). Continuing to increase the force (strongly repulsive regime), one could clearly distinguish the crystalline and amorphous domains, both on the phase image and on the topography images (see Fig. 24.1e, f). Thus, in order to visualize the mesostructure and to detect the spherical PEO domains, the AFM tip had to penetrate at least slightly the top PB-layer. Crystalline PEO domains were stiffer and thus could not be deformed as much as the PB matrix. This difference is responsible for contrast in the phase shift and can also explain the observed topography, a result in accordance with other studies on various block copolymer systems [44–46]. Thus, it can be concluded that low-force imaging is adequate for a correct determination of the exact surface topography of a sample, whereas imaging at somewhat elevated forces can be useful for compositional analysis of the surface and for the detection of “hidden” domains having different mechanical properties [47].

Another key point for soft materials is a proper interpretation of the observed phase contrast of soft heterogeneous samples. Local variations of the phase can be linked to dissipative processes due to the interaction between the tip and the soft sample. When the tip touches the surface, viscous forces acting against the tip motion contribute to the phase lag of the oscillator. A simple method for qualitative determination of the contribution of the zones with different mechanical properties in the contrast of the AFM images (both topographic and phase) is the comparison between the force–distance curves recorded at different locations on the sample (corresponding to domains of

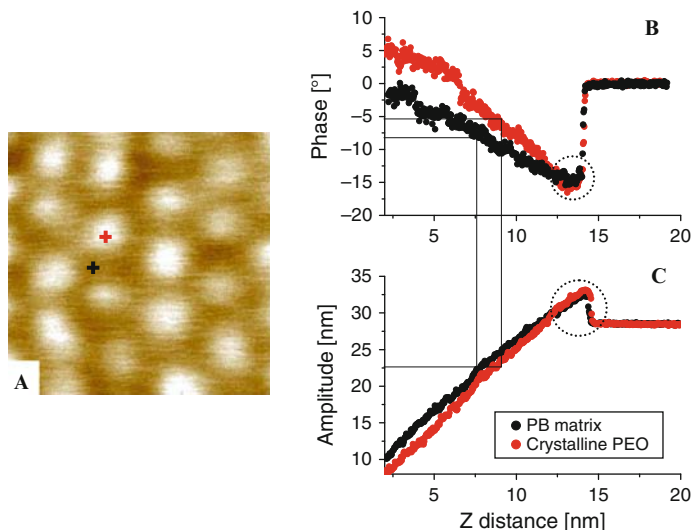


Fig. 24.2. Force–distance curves recorded during the approach of the AFM tip to the sample surface. **(A)** – $100 \times 100 \text{ nm}^2$ AFM phase image of crystallized HP 17 sample, **(B)** and **(C)** – phase shift and amplitude of the oscillating tip versus tip–sample distance, respectively. *Red curve* corresponds to the crystalline PEO domain (*red cross* on the AFM image) and *black curve* corresponds to the amorphous PB matrix (*black cross*)

different polymer chains) and the corresponding phase or topographic image. Such force–distance curves were recorded during the approach of the tip to the sample surface (Fig. 24.2b, c).

The two different approach force–distance curves correspond to the amorphous PB matrix and crystalline PEO domains, respectively. From the amplitude curves, we can estimate the contribution of viscoelastic properties of the different domains by comparing the piezo displacement over different domains at constant amplitude of about 23 nm (corresponding to the amplitude set-point A_{SP} during the imaging of the image in Fig. 24.1e, f). In order to maintain the amplitude of the tip constant, the piezo has to move about 2 nm deeper into the PB (softer) domain with respect to the piezo displacement over the crystalline PEO domain (stiffer). This value can be compared to the Z range of the image in Fig. 24.1e which is 3 nm. The corresponding phase contrast is about 3° .

Another interesting feature concerns the zones exhibiting a minimum or a maximum (*dotted circles*) of the phase and amplitude curves, respectively. These zones correspond to AFM imaging at the transition between attractive and repulsive regime. As a consequence of the interaction with the stiffer crystalline PEO domains, the phase lag of the oscillating tip is larger compared to the phase lag observed on the amorphous PB matrix. On the contrary, in

the strongly repulsive regime (at decreasing Z distance) the phase lag caused by the amorphous PB matrix is larger compared to the phase lag caused by the crystalline PEO domains. In conclusion, the oscillator parameters (\mathbf{A}_0 , \mathbf{A}_{SP} , and ω_0) together with the viscoelastic properties of the sample determine the imaging regime (attractive or repulsive).

24.3 Melting of Confined, Nanometer-Sized Polymer Crystals

24.3.1 Self-Assembly and Non-Periodic Patterns

The self-assembly of block copolymers into ordered mesostructures exhibiting a nanometer periodicity has already been extensively used to generate nano-structured systems [16,33–39]. Varying the properties of the blocks (composition, affinity, length, etc.) allows to control size, shape, and order of the resulting pattern. The periodicity of these structures is the same everywhere in the whole sample [48]. However, periodic patterns together with arbitrarily non-periodic features chosen by design cannot be obtained by such a self-assembly process. Using techniques with high spatial resolution, allowing for local manipulation on surfaces, one can overcome this problem and generate a predefined, non-periodic, lateral pattern with a tunable hierarchy of length scales. Here, the goal is to create changes in properties over areas as small as a few square nanometers. Such may be achieved by depositing molecules locally [49], by performing a chemical reaction in restricted geometries [50], or by controlling a phase transition in a small volume. Several pathways have been proven as feasible [49–55], most of them use a small probe like the tip of an atomic force microscope (AFM) which allows to interact locally with the sample: transfer of ink molecules with an AFM tip (dip-pen) [49], tip-induced polymerization [50], melting [51,52], nanomachining [53], modification of the electrical properties [54], etc. In contrast to approaches like the fabrication of nanoscale materials by co-extrusion of two reactive polymers [56], local modification of materials properties with a proximity probe provides an extra dimension in obtaining custom-made nano-structured non-periodic materials.

Here, ordered polymeric nanostructures and local surface manipulation are combined. In particular, it has been examined how nanometer-sized polymer crystals can be molten individually when they are confined in a spherical compartment (“cell”) surrounded by a non-crystallizable matrix of a block copolymer mesostructure (Fig. 24.3). Attention was paid to possible changes of the mesostructure induced by local melting of crystalline cells. The aim was to melt one single crystalline cell without perturbing its neighboring cells. To this end, one had to transfer an amount of thermal energy sufficient to melt a crystallite. In addition, one had to control that the heat source was exposed only to small areas of the surface in order to avoid that neighboring cells were molten as well.

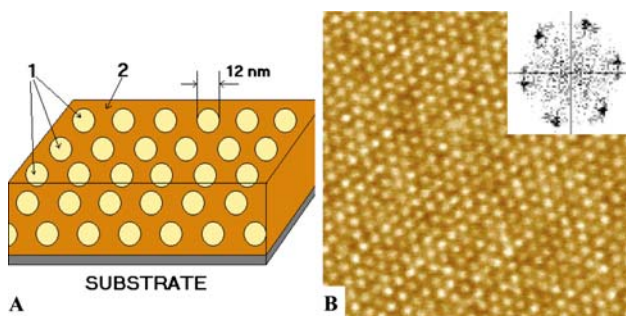


Fig. 24.3. The mesostructure of the studied PB_h -b-PEO block copolymer. (A) – Schematic drawing of the mesophase pattern in a thin film: 1 – crystalline PEO cells; 2 – liquid PB_h matrix; (B) – AFM phase image ($500 \times 500 \text{ nm}^2$) showing the sample after annealing and subsequent crystallization at -25°C for more than 20 h. The inset shows the corresponding Fourier transform, indicating a well-ordered pattern

The advantages of AFM measurements at elevated temperatures have been recognized for some time and several heating stages have been described in the literature [5,47,57,58]. In the field of polymers, AFM measurements at elevated temperatures were employed to study crystallization/melting processes and the growth of individual crystalline lamellae at a nanometer resolution. Such simultaneous in situ and real-time information could not be obtained with any other microscopic technique. In the present study, AFM was used to localize the heating of the sample surface to areas comparable to the size of the heated AFM tip. This AFM approach has the advantage that it allows both, manipulation and characterization (imaging) of the sample. The sample consisted of nanometer-sized individual crystals separated by an amorphous matrix of a hydrogenated poly(butadiene-block-ethylene oxide) (PB_h -b-PEO) mesostructure (Fig. 24.3). Crystallization of the PEO block in this particular system was already studied extensively [21, 59].

It was shown that nucleation is only possible at large undercoolings of about 60 K. The resulting crystals were each restricted to a single spherical cell of the mesostructure. All crystals were formed independently without any visible coupling between neighboring compartments. We emphasize that the huge difference of about 60 K between the temperature where a crystal can be formed and the temperature where it eventually melts provided a large temperature window where neither melting nor crystallization occurs. Consequently, in this temperature range a partially crystalline state of the sample can be locked in. However, although the amorphous PB_h matrix presents an enormous barrier hindering exchange of PEO blocks between neighboring cells, it cannot be excluded that a very small but non-zero probability exists for such an exchange. Thus, besides the possibility of local melting, the study focused on the influence of heating (annealing) on long-range order of the arrangement of cells and local changes at the level of individual cells.

24.3.2 The AFM Set-Up Employed for Local Heating

When using Tapping Mode atomic force microscopy (TM-AFM) one can estimate the “tapping force” (F) applied onto the sample surface during intermittent contact approximately by [60]

$$\mathbf{F} \sim k(\mathbf{A}_0 - \mathbf{p} \cdot \mathbf{A}_{\text{SP}}) \sim k \cdot \mathbf{A}_0(1 - \mathbf{p} \cdot \mathbf{r}_{\text{SP}}), \quad (24.1)$$

where k is the spring constant of the cantilever, \mathbf{A}_{SP} is the working amplitude set-point and \mathbf{p} is correction parameter which takes into account possible indentation of the tip into the sample. To avoid adhesive tip-sample interactions, large drive amplitudes of the oscillating probe were used. Typically, for the tip of the freely oscillating cantilever not in contact with the sample an amplitude set-point \mathbf{A}_0 equivalent to ~ 70 nm was chosen. The amplitude ratio $\mathbf{r}_{\text{SP}} = \mathbf{A}_{\text{SP}}/\mathbf{A}_0$ is thus related to the applied force. The higher the value of \mathbf{r}_{SP} the weaker is the force the tip applies onto the sample surface, at constant drive amplitude. In the following, \mathbf{r}_{SP} is taken as an indicator of the applied force.

As mentioned already, due to the lower surface tension of the majority PB_h block, the film surface was covered with a thin PB_h -layer of a few nm. Thus, a low-force imaging is adequate for a correct determination of the surface topography of the sample, whereas imaging at somewhat elevated forces can be useful for the analysis of the composition of the near-surface layer and for the detection of “hidden” PEO-domains exhibiting different mechanical properties [47]. As an important side-result, it is intriguing that at RT these hidden crystalline PEO crystals could not be destroyed by the mechanical impact of the AFM tip, even when \mathbf{r}_{SP} was decreased to 0.01. Upon increasing \mathbf{r}_{SP} , that is, re-imaging the scanned area at significantly lower tapping force, the obtained AFM images were fully identical with the initially taken images.

In order to transport locally enough thermal energy to melt a crystalline PEO cell, a commercial equipment was used (Fig. 24.4), which allowed to heat the sample independently from below via a hot stage (at a nominal temperature \mathbf{T}_H) and from above via a heated AFM probe (at a nominal temperature \mathbf{T}_L). Thus, the sample was sandwiched between two heaters of controlled temperatures. However, the exact temperature at the sample surface itself, which is the crucial parameter, could not be determined precisely [57] due to the different thermal conductivities of the various media involved (air, polymer, substrate, etc.) which caused thermal gradients difficult to measure.

24.3.3 Local Melting of Confined Polymer Crystals

In our first attempts to melt individual PEO cells, the film surface was approached with a hot tip while the heating from below the sample was switched off, that is, the sample stage was kept at RT. The nominal temperature of the tip (\mathbf{T}_L) was varied in a wide range from 55 to 120°C. In

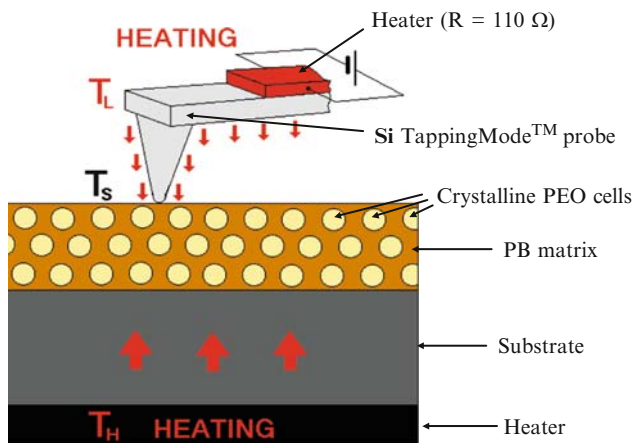


Fig. 24.4. Schematic drawing of the components of the heating stage

all cases, part of the crystalline cells, even when they were far away from the location of the tip, were unintentionally melted during such an approach. Crystalline cells melted not just directly under the tip but also over large areas around the tip and especially below the cantilever of the probe. At an elevated temperature $T_L = 80^\circ\text{C}$, about 30–40% of the initially all crystalline cells melted (Fig. 24.5). In part, this was certainly due to heat transfer via the heated air around the cantilever as opposed to heat transferred only via contact of the tip with the surface.

One may ask why did not all crystals melt? What is the reason that some crystals melted and some did not? To answer this question, one has to know that polymer crystallization at large undercoolings and within the restricted geometry of the spherical compartments resulted in a variety of metastable crystals which contained many defects and differed in stability, corresponding to different degrees of order related to their crystallinity. This difference in stability is reflected in melting temperatures T_{melting} distributed over a wide range of about 10°C . As one possibility for explaining this difference, one may consider that the mean size of the crystallites within these nano-spheres fluctuated, and thus the melting temperature, which is proportional to the size of the nano-crystals. As shown previously [21, 47], the width of this melting range depends on thermal history. The stability of polymer crystals (their perfection, their degree of order) can be improved by annealing close to the melting temperature. As a result, the melting range became narrower and the average melting temperature shifted to higher temperature [59] (see Fig. 24.6).

Taking advantage of this phenomenon of crystal improvement, which is exhibited by most polymer crystals, the samples were annealed at a temperature slightly below the melting point of the majority of the crystallites. Thereby, the number of unintentionally molten cells was minimized but as a drawback the long-range order of the mesophase pattern was lost (Fig. 24.7).

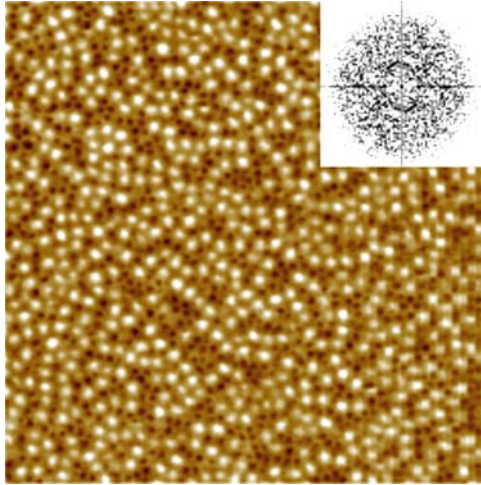


Fig. 24.5. AFM phase image ($1 \times 1 \mu\text{m}^2$) showing the sample after the approach of the hot AFM probe: $T_H = 25^\circ\text{C}$, $T_L = 80^\circ\text{C}$. The stiffer regions appear brighter than the softer regions. The three different phase contrasts correspond to stiff crystalline PEO cells (*bright*), the amorphous PB matrix (*gray*), and the molten PEO cells (*dark*). The inset shows the corresponding Fourier transform indicating the loss of order

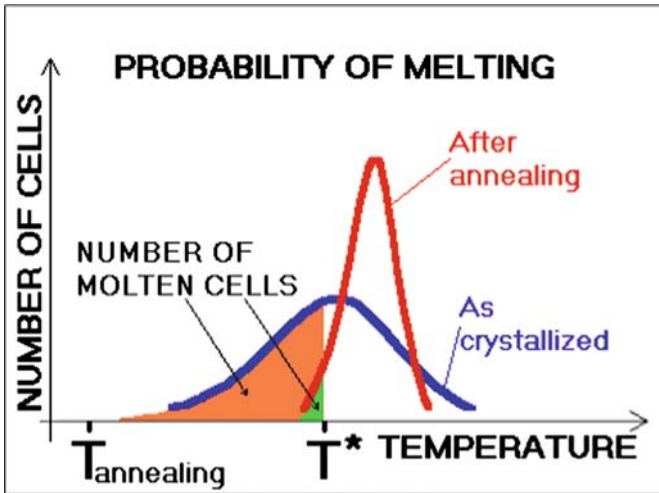


Fig. 24.6. Schematic representation of the changes in the distribution of the number $N(T)$ of PEO cells melting at a given temperature T , after annealing of the sample close to the stability limit of the cells (*dashed line*) with respect to the as crystallized sample (*full line*). When taking an AFM image at T^* all cells $N(T < T^*)$ will be molten

In order to investigate the reason for this loss of order, the temporal evolution of the mesophase pattern caused by annealing at $T_{\text{H}} = 36^{\circ}\text{C}$ (Fig. 24.8) was investigated. It became obvious that some cells disappeared almost instantaneously leading to an increase of the average distance between cells. Moreover, in the course of time, disorder increased. There were clear

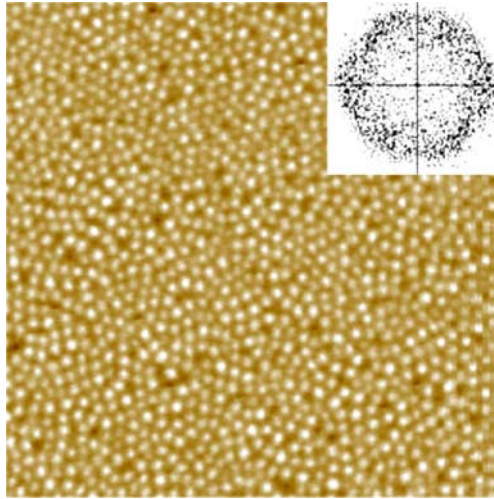


Fig. 24.7. AFM phase image ($1 \times 1 \mu\text{m}^2$) of a sample which, after crystallization, was annealed for 10 min at 36°C , taken with a hot AFM probe: $T_{\text{H}} = 36^{\circ}\text{C}$, $T_{\text{L}} = 70^{\circ}\text{C}$. The inset shows the corresponding Fourier transform indicating partially preserved order

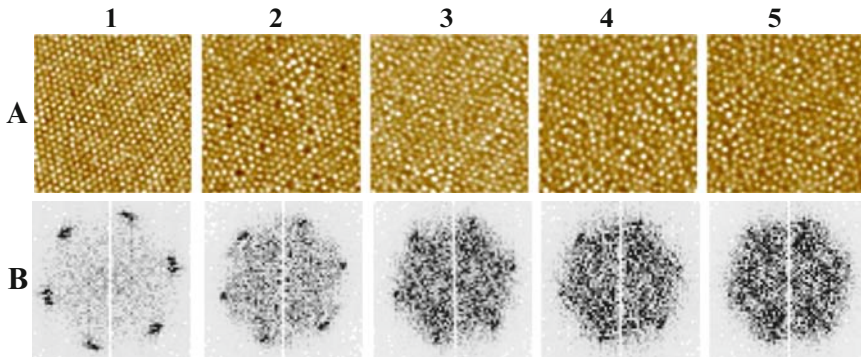


Fig. 24.8. Relaxation of the crystallized sample during in situ annealing close to the melting point. (A1) to (A5) – AFM phase images ($500 \times 500 \text{nm}^2$) after annealing for 0, 5, 20, 240, and 340 min, respectively, $T_{\text{L}} = \text{room temperature}$ (heating is off), $T_{\text{H}} = 36^{\circ}\text{C}$; (B1) to (B5) – corresponding Fourier transforms (FFT) of the AFM images. One can observe an almost instantaneous increase in the average distance between the cells and a continuous loss of the order of the pattern with time

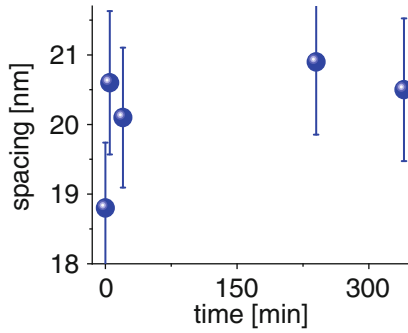


Fig. 24.9. The average distance (obtained from FFTs) between the cells versus time

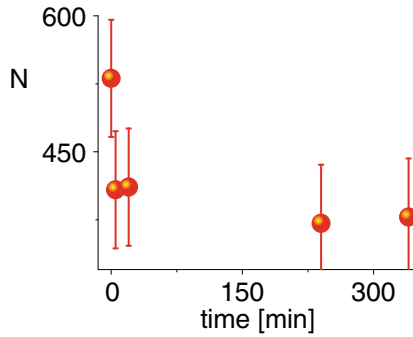


Fig. 24.10. Total number N of cells within the area of $500 \times 500 \text{ nm}^2$ versus time

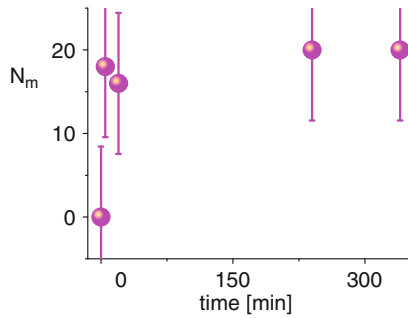


Fig. 24.11. The number of molten cells N_m (within the area of $500 \times 500 \text{ nm}^2$) versus time

indications that the distances between the PEO cells (Fig. 24.9), their total number per unit area (Fig. 24.10), and the number of molten cells (Fig. 24.11) changed with time.

There were also clear indications that the distances between the PEO cells and their number per unit area changed with the annealing temperature (Fig. 24.12). For higher temperatures and long annealing times (e.g., after

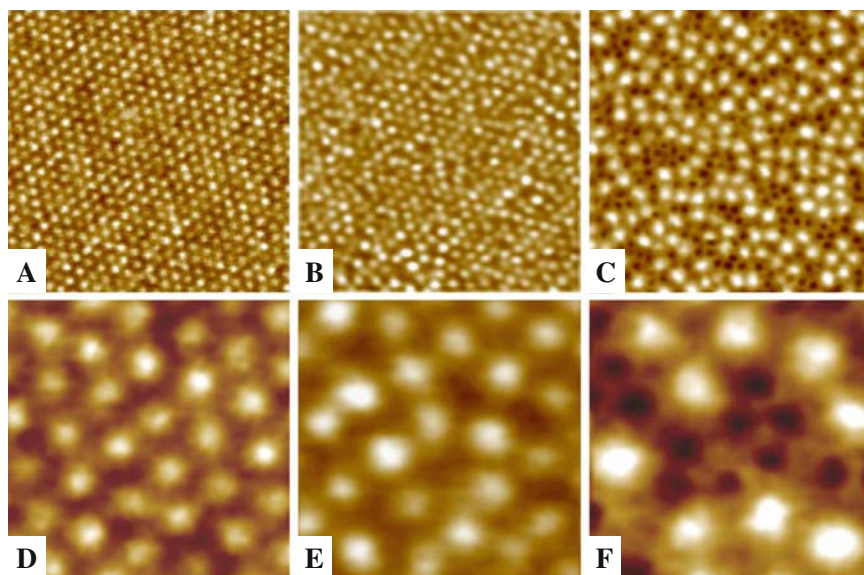


Fig. 24.12. AFM phase images showing a loss of mesophase order during annealing (**A**) – as crystallized sample, 100% of the PEO cells are crystalline and the average distance between them is 18.8 ± 1 nm; (**B**) – sample annealed for 20 min at 36°C , about 4% of the PEO cells are molten and the average distance between them increased to 20.1 ± 1.5 nm; (**C**) – sample annealed for ~ 5 h at 37.5°C about 20% of the PEO cells are molten, one can observe a bimodal distribution of the distances between the cells – the average distance between the crystalline ones increased to 27 ± 2 nm, while the average distance between the molten ones increased only to 21 ± 1 nm. The size of all images is 500×500 nm²; (**D**), (**E**), and (**F**) – represent 125×125 nm² sections of A, B, and C, respectively

about 5 h at $T_{\text{H}} = 37.5^\circ\text{C}$), a bimodal distribution of the average distances between the cells was found. There, the average distance between the crystalline cells increased from 19 ± 1 nm (for the as crystallized sample) to 27 ± 2 nm, while the distance between molten cells was about 21 ± 1 nm. This rearrangement, which caused the loss of order of the pattern, can most probably be explained by a limited but non-negligible diffusion of PEO blocks from molten cells toward crystalline ones.

Such diffusion was facilitated by the increased mobility of the polymer chains at elevated temperatures. This assumption was corroborated by the fact that a bimodal distribution of cell diameters appeared at comparatively high annealing temperatures, where the number of the molten cells was higher. There, more molten PEO blocks were available to “feed” the crystalline cells. As a consequence, re-organization resulted in the growth of the crystalline cells with material from molten cells, which, in turn, led to rearrangement of the pattern and the loss of its order. In summary, annealed samples were

stable (as defined by the low percentage of unintentionally molten cells) up to comparatively high temperatures (e.g., $T_H = 37^\circ\text{C}$).

It should be noted that tapping mode AFM and the chosen block copolymer system exhibit several advantages. Confinement of nanometer-sized PEO crystals in a soft PB matrix and the periodic arrangement of these PEO crystals were achieved automatically by an intrinsic self-assembly process. The thermal conductivity of the PB matrix is rather low (about 0.1 W mK^{-1} , this is, e.g., about three orders of magnitude lower than for metals). Such low thermal conductivity prevents significant heat propagation from the crystalline cell in contact with the tip to neighboring cells. Moreover, the viscoelasticity (softness) of the PB matrix dissipates mechanical energy and thus reduces the influence of deformation on neighboring cells. In addition, TM-AFM allows controlling penetration depth (by varying the applied force onto the sample surface), contact time, and contact area between the hot tip and the crystalline cells [44–46]. All these aspects together provide a sensitive and precise “dosage” of the heat transferred to an individual crystalline cell.

In order to locally increase the temperature T_C of an individual cell in contact with the tip above its melting temperature while the temperature T_S of the rest of the sample had to stay below the melting point of the crystalline material (Fig. 24.13). The necessary condition for melting a chosen cell is defined by the following inequality:

$$T_S < T_{\text{melting}} < T_C \quad (24.2)$$

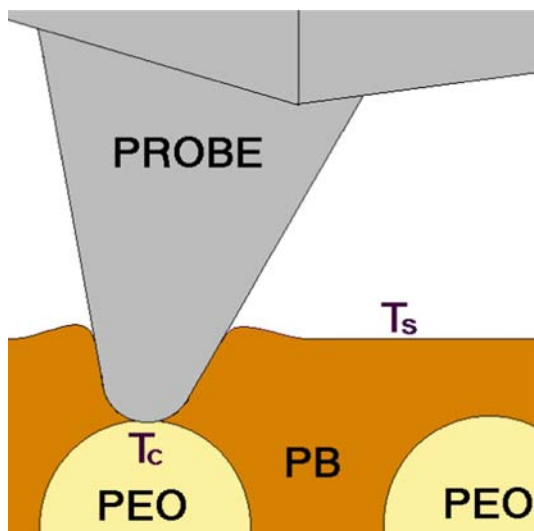


Fig. 24.13. Schematic drawing of tip–crystalline cell contact at high tapping force

The additional thermal energy, which has to be provided to melt a cell, depends on the difference between T_S and T_{melting} . The higher T_S the lower is the amount of additional thermal energy needed for melting. In chosen set-up, heating the whole probe did not allow to use high temperatures for T_L because of the risk to unintentional melting of cells which were not located directly under the AFM tip. Thus, rather high temperatures T_H were used. Under such conditions not much additional thermal energy was needed to melt the crystals. Comparatively, low nominal temperatures T_L of the probe could be used, which, in turn, avoided that the amount of heat transferred through the surrounding air was able to melt cells in the neighborhood of the probe.

In order to verify if selected individual crystalline cells could be melted on purpose, several experiments were performed at different temperatures of the heater underneath the sample (T_H , ranging from 33 to 39°C) and T_L (ranging from 55 to 70°C). First, a small area (e.g., $200 \times 200 \text{ nm}^2$) was scanned for periods of time between 5 min and 1 h. Zooming out afterwards (e.g., to $1 \times 1 \mu\text{m}^2$) showed that, regardless of the values of T_H and T_L , no crystals were melted via the hot tip. We did not observe any significant difference in the number of the very few molten cells inside and outside the initially scanned area. Thus, the exchanged thermal energy between the hot tip and the sample was insufficient to melt PEO cells on purpose. The failure of this melting attempt can be attributed to the lack of efficient heat transfer to the PEO crystals, which was hindered by the thin (about 5 nm thick) surface layer of the PB matrix.

To overcome this barrier, a much higher tapping force was applied than for the measurements described before where r_{SP} was limited to the range from 0.90 to 0.97. This corresponded to lowering the value of r_{SP} . A higher tapping force allowed the tip to penetrate more deeply into the sample (probably even to pass through the PB surface layer) and thus to approach the crystalline cells more closely (Fig. 24.13).

At the same time, this procedure also increased the contact time and contact area between the tip and the sample [43, 45, 46]. When r_{SP} was reduced down to 0.05, the crystalline cells located directly below the hot tip could be melted without perturbing the neighboring cells. During such melting, the imaging mode of TM-AFM was used but the size of the scanned area was only of the order of the size of the PEO cell (10 nm). In order to create custom-designed non-periodic patterns of larger sizes, the tip was displaced slowly over the sample (at about 50 nm s^{-1}) while tapping at low r_{SP} . A typical result can be seen in the phase images displayed in Fig. 24.14.

All molten cells were localized in a square-shaped pattern resulting from the displacement of the tip on the sample. This demonstrates that it is possible to “design” patterns with regular geometrical shape simply by moving a hot AFM tip over the surface of a crystalline diblock copolymer sample. Surprisingly, although the width of the molten regions was expected to be only about 10 nm (i.e., comparable to the width of the radius of curvature of the tip), the width of stripe was on the average about 70 nm (Fig. 24.14, left

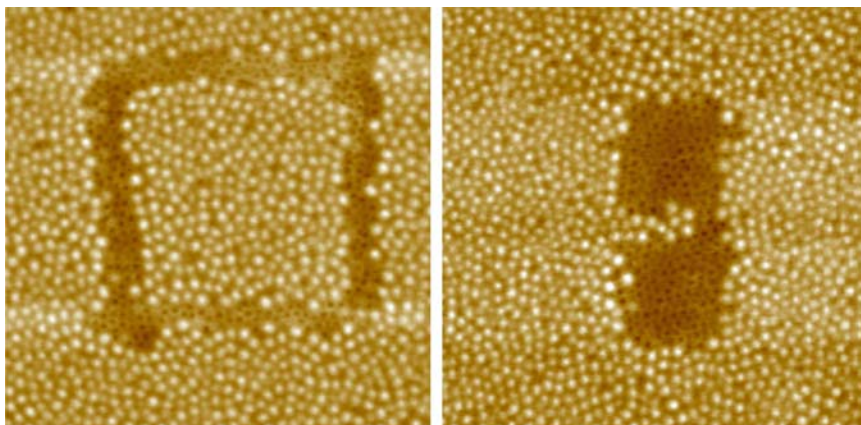


Fig. 24.14. AFM phase images ($1 \times 1 \mu\text{m}^2$) showing patterns sketched on the sample surface by local melting of the crystalline cells under the tip. The phase contrasts have the same meaning as explained in Fig. 24.5

image). This increased width of the melting trace can partially be attributed to the increase of the contact area, a consequence of the conical shape of the tip and comparatively deep penetration into the sample at low r_{SP} during “writing” (localized melting).

It should be mentioned that in the topography images the molten areas appeared to be lower than the surrounding crystalline cells (Fig. 24.15). However, for the chosen imaging parameters, the height image also contained information about the penetration of the tip into the sample. Crystalline cells are much stiffer than molten cells. Thus, at a given r_{SP} , the tip penetrated deeper into molten cells. The relative penetration depth for molten cells compared to crystalline cells was determined by taking cross-sections across partially molten samples. Independent of how cells were molten, either by annealing the sample close to the average melting point of the sample or by deliberately melting cells with a hot tip at high tapping force, such a comparison yielded the same penetration depth contrast. In the example shown in Fig. 24.15 taken with $r_{\text{SP}} = 0.92$, this contrast is about 4 nm.

The boundaries of the resulting molten traces were rather well defined and not frayed out as it might be expected to happen by lateral diffusion of heat within the film. Such diffusion of heat would have also allowed melting surrounding cells, which were not in intimate contact with the hot tip. In part, the relative sharpness of the boundaries was certainly due to the limited thermal conductivity of the polymers. The amount of lateral heat transfer was insufficient to induce melting of cells about one cell diameter away from the tip. However, there exists an additional reason for the relative sharpness of the boundaries. Although small and insufficient to melt a crystal, the thermal energy supplied to crystals at the border to the molten region was sufficient to increase the mobility of the molecules within the crystallites. Such mobility

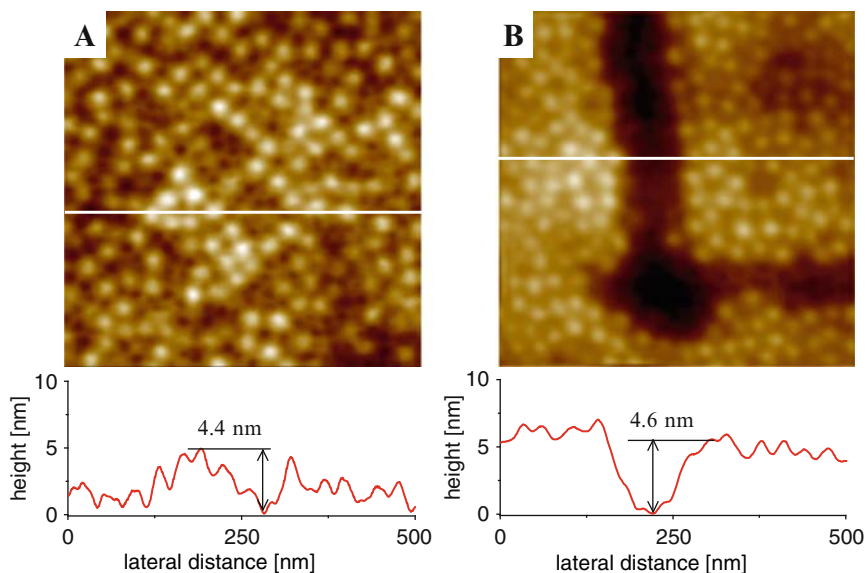


Fig. 24.15. Influence of the tapping force used in imaging mode on the measured topography on samples consisting of molten (soft) and crystalline (hard) cells. Topography images including sections for: **(A)** sample where some of the crystalline cells were unintentionally molten by heating the as-crystallized sample to $T_{\text{H}} = 37.5^{\circ}\text{C}$. (The difference in height between the molten and the crystalline cells is about 4 ± 1 nm); **(B)** sample where “local melting” was performed. (The difference in height between the molten and the crystalline cells is also about 4 ± 1 nm). Both images were obtained at the same $r_{\text{SP}} (= 0.92)$. Because the depth variations between the molten and the crystalline cells are comparable in both cases, this indicates that the apparent depth of the molten regions is not due to damage of the sample. These differences in height appear only because of a difference in penetration of the AFM tip as a consequence of the largely different modules of the molten and crystalline cells

led to crystal improvement and also to slow diffusion of PEO chains from the molten cells toward the crystalline ones. As a consequence of such reorganizations, better-ordered and bigger crystals were formed, indicated by a narrowing of $N(\mathbf{T})$ and a shift of the position of the melting peak to higher temperatures (see Fig. 24.6). Such processes also caused a “self-stabilization” of the cells at the borders of molten areas. Therefore, there is a strong tendency for the formation of a “fence” of more stable crystalline cells of higher T_{melting} surrounding molten areas. In some cases, we could even observe and distinguish these better-ordered crystals located at the borders due to the increased brightness which indicates a higher stiffness and thus larger and better-ordered crystals (Fig. 24.16).

The ultimate goal was to localize the heat transfer between tip and sample over areas comparable to the size of one single crystalline cell. Thus, the study

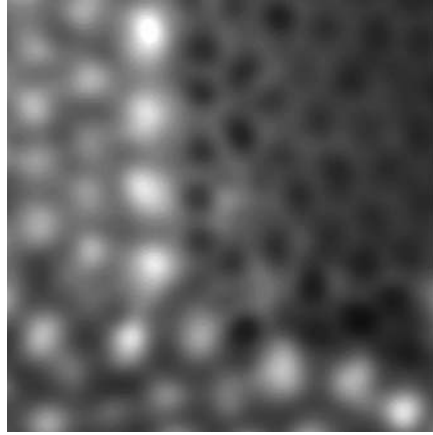


Fig. 24.16. AFM phase image ($200 \times 200 \text{ nm}^2$) focusing on the crystalline PEO cells at the borders of the molten region. The brighter color of these PEO cells may correspond to an increased size or better-ordered crystals inside these cells. In B and C, $r_{\text{SP}} = 0.07$ during the melting process and $r_{\text{SP}} = 0.92$ during imaging, $T_{\text{H}} = 37.5^\circ\text{C}$, $T_{\text{L}} = 70^\circ\text{C}$

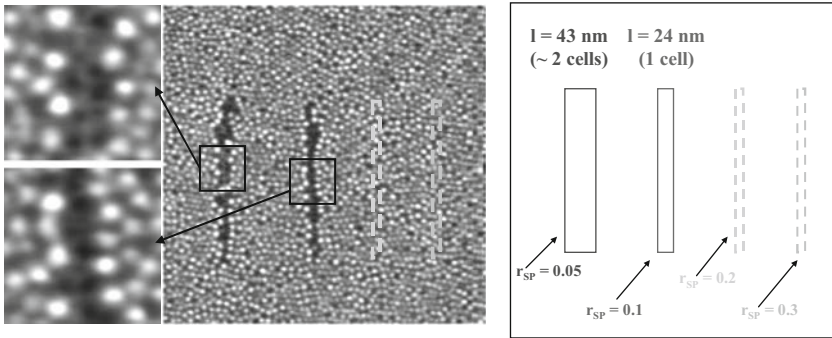


Fig. 24.17. AFM phase image ($1.3 \times 1.3 \mu\text{m}^2$) and scheme of the corresponding pattern demonstrating the dependence of the width of the molten trace with respect to the tapping force (indentation). Four runs at different values of r_{SP} are shown (from left to right): 0.05 – wider trace, average width $w \approx 43 \text{ nm}$; 0.1 – narrower trace, average width $w \approx 24 \text{ nm}$; 0.2 and 0.3 – marked with *dashed lines*; $T_{\text{H}} = 36^\circ\text{C}$, $T_{\text{L}} = 70^\circ\text{C}$.

As insets $200 \times 200 \text{ nm}^2$ sections of (A) showing molten traces at r_{SP} of 0.05 and 0.1 with average width of about 43 nm and 24 nm, respectively, are shown

focused on how the width of the molten trace depended on the penetration depth of the tip into the sample, that is, on the value of r_{SP} . At the conditions chosen in the experiments shown in Fig. 24.17, molten traces were only observed for the two lowest values of r_{SP} .

Interestingly, the widths of these two visible traces were distinctly different. The average width at $r_{\text{SP}} = 0.1$ was about 24 nm, which matches the periodicity of the mesophase pattern, while at $r_{\text{SP}} = 0.05$ the average width of the trace was about twice as large (see Fig. 24.17).

Based on these observations, the following three conclusions can be drawn: First, there exists a threshold value of r_{SP} below which the contact time and contact area are high enough to allow the transfer of sufficient thermal energy to melt crystalline cells. Second, by varying of the penetration depth, the width of the molten trace can be controlled. Third, comparing the results from Figs. 24.14 and 17.17 (Fig. 24.14, left image: $T_{\text{H}} = 37.5^{\circ}\text{C}$, $r_{\text{SP}} = 0.07$, Fig. 24.17: $T_{\text{H}} = 36^{\circ}\text{C}$, $r_{\text{SP}} = 0.05$), a decrease of the average width of the molten trace from about 70 nm to about 43 nm can be observed. At lower temperatures, T_{H} of the sample heater, more thermal energy has to be transferred from the tip to the sample in order to melt a crystalline PEO cell. In other words, at constant T_{L} and r_{SP} but lower T_{H} , the same amount of thermal energy provided via the tip ($T_{\text{L}} = 70^{\circ}\text{C}$ in Figs. 24.14 and 17.17) allows melting of fewer crystalline cells. In an attempt to describe our observations one arrives at the following qualitative relation:

$$w \sim DT_{\text{L}}/DT_{\text{H}} \cdot r_{\text{SP}}, \quad (24.3)$$

where w is the average width of the molten trace, $\Delta T_{\text{H}} = T_{\text{melting}} - T_{\text{H}}$ and $\Delta T_{\text{L}} = T_{\text{L}} - T_{\text{melting}}$. This relation allows designing patterns with a selectable width of the traits.

In addition, as the process of melting is reversible, it was also possible to perform similar experiment in a complementary way. Molten cells could be crystallized individually, for example, using an AFM tip as a “cold finger” in order to individually nucleate crystallization in a cell. Such induced nucleation was attempted with a highly asymmetric polyethylene-b-poly(styrene-*r*-ethylene-*r*-butene) diblock copolymer (E/SEB) with molecular weight of the polyethylene (E) and poly(styrene-*r*-ethylene-*r*-butene) (SEB) blocks of 9,000 and 55,000 g mol⁻¹, respectively, corresponding to a E weight fraction of 14%. The melting and crystallization temperatures of the E block were found to be around 100°C and 60°C, respectively, while the matrix remains above its glass transition (T_{g} of the SEB block is about 25°C) during crystallization process. We found that nucleation can indeed be induced in the E spheres (already established in the melt) by cooling each sphere individually under its crystallization temperature with a cold AFM tip. The sample (T_{H}) was kept at 72°C, well above the crystallization temperature for the majority of the E spheres (60°C). Thus, only in very small number of the polyethylene compartments nucleation and subsequent crystal growth could occur. Then we routinely imaged (r_{SP} was about 0.95 and the scan rate was about 3 Hz) an area of $1.5 \times 1.5 \mu\text{m}^2$ while keeping the AFM tip (T_{L}) at RT. Upon completion of five subsequent scans (which took about 400 s) over the same area, we zoomed out to $3 \times 3 \mu\text{m}^2$ and acquired an image at the same imaging conditions (Fig. 24.18). As a result, we observed that more

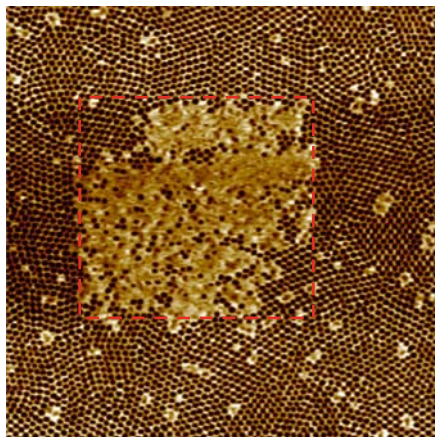


Fig. 24.18. AFM phase image ($3 \times 3 \mu\text{m}^2$) demonstrating that the probability for nucleation in the E spheres (at $T_{\text{H}} = 72^\circ\text{C}$) was enhanced dramatically by contact with a cold AFM tip ($T_{\text{L}} = \text{RT}$) for long enough time. The area of the initial small size scan is outlined with the dashed line. The phase contrasts has the same meaning as in Fig. 24.6

than 88% of the E spheres within the initial small scan area were crystallized, while outside that area nucleation occurred in less than 5% of the E spheres. In order to confirm the latter number, we have checked locations as far as $60 \mu\text{m}$ away (in each direction) from the initial scan area, that is, areas which were not even under the cold cantilever when the initial scans were performed. This observation clearly demonstrated that the probability for nucleation was dramatically enhanced only for those E spheres which underwent direct contact with the cold AFM tip for long enough time. Thus, the direct cooling of the individual polyethylene sphere below the global sample temperature, most probably assisted by the mechanical impact of the oscillation AFM tip, allowed only selected polyethylene domains to be crystallized.

These experiments demonstrate that melting of nanometer size polymer crystals via a hot AFM tip is feasible. Moreover, these results indicate that by providing limited thermal energy, which is only increasing polymer mobility within the crystallites but insufficient to melt them, one also can improve the degree of order (crystallinity) of individual crystals in a controlled fashion. In addition, as the process of melting is reversible, it will also be possible to perform the experiments in a complementary way. Molten cells could be crystallized individually, for example, using a cold tip to individually nucleate crystallization in a cell. However, long-range order and long-time stability of the array of cells are significantly affected by the non-zero probability of diffusion of PEO block across the incompatible PB-matrix. It should be possible to avoid or to lower such diffusion by either cross-linking the polymer matrix surrounding the crystalline domains or separating the domains further by using

longer blocks forming the matrix. Thus, the combination of the advantages of block copolymers (ordered self-assembled mesostructures with a tunable nanometer length scale) with the possibility of TM-AFM to interact locally with a sample enables us to superpose periodic and non-periodic structures, a new way for the integration of a hierarchy of length scales in surface patterns. This approach also represents a powerful tool to manipulate mechanical properties in small volumes of materials and opens up an additional possibility for surface engineering controlled on a nanometer scale.

24.4 Lowering the Crystal Nucleation Barrier by Deforming Polymer Chains

Besides locally melting crystalline structures, AFM also allows to generate well-defined crystallization patterns. This can be achieved by mechanically deforming the conformations of semi-crystalline polymers while these molecules are still amorphous. Such deformations can be introduced by moving an AFM tip across the surface of these polymers. However, if the polymers are in their molten state, the deformed chains will rapidly relax and attain equilibrated conformations. Therefore, deformations have to be generated in the glassy state. An often employed approach, complementary to deformations induced via an AFM tip, is the transfer of polymers by rubbing a solid pin (this “reservoir of molecules” is glassy or partially crystalline) onto a solid substrate.

The objective is to demonstrate that amorphous *i*-PS chains can be oriented, either by generating a film via friction transfer or by applying an external friction force on polymers in an amorphous isotropic film. Subsequently, the consequences of such preferential orientation on crystallization will be established. In the first case, a friction-transfer approach (moving a pin of solid *i*-PS across a smooth silicon counterface) was used to deposit a molecularly thin layer of highly oriented *i*-PS chains. In the latter case, an AFM tip is moved at room temperature (i.e., well below the glass transition of *i*-PS) on the surface of a spin-coated *i*-PS film. After a subsequent annealing step, the resulting crystallization patterns were analyzed. The goal was to understand how the friction-induced chain orientations influence crystal nucleation and the orientation of the resulting crystalline lamellae.

24.4.1 Stretched Chains Resulting from Friction Transfer

During friction transfer, a thin amorphous film will be produced which consists of polymer chains preferentially aligned in the rubbing direction. Due to their non-uniform orientation, such molecules have a high tendency to crystallize, that is, the energy barrier for nucleation is significantly reduced. As an example, one may consider the generation of oriented crystals of *i*-PS in films prepared by friction transfer [61]. Polystyrene (PS) is a widely used

engineering thermoplastic with good thermal stability. Three steric isomers of polystyrene exist which include the crystallizable isotactic and syndiotactic forms, and the non-crystallizable atactic polystyrene. The frictional properties of atactic polystyrene (amorphous) have been extensively investigated in the literature, contrary to *i*-PS. Surface plastic deformation of atactic polystyrene can appear under nano and macro tribological conditions. The formation of abrasive wear during sliding of a hard ball on atactic polystyrene, with macroscopic cracks [62] or meandering groovelike structures [63] perpendicular to the sliding direction, was observed. Rubbed films also exhibit molecular reorientation, with the vinyl group perpendicular to the rubbing direction, the planes of the phenyl side groups being reoriented perpendicular to the para direction [64]. Polarization-modulation infrared reflection-absorption spectroscopy (PM-IRRAS) has shown that PS chains deposited by friction transfer exhibit an anisotropic orientation of their main backbone preferentially perpendicular to the friction direction [64, 65].

The influence of stereochemistry of polymers on a variety of thermal and morphological factors has motivated many investigations on conformation and crystallization kinetics [66]. In polymer processing, the final properties depend directly on the morphology, in particular on anisotropic crystalline structures. Thus, controlling macroscopic orientation of small molecular and polymeric materials represents an important step for the improvement of their physical properties [67–69]. In most cases of polymer crystallization, it is difficult to avoid nucleation as it may be too fast or too many nuclei may appear simultaneously. However, *i*-PS crystallizes slowly even under high cooling rate and thus can easily be obtained in an amorphous state by quenching below the glass transition temperature [70, 71]. When *i*-PS is crystallized from molten thin films, it was possible to observe crystal morphology changes with crystallization temperature [72]. Above 200°C, crystals grew into hexagonal lamellar plates, while around 195°C rounded hexagonal patterns, at 180°C circular disc-like structures and below 170°C 2D spherulites were found [72, 73].

So far, the possible correlation between friction (i.e., the sliding direction) and the resulting orientation of polymer crystals was not yet studied in much detail. In particular, the effect of a mechanical stress on the direction of crystal growth in *i*-PS has not been investigated. In bulk samples, the application of extensional deformation or shear flow can indeed have a strong influence on chain conformation and polymer crystallization [74, 75]. Such oriented polymer melts can generate highly anisotropic crystalline features, such as “shish-kebab” and trans-crystalline superstructures [76]. Shish-kebab structures consist of an elongated long central core (shish) onto which periodically lamellar crystals (kebab) are attached [77, 78]. Different methods for obtaining oriented polymer chains can be used such as stretching, mechanical rubbing, or via a liquid crystalline precursor state [79–88, 90–92]. The possibility to orient the polymer backbone by friction has been employed by Tabor [93] and Wittmann and Smith [94].

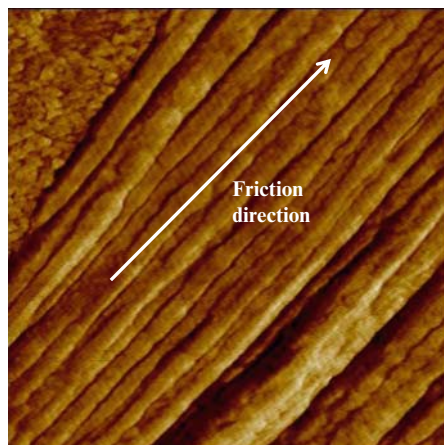


Fig. 24.19. Morphologies of i-PS layer obtained by friction transfer at low load (near to zero), a sliding velocity of 5 cm s^{-1} at $T = 95^\circ\text{C}$ ($T \sim T_g$) with a cyclohexanone drop in the contact region (AFM phase image: $1.23 \times 1.23 \mu\text{m}^2$)

The friction-transfer technique is interesting to obtain thin films of insoluble polymers and also to orient polymer chains during such deposition. However, the influence of such preferential chain orientation on the probability of crystal nucleation and the possibility of controlling the crystal growth direction was never investigated.

Transfer of a thin layer of i-PS was achieved by moving this solid i-PS pin (3 mm in diameter, obtained by compression-molding) on a smooth silicon wafer at a controlled rate, temperature, and applied pressure, using a pin-on disk tribometer. Sliding at a temperature close to T_g in the presence of a very small quantity of cyclohexanone in the contact region between pin and substrate resulted in a thin transferred layer of i-PS, which exhibited clear signs of alignment of extended fiber-like objects. This transferred i-PS film was observed by AFM after rubbing (Fig. 24.19). Ribbons of i-PS (width between 60 and 90 nm, thickness of approximately 17 nm, comparable to the radius of gyration of this polymer) aligned along the friction direction were observed. This clear orientation of the observed surface morphology indicated strong deformations of the transferred i-PS chains.

Isothermal crystallization of the transferred film was performed by annealing the transferred film for 12 h at 210°C in an inert nitrogen atmosphere. The resulting morphology of the crystalline structure was again determined at ambient conditions by AFM in the tapping mode. Such isothermal crystallization at 210°C of the transferred i-PS layer induced the growth of crystals oriented along the sliding direction as shown in Fig. 24.20. It is well known that applying an external extensional or shear field can have a strong impact on the crystallization behavior of semi-crystalline polymer [95, 96]. In the present case, frictional shear caused extensive deformation of polymer chains

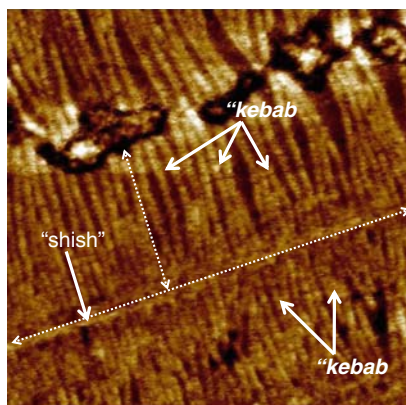


Fig. 24.20. AFM image for an i-PS film prepared by friction transfer and subsequently crystallized for 12 h at 210°C (size: 900 × 900 nm²)

in the sliding direction. As a consequence, the crystals generated from this transferred polymer layer displayed a clear shish-kebab pattern with a longitudinal axis (shish) preferentially aligned in the sliding direction similar to the situation of shish-kebab structures in the case of melt extrusion [97, 98]. The lamellae (kebab) exhibited a constant width (16 nm), a value fully consistent with the height of flat-on i-PS crystals grown from spin-coated thin films [72]. Contrary to these flat-on crystals obtained in spin-coated films, friction-transfer rather favors the formation of shish-kebab crystals formed by edge-on lamellae. However, if the film was deposited at a higher temperature (150°C) in the presence of cyclohexanone, the probability for the formation of shish-kebab structures was strongly reduced. As illustrated in Fig. 24.21, under such conditions preferentially flat-on hexagonal crystals were observed. Only very few shish-kebab crystals were formed.

24.4.2 Stretched Chains Resulting from Rubbing with an AFM Tip

Above results strongly suggest that friction-transfer allows aligning not only the transferred polymer chains but also the subsequently forming crystalline domains within these transferred films. Thus, a relation seems to exist between chain extension and nucleation probability for edge-on lamellae emanating from a shish backbone. Based on this hypothesis, in a second series of experiments, polymer chain conformations were deformed locally and in a controlled way by moving an AFM tip across the surface of spin-coated i-PS films [99].

In a first attempt, repeated scratching with an AFM tip in contact at a moderate applied load favored plastic deformation and wear of the amorphous thin film (Fig. 24.22). After isothermal crystallization at 210°C, exactly at the edge of the scratched area, an extremely high nucleation density of edge-on

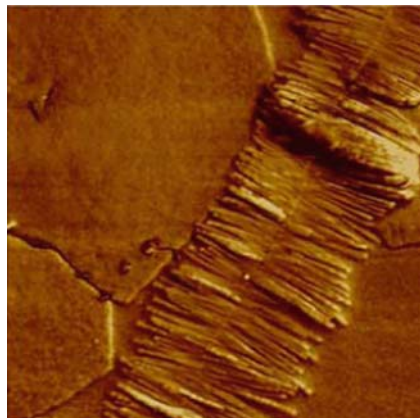


Fig. 24.21. AFM image of i-PS crystals, showing the coexistence of hexagonal flat-on crystals and shish-kebab crystals (i-PS film prepared by friction transfer at 150°C in the presence of cyclohexanone in the contact region) and crystallized for 12 h at 220°C (size $4 \times 4 \mu\text{m}^2$)

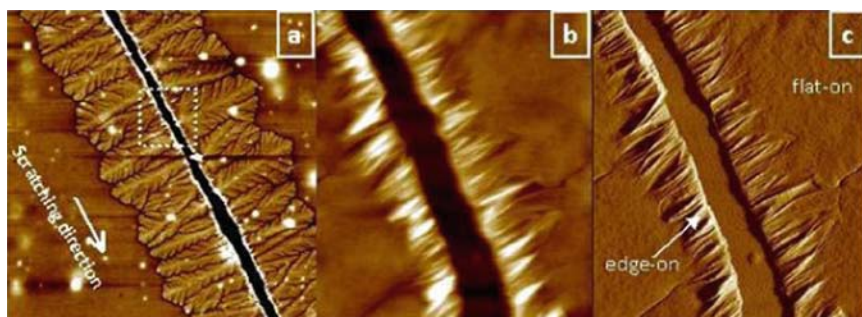


Fig. 24.22. AFM images showing the orientation of the i-PS crystals lamellae obtained after isothermal crystallization of the scratched thin film (repeated scratching along the same line). (a) Topographic image (size: $45 \times 45 \mu\text{m}^2$; height range: 30 nm). The branched features originate from flat-on lamellar crystals; (b, c) topographic and phase mode image of the region indicated by the *dotted box* in (a) (size: $5 \times 5 \mu\text{m}^2$; height range: 60 nm) representing edge-on lamellae, continued by flat-on lamellae, oriented perpendicular to the scratching direction

crystals was observed (Fig. 24.23). The nucleation density was at its maximum possible value. The distance between edge-on lamellae directly at the edge of the scratched region was only about 15 nm which is comparable to the thickness of a lamellar crystal. Intriguingly, all crystals grew perpendicular to the scratching direction. However, at a distance of about $1 \mu\text{m}$ from the edge of the scratched region, a transition from edge-on to flat-on lamellae occurred (Fig. 24.22). The origin of this transition is so far not known.

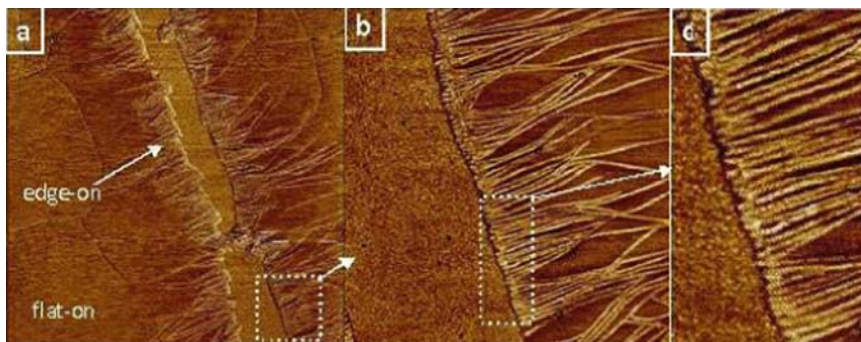


Fig. 24.23. AFM phase images demonstrating the high nucleation density of edge-on lamellae at the edge of the scratched region; Sizes: (a) $6 \times 6 \mu\text{m}^2$, (b) $1.5 \times 1.5 \mu\text{m}^2$, (c) $570 \times 250 \text{nm}^2$

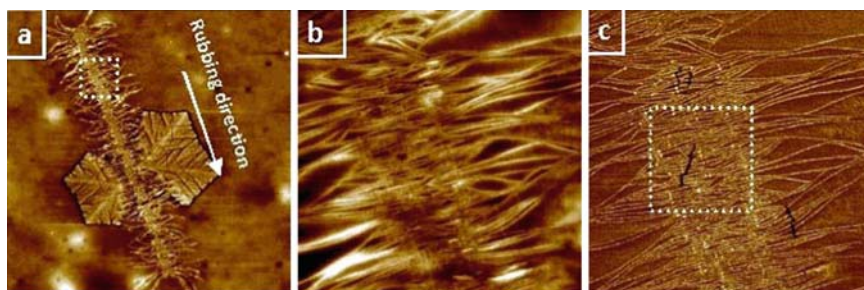


Fig. 24.24. AFM image showing the orientation of the i-PS crystalline lamellae obtained after isothermal crystallization of the rubbed thin film (one scratch on the same line at very low applied load). (a) Topographic mode ($40 \times 40 \mu\text{m}^2$ of size and 20 nm in height); (b) topographic mode ($20 \times 20 \mu\text{m}^2$ of size and 20 nm in height); (c) topographic and phase mode of the region indicated by the *dotted box* in B ($2 \times 2 \mu\text{m}^2$ of size and 13 nm in height) representing the oriented edge-on lamellae perpendicular to the *rubbed line*

In order to investigate the influence of the applied force and to avoid plastic deformation, the load on the tip was reduced. Soft rubbing of the polymer surface with an AFM tip (only once in a single direction) at a very low applied load favors the elastic deformation of the polymers on the amorphous thin film. Under such conditions, very little plastic deformation occurred as demonstrated by no more than very weak changes in topography of the rubbed film (Fig. 24.24). Interestingly, even without any visible deformation of the film, the AFM tip nonetheless was obviously causing changes in the chain conformations. This became visible when isothermally crystallizing the rubbed film. Exclusively in the rubbed region edge-on lamellae appeared (Fig. 24.24). The distance between these lamellae was again very small, indicating that even soft rubbing was sufficient to significantly lower the nucleation barrier. In contrast to hard rubbing (scratching), the edge-on lamellae did not switch

to flat-on lamellae (except for a single event seen in Fig. 24.24a). In addition, it also became obvious that rubbing controlled the orientation of these edge-on lamellae. As can be clearly seen from Fig. 24.24b, in the rubbed region all lamellae were oriented perpendicular to the rubbing direction. Outside this region, the direction of these lamellae started to change and eventually most lamellae stopped to grow as they were coalescing with other lamellae.

These few exemplary results already impressively demonstrate that an AFM tip, even when it is interacting only rather little with the surface, is able to affect the conformations of polymers. Of course, such deformation will remain and affect the nucleation probability for crystal growth only if the polymers are sufficiently immobile so that their conformations cannot relax back to the equilibrium state.

24.5 Conclusions: Controlling Polymer Properties at a Molecular Scale

In summary, using the advantages of the atomic force microscopy (non-destructive technique, high spatial resolution, possibility to perform in situ studies at elevated temperatures, etc.) for real-time observation of polymer samples, crystallization, and melting processes in nanometer-sized polymer crystals confined in geometries such as the micro-phase separated asymmetric diblock copolymers can be studied in detail. The combination of the advantages of block copolymers (ordered self-assembled mesostructures with a tunable nanometer length scale) with the possibility of TM-AFM to interact locally with a sample represents a powerful tool to study melting kinetics of confined polymer nano-crystals. Besides the independent melting (without any correlation between the individual mesophase compartments), it is also possible to observe, although the amorphous matrix presents an enormous barrier hindering exchange of crystallizable blocks between neighboring compartments, a non-zero probability for exchange of material between the molten and the crystalline mesophase compartments (crystals were feeding with material from the molten compartments). Moreover, atomic force microscopy can be used to melt confined nanometer-sized polymer crystals “on demand” by using hot AFM probe (high temperature AFM). This approach allows manipulating mechanical properties of materials in small volumes and opens up an additional possibility for surface engineering controlled on a nanometer scale. A possibility for local modification of materials properties with heated proximity probe provides an extra dimension in obtaining custom-made nano-structured non-periodic materials. In addition, as the process of melting is reversible, it will also be possible to perform the experiments in a complementary way. Molten cells could be crystallized individually, for example, using a cold tip to individually nucleate crystallization in a cell.

As AFM technique progresses toward faster scan rates, easier use at elevated temperatures and lower imaging forces (lower noise), one can imagine a

possible extension toward the investigation of the nucleation process in confined geometries or a local melting technique in combination with well-aligned crystallized cylindrical mesophase domains. We expect that further systematic studies of this kind will help us to obtain even more profound results on how polymer crystals are nucleated, how they grow and to envisage various applications of those processes. The above demonstrated correlation between chain deformation induced by rubbing polymers and an increased nucleation probability represents a first step in this direction.

Thus, one can doubtlessly conclude that AFM represents a powerful tool for manipulating polymers and their mechanical properties in small volumes. This opens up an additional possibility for surface engineering with a control on a nanometer scale.

References

1. G. Binnig, C.F. Quate, C. Gerber Phys. Rev. Lett. **56** 930 (1986)
2. T.J. McMaster, J.K. Hobbs, P.J. Barham, M.J. Miles Probe Microsc. **1** 43 (1997)
3. J.K. Hobbs, T.J. McMaster, M.J. Miles, P.J. Barham Polymer **39** 2437 (1998)
4. C. Harrison, D.H. Adamson, Z. Cheng, J.M. Sebastian, S. Sethuraman, D.A. Huse, R.A. Register, P.M. Chaikin Science **290** 1558 (2000)
5. R. Pearce, G.J. Vancso Polymer **39** 1237 (1998)
6. Y.K. Godovsky, S.N. Magonov Langmuir **16** 3549 (2000)
7. Y.K. Godovsky, S.N. Magonov Polym. Sci. Ser. A **43** 647 (2001)
8. C. Basire, D.A. Ivanov Phys. Rev. Lett. **85** 5587 (2000)
9. J.K. Hobbs, M.J. Miles Macromolecules **34** 353 (2001)
10. J.K. Hobbs, A.D.L. Humphris, M.J. Miles Macromolecules **34** 5508 (2001)
11. R. Pearce, G.J. Vancso J. Poly. Sci. B **36** 2643 (1998)
12. L.G.M. Beekmans, D.W. van der Meer, G.J. Vancso Polymer **43** 1887(2002)
13. G. Reiter, G. Castelein, P. Hoerner, G. Riess, A. Blumen, J.U. Sommer Phys. Rev. Lett. **83** 3844 (1999)
14. A.J. Ryan, W. Hamley, W. Bras, F.S. Bates Macromolecules **28** 3860 (1995)
15. Y.L. Loo, R.A. Register, A.J. Ryan Phys. Rev. Lett. **84** 4120 (2000)
16. C. De Rosa, C. Park, E.L. Thomas, B. Lotz Nature **405** 433 (2000)
17. Y.L. Loo, R.A. Register, D.H. Adamson Macromolecules **33** 8361 (2000)
18. Y.L. Loo, R.A. Register, D.H. Adamson J Poly Sci B **38** 2564(2000)
19. H.L. Chen, S.C. Hsiao, T.L. Lin, K. Yamauchi, H. Hasegawa, T. Hashimoto Macromolecules **34** 671 (2001)
20. L. Zhu, S.Z.D. Cheng, B.H. Calhoun, G. Ge, R.P. Quirk, E.L. Thomas, B.S. Hsiao, F. Yeh, B. Lotz Polymer **42** 5829 (2001)
21. G. Reiter, G. Castelein, J.U. Sommer, A. Röttele, T. Thurn-Albrecht Phys. Rev. Lett. **86** 5918 (2001)
22. A. Taden, K. Landfester Macromolecules **36** 4037 (2003)
23. J.D. Hoffman, G.T. Davis, J.I. Lauritzen, in *Treatise on Solid State Chemistry*, ed. by N.B. Hannay. Crystalline and Noncrystalline Solids, vol 3 (Plenum Press, New York, 1976), pp. 497–614
24. J.D. Hoffman, R.L. Miller Polymer **38** 3151 (1997)
25. H.D. Keith, F.J. Padden J. Polym. Sci. **39** 101 (1959)

26. A. Keller *J. Polym. Sci.* **39** 151 (1959)
27. F.P. Price *J. Polym. Sci.* **39** 139 (1959)
28. A. Keller *Philos. Mag.* **2** 1171 (1957)
29. P.J. Barham, R.A. Chivers, A. Keller, J. Martinez-Salazar, S.J. Organ *J. Mater. Sci.* **20** 1625 (1985)
30. S.J. Organ, A. Keller *J. Mater. Sci.* **20** 1571 (1985)
31. G. Strobl *Eur. Phys. J. E* **3** 165 (2000)
32. M. Muthukumar, C.K. Ober, E.L. Thomas *Science* **277** 1225 (1997)
33. I.W. Hamley *The Physics of Block Copolymers* (Oxford University Press, Oxford, 1998)
34. F.S. Bates, G.H. Fredrickson *AIP Phys. Today* **2** 32 (1999)
35. T. Goldacker, V. Abetz, I. Stadler, Y. Erukhimovich, L. Leibler *Nature* **398** 137 (1999)
36. U. Jeong, H.C. Kim, R.L. Rodriguez, L.Y. Tsai, C.M. Stafford, J.K. Kim, C.J. Hawker, T.P. Russell *Adv. Mater.* **14** 274 (2002)
37. G. Krausch, R. Magerle *Adv. Mater.* **14** 1579 (2002)
38. O. Ikkala, G. Ten Brinke *Science* **295** 2407 (2002)
39. G. Riess *Prog. Polym. Sci.* **28** 1107 (2003)
40. L. Wang *Appl. Phys. Lett.* **73** 3781 (1998)
41. L. Nony, R. Boisgard, J.P. Aimé *J. Chem. Phys.* **111** 1615 (1999)
42. J. Tamayo, R. García *Langmuir* **12** 4430 (1996)
43. R. García, R. Pérez *Surf Sci Rep* **47** 197 (2002)
44. A. Knoll, R. Magerle, G. Krausch *Macromolecules* **34** 4159 (2001)
45. F. Dubourg, J.P. Aimé, S. Marsaudon, R. Boisgard, P. Leclère *Eur. Phys. J. E* **6** 49 (2001)
46. F. Dubourg, S. Kopp-Marsaudon, P. Leclère, R. Lazzaroni, J.P. Aimé *Eur Phys J E*, **6** 387 (2001)
47. D. Ivanov, S. Magonov in *Polymer Crystallization: Observations, Concepts and Interpretations*, ed. by J.-U. Sommer, G. Reiter (Springer, Heidelberg, 2003), pp. 98–130
48. N. Rehse, A. Knoll, M. Konrad, R. Magerle, G. Krausch *Phys. Rev. Lett.* **87** 035505 (2001)
49. J.H. Lim, C.A. Mirkin *Adv. Mater.* **14** 1474 (2002)
50. Y. Okawa, M. Aono *Nature* **409** 683 (2001)
51. G. Binnig, M. Despont, U. Drechsler, W. Häberle, M. Lutwyche, P. Vettiger, H.J. Mamin, B.W. Chui, T.W. Kenny *Appl. Phys. Lett.* **74** 1329 (1999)
52. P. Vettiger, M. Despont, U. Drechsler, U. Durig, W. Häberle, M. Lutwyche, H.E. Rothuizen, R. Stutz, R. Widmer, G.K. Binnig *IBM J. Res. Dev.* **44** 323 (2000)
53. T.H. Fang, W.J. Chang *J. Phys. Chem. Solids* **64** 913 (2003)
54. Y.W. Cneh, J.C. Li, Z.Q. Xue *Appl. Phys. A* **77** 379 (2002)
55. R. Schmidt, G. Haugstad, W.L. Gladfelter *Langmuir* **19** 898 (2003)
56. H. Pernot, M. Baumert, F. Court, L. Leibler *Nat. Mater* **1** 54 (2001)
57. J. Hobbs *Polymer Crystallization: Observations, Concepts and Interpretations*, ed. by J.-U. Sommer, G. Reiter. *Lecture Notes in Physics*, (Springer, Heidelberg, 2003) pp. 82–97
58. H. Schönherr, L.E. Bailey, C.W. Frank *Langmuir* **18** 490 (2002)
59. A. Röttele, T. Thurn-Albrecht, J.U. Sommer, G. Reiter *Macromolecules* **36** 1257 (2003)
60. R. Höper, T. Gesang, W. Possart, O.D. Hennemann, S. Boseck *Ultramicroscopy* **60** 17 (1995)

61. K. Jradi, S. Bistac, M. Schmitt, G. Reiter, *Polymer* **50** 3724 (2009)
62. T. Aoiike, H. Uehara, T. Yamanobe, T. Komoto *Langmuir*, **17** 2153 (2001)
63. S.W. Lee, J. Yoon, H.C. Kim, B. Lee, T. Chang, M. Ree *Macromolecules* **36** 9905 (2003)
64. T. Elzein, V. Kreim, S. Bistac *Journal of Polymer Science, Part B* **44** 1268 (2006)
65. T. Elzein, V. Kreim, S. Ghorbal A Bistac *Journal of Polymer Science, Part B* **44** 3272(2006)
66. C. Wang, C.C. Lin, L.C. Tseng *Polymer* **47** 390 (2006)
67. H. Chen, H. Xu, P. Cebe *Polymer* **48** 6404 (2007)
68. H. Xu, P. Cebe *Polymer* **46** 8734 (2005)
69. T. Liu, J. Petermann *Polymer* **42** 6453 (2001)
70. A. Mamun, S. Umemoto, N. Ishihara, N. Okui *Polymer* **47** 5531 (2006)
71. A.A. Minakov, D.A. Mordvintsev, R. Tol, C. Schick *Thermochimica Acta* **442** 25 (2006)
72. K. Taguchi, H. Miyaji, K. Izumi, A. Hoshino, Y. Miyamoto, R. Kokawa *Polymer* **42** 7443 (2001)
73. M. Yamato, T. Kimura *Sci. Technol. Adv. Mater.* **7** 337 (2006)
74. H. Janeschitz-Kriegl, E. Ratajski, M. Stadlbauer *Rheol. Acta* **42** 355 (2003)
75. M. Stadlbauer, H. Janeschitz-Kriegl, G. Eder, E. Ratajski *J. Rheol.* **48** 631 (2004)
76. H. Fukushima, Y. Oqino, K. Nishida, T. Kanaya *Polymer* **46** 1878 (2005)
77. J.K. Hobbs, A.D.L. Humphris, M.J. Miles *Macromolecules* **34** 5508 (2001)
78. K. Wang, F. Chen, O. Zhang, O. Fu *Polymer* **49** 4745 (2008)
79. T. Granier, E.L. Thomas, D.R. Gagnon, F.E. Karasz, R.W. Lenz *J. Polym. Sci. B:Polym. Phys.* **24** 2793 (1986)
80. D.D.C. Bradley *J. Phys. D: Appl. Phys.* **20** 1389 (1987)
81. D. Chen, M.J. Winokur, M.A. Masse, F.E. Karasz *Polymer* **33** 3116 (1992)
82. P.F. van Hutten, J. Wildeman, A. Meetsma, G.J. Hadzioannou *Am. Chem. Soc.* **121** 5910 (1999)
83. C.Y. Yang, K. Lee, A.J. Heeger *J. Mol. Struct.* **521** 315 (2000)
84. R. Resel, R. Kiebooms, D. Vanderzande, F. Stelzer *Monatsh. Chem.* **132** 433 (2001)
85. S.H. Chen, A.C. Su, H.L. Chou, K.Y. Peng, S.A. Chen *Macromolecules* **37** 167 (2004)
86. S.H. Chen, A.C. Su, S.R. Han, S.A. Chen, Y.Z. Lee *Macromolecules* **37** 181 (2004)
87. S.H. Chen, C.H. Su, A.C. Su, S.A. Chen *J. Phys. Chem. B* **10** 8855 (2004)
88. T.W. Hagler, K. Pakbaz, J. Moulton, F. Wudl, P. Smith, A. Heeger *J. Polym. Commun.* **32** 339 (1991)
89. M. Hamaguchi, K. Yoshino *Jpn. J. Appl. Phys.* **34** L712 (1995)
90. P. Dyreklev, M. Berggren, O. Inganas, M.R. Andersson, O. Wennestrom, T. Hjertberg *Adv. Mater.* **7** 43 (1995)
91. M. Grell, D.D.C. Bradley *Adv. Mater.* **11** 895 (1999)
92. A. Bolognesi, C. Botta, M. Martinelli, W. Porzio *Org. Electron.* **1** 27 (2000)
93. C.M. Pooley, D. Tabor *F.R.S Proc. R. Soc. Lond. A.* **329** 251 (1972)
94. J.C. Wittmann, P. Smith *Nature* **352** 414 (1991)
95. K. Wang, F. Chen, O. Zhang, O. Fu *Polymer* **49** 4745 (2008)
96. S. Yan, J. Petermann *Polymer* **41** 6679 (2000)
97. J.A. Kornfield, G. Kumaraswamy, A.M. Issaian *Ind. Eng. Chem. Res.* **41** 6383 (2002)

98. G. Kumaraswamy, J.A. Kornfield, F. Yeh, B.S. Hsiao *Macromolecules* **35** 1762 (2002)
99. K. Jradi, S. Bistac, M. Schmitt, A. Schmatulla, G. Reiter, *Eur. Phys. J. E* **29** 383 (2009)

Application of Contact Mode AFM to Manufacturing Processes

Michael A. Giordano and Steven R. Schmid

Summary. A review of the application of contact mode atomic force microscopy (AFM) to manufacturing processes is presented. A brief introduction to common experimental techniques including hardness, scratch, and wear testing is presented, with a discussion of challenges in the extension of manufacturing scale investigations to the AFM. Differences between the macro- and nanoscales tests are discussed, including indentation size effects and their importance in the simulation of processes such as grinding. The basics of lubrication theory are presented and friction force microscopy is introduced as a method of investigating metal forming lubrication on the nano- and microscales that directly simulates tooling/workpiece asperity interactions. These concepts are followed by a discussion of their application to macroscale industrial manufacturing processes and direct correlations are made.

Key words: Abrasive machining, Atomic force microscope, Chemical mechanical polishing, Contact mode, Hardness testing, Indentation, Lubrication, Manufacturing, Nanolithography, Plowing, Polymer processing, Powder processing.

25.1 Introduction

The atomic force microscope (AFM) has a long history of use as a research tool for a wide variety of industries. In recent years, the applications of AFM in microelectromechanical systems (MEMS) and biological applications have become well developed, as has the use of the AFM to conduct evaluations of boundary lubricants, evaluate material properties, and perform fundamental measurements of friction. Of course, the AFM is most widely used to image surfaces in tapping mode; perhaps the industry application where the AFM demonstrates its use most convincingly is in the imaging of integrated circuits; there are many surface defects such as “haze” that are visible with the naked eye but difficult to discern with conventional profilometry, whereas the AFM allows measurement, quantification, and differentiation of such surfaces.

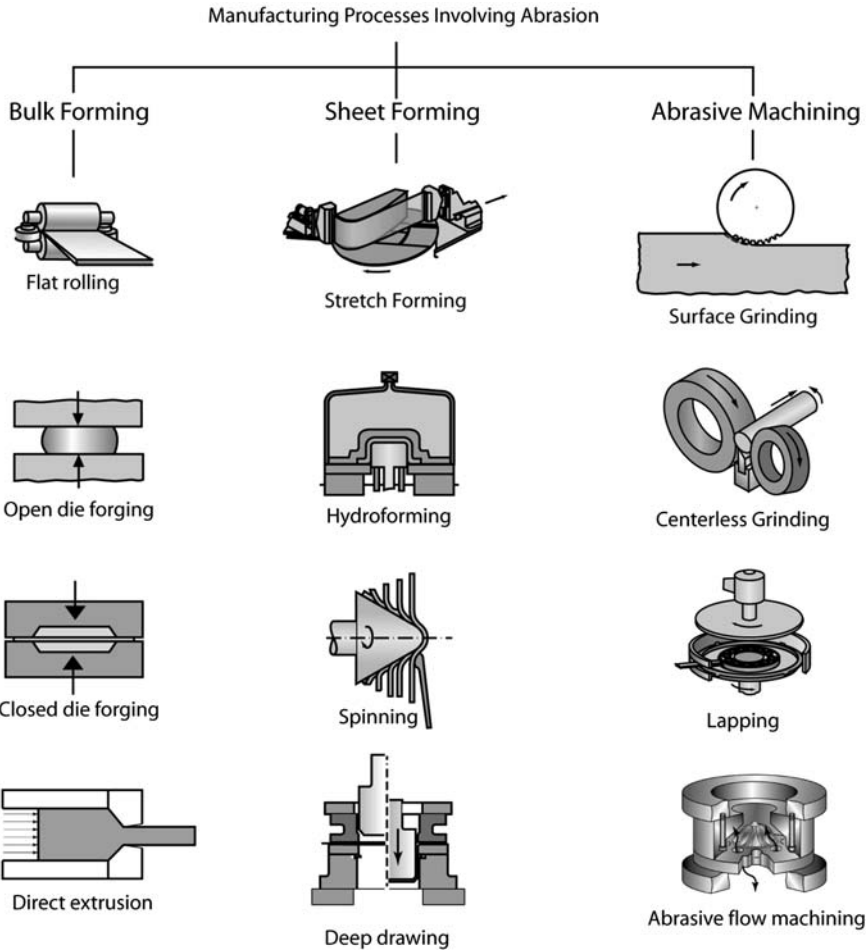


Fig. 25.1. Manufacturing processes involving asperity contact between tooling and workpiece [1]

In contact mode, the AFM is useful in the simulation of a fundamental problem that is widely encountered in manufacturing processes: the plowing of a hard asperity into a softer opposing surface. Such circumstances are widely encountered, as shown in Fig. 25.1. In forming operations, it is not uncommon that an asperity on a tool surface will lead to plowing in the workpiece; the plowing can change direction in rolling as the asperity passes the neutral plane. Machining and abrasive finishing operations rely on abrasion for material removal and processes such as lapping and abrasive flow machining depend on individual asperity interactions.

Manufacturing operations have not been exhaustively simulated by the AFM to date. Real manufacturing operations usually involve high

temperatures, rough and dynamic surfaces, high strain rates, substrate strains that soften asperities, and demanding lubrication regimes [2]. Hydrodynamic lubrication is to be avoided because of orange peel, whereas boundary or mixed lubrication with significant interactions between lubricant, tooling, and work-piece is the normal condition. It is difficult for the AFM to serve as more than a bench scale simulation tool, but this has proven worthwhile in a number of applications to date and the AFM has helped researchers improve their fundamental understanding of many manufacturing processes.

This chapter reviews the capabilities of the AFM in contact mode as applicable to manufacturing processes. Numerous researchers have developed AFM capabilities for situations far removed from manufacturing practice, but this does not diminish the value of the fundamental approaches.

25.2 Review of Atomic Force Microscope Capabilities Relevant to Manufacturing

This section describes the test methods applicable to manufacturing. Hardness tests by indentation, scratch tests, and wear tests all have direct correlations to many manufacturing processes. Although many researchers have focused their work on thin films and coatings, manufacturing applications must deal with the presence of lubricants, work-hardened surfaces, high temperatures and pressures, and nonidealized situations. Even so, hardness, scratch, and wear tests have direct applicability to manufacturing situations.

25.2.1 Evaluation of Mechanical Properties

Hardness Testing

Hardness is the resistance of a surface to plastic deformation and is traditionally defined as the pressure beneath a blunt indenter sufficient to cause plastic flow [3]. Slip-line field models suggest that the hardness is three times the flow stress of a rigid, perfectly plastic material [4]. Materials with little strain hardening or metals that have undergone significant cold work are closely approximated by rigid perfectly plastic behavior, but this is not the general case. Often, experimental characterization of a material with a given processing history can give correlations between hardness and flow strength [1].

Hardness as a concept must be distinguished from hardness number, which is a geometry and load-dependant measure of hardness. It has long been known that indentation techniques can be used to measure the hardness of a material [3, 5–12]. In 1812, Mohs [13] was the first to rank materials according to *relative* hardness. Those which would scratch other materials were labeled harder than the rest, and the Mohs scale is still commonly used in geology. Modern

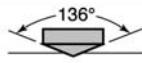

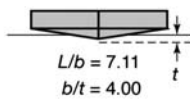
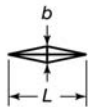


Test	Indenter	Shape of indentation		Load, P
		Side view	Top view	
Vickers	Diamond pyramid			1–120 kg
Knoop	Diamond pyramid			25 g–5 kg
AFM	Diamond pyramid			10 nN–10 μ N

Fig. 25.2. Indentation tip geometries [1]

hardness testing has advanced greatly since that time, has become standardized internationally, and has specialized conditions for different materials. The most common microscale hardness tests are the Vickers and Knoop tests. The main differences between these tests are the load applied and the shape of the indenter involved, as illustrated in Fig. 25.2. The Knoop test is based on a rhomboidal pyramid-shaped indenter, whereas the Vickers indenter has a square base. The diamond tip on an AFM probe is a triangular pyramid, which is similar to one-half of a Knoop indenter, properly referred to as a Berkovich indenter, which is a less common hardness test in practice.

Hardness can be determined from [3]:

$$H = \frac{P_{\max}}{A_c} \tag{25.1}$$

where P_{\max} is the maximum load applied during the hardness test and A_c is the projected area of contact. Indentation experiments can be conducted on a small length scale (a few tens of nanometers to a few micrometers of penetration depth) in an AFM. Experimental error can be introduced through inaccurate determination of the applied load and the area of contact during the hardness test. Determination of the applied load is discussed below, but requires accurate knowledge of the cantilever stiffness, which can be determined through beam theory [14–19], finite element analysis [20–23], static loading [16,24–31], dynamic response [15,17,32,33], pendulum analysis [16,24], thermal spectrums [34], or experimentally [35]. Reviews of selected methods are presented by Burnham et al. [36], Cumpson et al. [30], and Gibson et al. [37,38]. All of these approaches have errors and uncertainties associated with them. For example, as can be seen in Fig. 25.3, there is significant adhesive used to affix a diamond tip to a stainless steel cantilever; the composite beam stiffness is difficult to calculate.

One of the most difficult aspects of calculating hardness in practice is the determination of the actual contact area between the indenter and the material

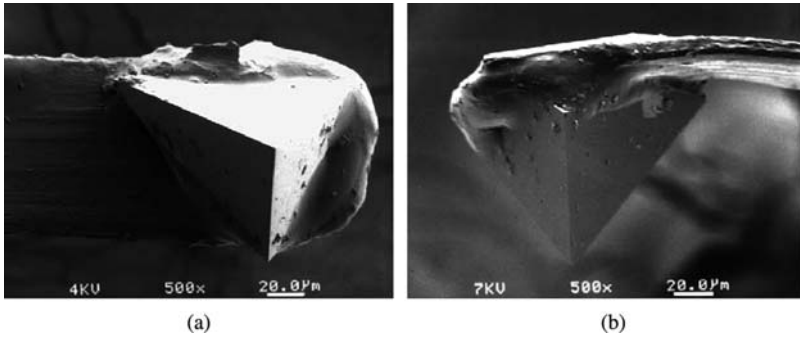


Fig. 25.3. Typical diamond indentation probe for AFM use. (a) Tip view and (b) leading edge view [39]

surface. For a true hardness calculation, the projected area of contact under load should be used. Since measurement of the area generally occurs after unloading, this measurement is compromised by any recovery in the material, friction, etc., as discussed below.

The hardness test currently supported by AFM control software is user-friendly and typically requires the following procedure [40]:

1. Image the desired area in TappingMode with a silicon tip.
2. Equip the AFM with a diamond-tipped probe.
3. Engage the surface in TappingMode.
4. Set values for indentation parameters (most importantly indentation set point).
5. Set values for automatic indentation parameters (grid size, spacing, set point increase, etc.).
6. Perform the indentation while capturing force–displacement data.
7. Retract the probe and remove the diamond-tipped cantilever.
8. Re-engage surface and image the indentation in TappingMode with a silicon tip.

This procedure results in force–displacement data acquired during the hardness test and the shape of the residual impression left in the surface. The force exerted on the sample is given by:

$$P_{\max} = S\delta k \quad (25.2)$$

where P_{\max} is the load and usually expressed in nano- or micronewtons, S is the user-specified $(A - B)/(A + B)$ photodetector set point in units of volts, δ is the sensitivity of the piezotube, in units of length per volts, and k is the normal cantilever stiffness, in force per displacement. The area is directly obtained from the tapping mode image of the residual indent, and hardness is then calculated with (25.1).

Most hardness test probes for use in an AFM consist of a single crystal diamond tip adhesively bonded to the end of a stainless steel cantilever as seen in Fig. 25.3. Stiffness and resonating frequency of the probe can be controlled by changing the length of the cantilever and the position of the diamond tip on the cantilever.

Results from AFM indentations typically display an indentation size effect (ISE). Early observers of this phenomenon include Onitsch [41], Maugis et al. [42], Pashley et al. [43], Pharr and Oliver [44], Gane and Bowden [45], and manufacturing researchers investigating grinding processes long before micro/nano-testing methods were available. Marshall et al. [46] and Backer et al. [47] commented on the unusually high forces needed during the grinding process. Marshall et al. expressed the specific energy in grinding as

$$u = \frac{\pi DNH}{12vbd} \quad (25.3)$$

where u is the energy per unit volume, D is the diameter of the grinding wheel, N is the wheel speed in rotations per minute, v is the table velocity, b is the width of the specimen, and d is the depth of cut. Marshall et al. found that for depths of cut larger than $25\ \mu\text{m}$, the specific energy decreased. These results agreed with single point cutting experiments which showed that the required energy needed to remove material decreased for larger depths of cut.

According to chip formation models by Merchant [48], the shear stresses present during grinding were much greater than the yield strength of the bulk workpiece material. In addition, it was found that these shear stresses increased as the size of the grinding wheel grains decreased. To explain these results, a size effect was suggested similar to that found in other material tensile tests, such as very thin wires [49], piano wire, and glass fibers [50].

Meyer [51] formulated the following relationship to explain the size effect observed in indentation:

$$P = kd^n \quad (25.4)$$

where P is the applied load, d is the diameter of the spherical indenter (or half the diagonal of a Vickers indenter), and k and n are constants. McHargue [52] offers a relationship for hardness:

$$H = ch^{m-2} \quad (25.5)$$

where H is the measured hardness, h is the indentation depth, and c and m are constants.

McHargue offers a thorough explanation for the reasons that an ISE exists, including:

- The indentation images in tapping mode may not reflect the projected area of contact in the hardness test because of elastic and viscoelastic recovery of the surface [53–57]. For example, indentations of PMMA exhibit severe recovery, and the projected triangle will have severely curved sides.

- The indentation shape can be affected by surface residual stresses.
- Depending on the depth of penetration and the shape of the indenter, it is possible for friction between the indenter tip and substrate to affect projected area. Blunt indenters have dead metal zones beneath them and are less susceptible to friction [4]. Since diamond-tipped cantilevers have a finite tip radius, they can act blunt or sharp depending on the depth of penetration achieved. Thus, compensating for friction is a challenging problem in hardness testing [58–61].
- Slip-line field theory provides the fundamental basis for correlating hardness to flow stress [4], but slip-line field theory is difficult for strain hardening materials, although it has been attempted [62]. Thus, postyield behavior of the material can make the correlation of hardness experiments to material mechanical properties difficult.
- The substrate needs to be around an order of magnitude deeper than the probe penetration in order to accurately measure hardness [1], and this is sometimes not possible with thin coatings or surface layers. The AFM is valuable because penetration depth is smaller than in Knoop or microVickers tests and even smaller than that achieved in nanoindenters, but for thin layers, this remains a concern.
- While normally disregarded, crystal orientation or material anisotropy must be taken into account for small-scale indentations. In addition, it is often unclear how deep a particular grain extends into a surface, so that hardness measurements within a single grain can vary considerably [63].
- Over time, sharp indenters will encounter wear and be flattened, as seen in Fig. 25.4. Without correction, hardness seems to depend on indentation depth for Si(111), but when a correction is added to take into account true tip shape, it is seen that hardness is independent of depth [65]. It has been found that the rounding of the indenter tip needs to be taken into account only when the indentation depth is less than 50 nm [66].
- For indentations that span multiple grains, grain boundaries can cause variations in plastic behavior due to the associated restrictions of dislocation movement.
- For materials that exhibit crystalline and noncrystalline structures, commonly observed in many polymers, hardness values will vary depending on which structure is tested.

It is not surprising that a large number of researchers have investigated nano- and microhardness tests and have offered various explanations for the causes of ISE. For example, lubricants have been found to affect the ISE measurements [52, 53, 60, 67–71] and has been modeled by numerous researchers [72–74]. Also, Bhushan and Koinkar [75] developed a modified AFM to record indentations as shallow as 1 nm that allowed hardness measurements of thin films and materials at the nanometer length scale. In their study, Bhushan and Koinkar found that the hardness of Si(111) increased as the penetration depth and load decreased. They attributed this trend to phase transformations beneath

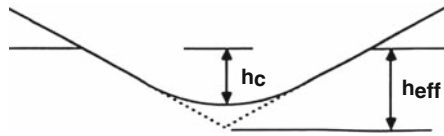


Fig. 25.4. Dulling of probe tip results in an increased tip diameter [64]

the indenter and to films present on the sample surface. Similar results have been obtained by Qian et al. [76] who found that hardness values could vary 10–30% between the micro- and nanoscale.

The presence of an ISE has a fundamental influence on cutting and plowing mechanics, and is therefore a phenomenon of significant importance to manufacturing processes.

Scratch Testing

There are many circumstances in which a soft material comes into contact with the asperities of hard tooling. Metal rolling, forging, extrusion, sheet-metal forming, machining, grinding, ultrasonic machining, abrasive flow machining, and chemical mechanical polishing (CMP) are only a few examples where surface asperities interact with an opposing material in an abrasive manner [77]. The act of plowing is inherently a plastic deformation process and leads to elevated friction, wear, and temperatures. These result in a degradation of surface finish and possible production of wear debris. In some cases, however, plowing is desirable and even necessary (such as in rolling, impression forging, and grinding), whereas in other cases, it is undesirable (as in the die region of forging, extrusion, sheet forming, and drawing operations). Regardless of whether plowing is desirable or not, it is an important phenomenon that has significant effect on product quality, tool life, and hence economics of manufacturing processes.

The use of scratch testing to obtain material flow stress is discussed here, whereas extension to wear is discussed at the end of Sect. 25.2.1. Scratch testing presents additional experimental challenges compared with hardness testing, but is valuable for the following reasons:

- Plowing directly simulates abrasive friction which is a key factor in many manufacturing processes.
- In manufacturing, asperities penetrate into workpieces by a depth that is approximately one-half of the final surface roughness. This depth ranges from around 250 nm for bulk deformation processes, around 50 nm for machining, and as little as 25 nm or so for grinding. AFM scratch tests are in this range and are thus well suited for directly simulating abrasive phenomena in manufacturing processes.
- Contact is restricted to a single asperity and its geometry is well defined due to the precision of AFM probes.

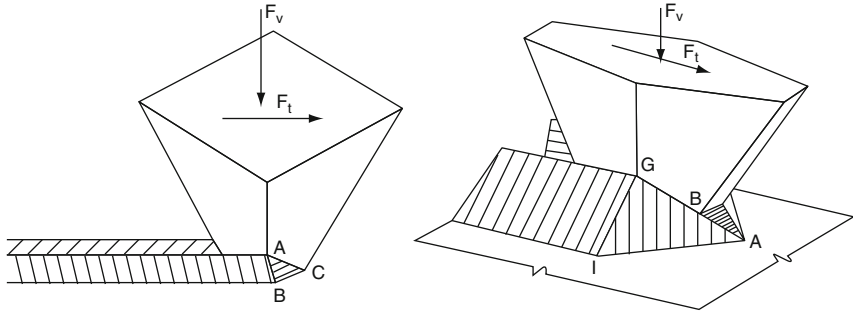


Fig. 25.5. (a) Model of plowing asperity used by DeVathire et al. [78] and Azarkhin and Richmond [79], and (b) model of plowing asperity used by Azarkhin et al. [83]

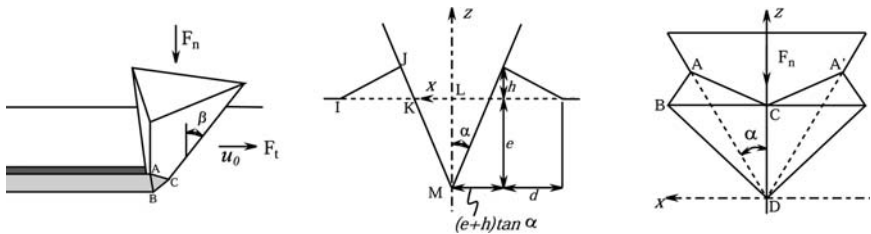


Fig. 25.6. Model of plowing asperity used by Hector and Schmid [85]

- The plastic deformation zone in plowing projects ahead of the indenter, as opposed to hardness testing where the plastic deformation zone projects below the indenter. Thus, thinner surface layers can be investigated with scratch testing compared with hardness testing. This also allows for the mechanical testing of ultrathin coatings that are too thin for hardness tests.
- The effectiveness of lubricants can be evaluated in a controlled manner under a single asperity sliding contact.

Mathematical Models of Plowing Indenters

It should be recognized that simulation of a plowing indenter is a fundamental problem with wide applicability to almost all manufacturing processes. The fundamental modeling of a single plowing asperity was performed by DeVathire et al. [78], Azarkhin and Richmond [79–82] who began with the geometry seen in Fig. 25.5, modified by Azarkhin et al. [83] and Azarkhin and Devenpeck [84], then further modified by Hector and Schmid [85] to fit the geometry of an AFM indenter, seen in Fig. 25.6. Azarkhin and Richmond [79] expressed the work done in plowing as the sum of energy dissipation due to

shear and friction, or

$$W = 2|\bar{u}_1 - \bar{u}_0| \frac{\sigma_0}{\sqrt{3}} (\langle ABD \rangle + \langle BCD \rangle) + 2m \frac{\sigma_0}{3} |\bar{u}_1| \langle ACD \rangle \quad (25.6)$$

where W is the rate of work, m is the friction factor, \bar{u}_0 is the relative speed between the indenter and the surrounding substrate (that is unaffected by plowing), \bar{u}_1 is sliding velocity between the indenter and the material in the tetrahedron ABCD, and the angular brackets denote area. Usually, this equation would be minimized to find the solution, but Azarkhin and Richmond state that this technique does not apply to problems with free surfaces [86] due to a lack of kinetic admissibility. Instead, Azarkhin and Richmond suggested a select surface geometry over which to minimize (25.6). They suggested using a function

$$\Phi = \int_S u_n^2 dS \quad (25.7)$$

where u_n is the normal component of surface velocity and S is the free surface. Two important results obtained from this model are given below:

1. Ridge geometry is related to friction factor.
2. Penetration depth is related to material flow stress and is insensitive to friction for sharp indenters.

Once the geometry and velocity fields have been obtained through minimization of (25.7), the plowing force can be obtained through conservation of energy. The power input is the product of workpiece material velocity around the indenter, u_0 , and the indenter plowing force, F_t , or

$$F_t = \frac{W}{u_0} = 2 \frac{|\mathbf{u}_1 - \mathbf{u}_0|}{u_0} \frac{\sigma_0}{\sqrt{3}} (\langle ABD \rangle + \langle BCD \rangle) + 2m \frac{\sigma_0}{\sqrt{3}} \frac{|\mathbf{u}_1|}{u_0} \langle ACD \rangle \quad (25.8)$$

Also, the normal force can be estimated by combining the vertical sliding component of the frictional stresses on surface ACD and the normal stresses on the projection of ACD, or

$$F_n = 2 \frac{\sigma_0}{\sqrt{3}} \mathbf{k} \cdot (\mathbf{AC} \times \mathbf{CD}) + 2m \frac{\sigma_0}{\sqrt{3}} (\mathbf{k} \cdot \mathbf{u}_1) \langle ACD \rangle \quad (25.9)$$

where \mathbf{k} is the unit vector in the z direction. Equations (25.8) and (25.9) are of particular value since they provide estimates of the workpiece strength and friction factor if the forces are known. This is accomplished by writing

$$m = \frac{\zeta F_t - \phi F_n}{\psi F_n - \eta F_t} \quad (25.10)$$

$$\sigma_0 = \frac{F_t}{\phi + m\psi} \quad (25.11)$$

where, from (25.8) and (25.9), the following expressions have been defined:

$$\zeta = \frac{2}{\sqrt{3}} \mathbf{k} \cdot \langle \text{ACD} \rangle \quad (25.12)$$

$$\phi = \frac{2}{\sqrt{3}} \frac{|\mathbf{u}_1 - \mathbf{u}_0|}{u_0} (\langle \text{ABD} \rangle + \langle \text{BCD} \rangle) \quad (25.13)$$

$$\psi = \frac{2}{\sqrt{3}} \frac{|\mathbf{u}_1|}{u_0} \langle \text{ACD} \rangle \quad (25.14)$$

$$\eta = \frac{2}{\sqrt{3}} (\mathbf{k} \cdot \mathbf{u}_1) \langle \text{ACD} \rangle \quad (25.15)$$

Other models of an asperity contacting a surface have been proposed by Komvopolulos et al. [87] and Polycarpou and Etsion [88]. The reader is directed to their works for an in-depth discussion of their modeling techniques.

Experimental Approaches

There are two principal methods used to perform scratch tests. The first is to use the automatic scratch feature incorporated into AFM control software. This approach typically requires the following procedure:

1. Image the desired area in tapping mode with a silicon tip.
2. Equip the AFM with a diamond-tipped probe.
3. Engage the surface in tapping mode.
4. Set values for scratch parameters (most importantly deflection set point).
5. Set values for automatic scratch parameters (number of scratches, set point increase, etc.).
6. Perform the scratches while capturing force–displacement data.
7. Retract the probe and remove the diamond-tipped cantilever.
8. Re-engage the surface and locate and image the scratches in TappingMode with a silicon tip.

Another method of performing scratches is to use a C-programming language-based macro to directly control tip behavior in nanolithography mode. Using a macro to manage process parameters allows for full customization and control of the scratching procedure. This method of scratch testing usually follows the following steps:

1. Image the desired area in tapping mode with a silicon tip.
2. Equip the AFM with a diamond-tipped probe.
3. Engage the surface in tapping mode.
4. Switch to nanolithography mode.
5. Load and execute the desired macro, simultaneously capturing force and displacement data.

6. Once the macro has finished, retract the probe and replace the diamond-tipped cantilever with a silicon tapping probe.
7. Re-engage surface, locate and image the scratch.

It should be noted that the software for the AFM does not record friction force during scratch tests, nor is it recorded in nanolithography mode. Instead, a signal access module must be used to obtain these measurements. Another complication as discussed by Bhushan [89] is the inevitable error associated with the mounting of the photosensors of the AFM, so the normal and torsional deflections need to be decoupled using the approach of Hector and Schmid [85].

The previously discussed plowing models are applied to AFM experiments in the following manner. The normal and the plowing forces are directly measured during a plowing experiment. The plow track is then measured in tapping mode, and depth of cut, ridge height, and ridge width are recorded. As shown by Azarkhin and Richmond [79], the ratio of ridge height to depth of cut is enough to determine the friction factor, m . Combined with the optimum geometry and kinematic results, (25.7) is a function of σ_o only. Equation (25.11) then gives flow stress from the measured value of F_t . The results of such scratch tests are discussed in Sect. 25.3.2 as applied to bulk and sheet forming.

Wear Testing

In most manufacturing processes, asperity to surface contact does not occur in a single direction for only one cycle, as simulated by scratch testing. For example, when sheet metal is rolled, asperities on the tooling plow the work-piece, then reverse the direction after the neutral point and then plow again along the same plow track but in the opposite direction. Abrasive finishing operations rely on repeated scratching of a surface by many asperities. These processes can be simulated in an AFM by performing various types of wear tests. In wear tests, the probe is rastered either over a line or a prescribed area. Naturally, the orientation of the diamond tip is an important concern for these tests. When a sharp edge of the pyramid is aligned with the plowing direction, as indicated in Fig. 25.7a, motion in the reverse direction causes a blunt face to be the leading edge. For more consistent results between the two directions, an orientation as illustrated in Fig. 25.7b is preferred [90, 91].

Line Tests

Line wear tests are performed by rastering a probe back-and-forth over the same line on the surface of a sample. In the AFM software, this is achieved by disabling the slow scan axis of the cantilever. Repeating this process mimics nanoscale wear by repeated asperity interactions. By varying the load applied

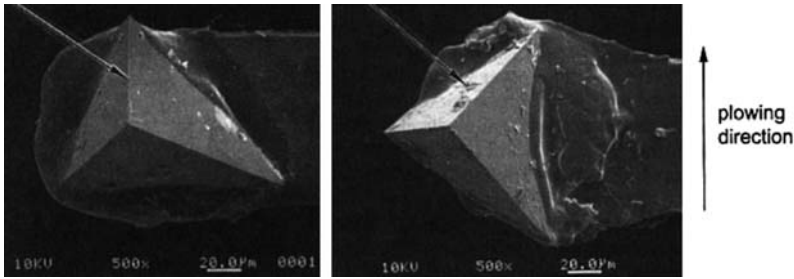


Fig. 25.7. Difference in orientation of diamond on cantilever tip [90]

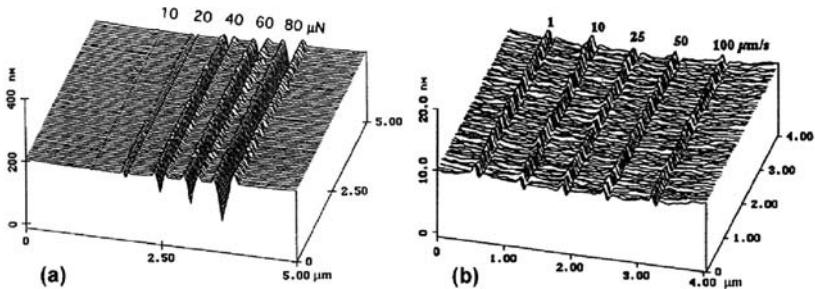


Fig. 25.8. Line wear tracks in (a) Si(111) under increasing load [75] and (b) in Si(100) under increasing velocity [92]

and number of repetitions performed, one can observe the changing wear mechanisms as surface material is removed.

This process was illustrated by Bhushan and Koinkar [75] using Si(111) samples scratched for 10 cycles under varying loads (Fig. 25.8a). As expected, the resultant scratches show that an increasing load results in an increasing wear depth. Scratch speed has also been investigated; Bhushan and Sundararajan [92] varied the scratch velocity on an Si(100) sample and found that the wear depth is independent of scratch speed. The wear tracks at 1, 10, 25, 50, and $100 \mu\text{m s}^{-1}$ for a normal load of $80 \mu\text{N}$ can be seen in Fig. 25.8b. It has been proposed that the increased frictional effects expected at higher velocities is minimal due to the presence of the large amount of substrate material readily available to dissipate the extra heat produced.

These types of tests are common and have been performed in the magnetic storage industry as well as on freshly cleaned, coated, and implanted Si(111) substrates [93–100]. Miyamoto et al. [93] performed multipass scratch testing on a number of different materials including amorphous carbon films, fluorinated amorphous carbon films, silicon-containing amorphous carbon films, and fluorinated silicon-containing amorphous carbon films. They concluded that the fluorination treatment provided the highest strength film and was successful in preventing wear.

Area Tests

The wear behavior and properties of a material are closely related to the characteristics of the surface. While sometimes significant at the macroscale, issues such as the presence or absence of surface contaminants, oxide layers, Beilby layers, and Langmuir-Blodgett films are of greater importance and can significantly affect test results on the nanoscale.

Area wear tests can be performed using an AFM by bringing the probe tip into contact with the surface and rastering the indenter over a given area so that wear patterns and mechanisms can be observed. The result of this test is a depressed region in which the tip has removed successive layers of material with each pass. The surrounding material, presence or absence of chips, and piled ridges at the sides of the wear patch provide valuable information to microscale wear characteristics and mechanisms of the material.

From Kaneko's map of wear modes [101], it is suggested that most AFM wear tests will take place in the "plowing" region since the degree of penetration of a diamond AFM tip is relatively small. Plowed shoulders, however, have undergone cold work and as a result are stronger and less brittle than the original material and can be more easily removed as chips in subsequent passes of the diamond probe.

Although the wear mode map suggests that AFM wear tests will remain in the "plowing" region, it has been observed that under certain circumstances, cutting during area wear tests can occur. Koinkar and Bhushan [102] performed area wear tests on an Si(111) surface and produced the wear patches shown in Fig. 25.9. Further measurements suggested that the volume of the debris was much smaller than the volume of the wear patch itself. From this, it was deduced that the debris created during the area wear test was "loose" and not attached to the surface, and the loose debris is removed or moved during imaging. Figure 25.10 shows SEM images of the wear patch and particles and provides insight regarding the morphology of the wear debris.

Under light loads, small wear particles were found outside the wear patch, but as the load increased, so did the number of ribbon-like chips. These chips

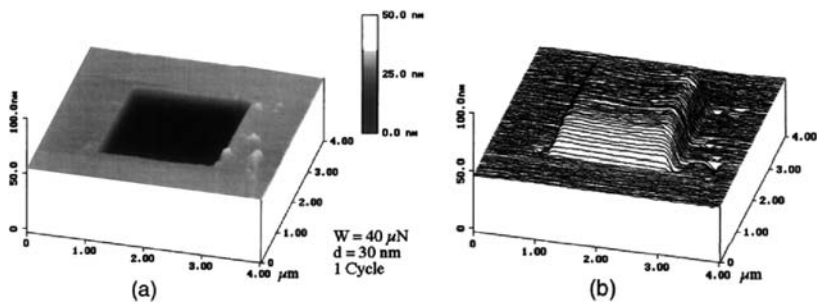


Fig. 25.9. Typical (a) height and (b) inverted height data of a wear patch on Si(111) as a result of an area wear test [102]

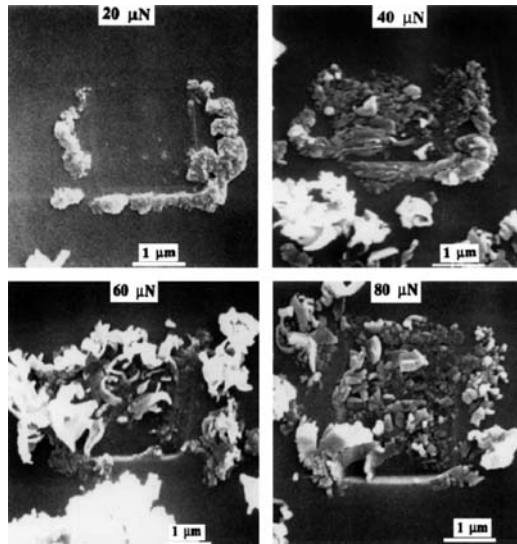


Fig. 25.10. SEM images of wear patches created with increasing load and their associated debris [102]

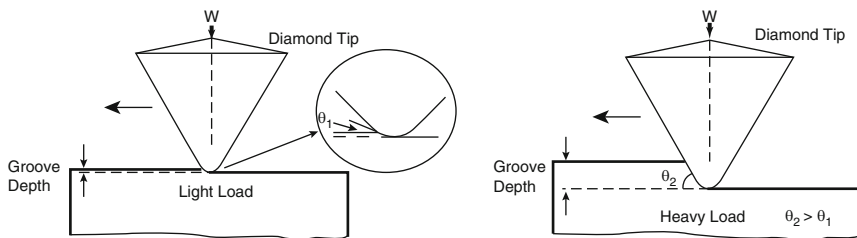


Fig. 25.11. Interaction between AFM tip and sample under (a) low loads and (b) high loads [102]

are the direct result of cutting. In order to explain the transition from the plowing regime to cutting, Chiou and Kato [103] investigated the area of contact between the surface and AFM tip. Hertz analysis of the contact area was used to determine that the yield stress of the surface is exceeded and plastic deformation readily occurs during area wear tests [104]. With help from the work of Hokkirigawa and Kato [105], the degree of penetration was calculated for the various loads used, and it was determined that it was the change in degree of penetration and attack angle between the tip and the surface that caused the transition from plowing to cutting. Figure 25.11 shows the changing interaction between tip and sample under light and heavy loads.

Area wear tests have also been performed by numerous researchers on a variety of materials. Schiffmann [106] studied the wear mechanisms of metal-containing amorphous hydrocarbon thin films and developed models based on

load and time dependence, Tago et al. [107] and King [108] used amorphous carbon films as well; Pidduck and Smith [109] performed tests on antiwear films to be used in engine valve trains; Kasai et al. [110] studied self-assembled perfluorosilane monolayers; Meyer et al. [111] investigated the wear properties of L-B films; Maw et al. [112] used SiO₂, as did Katsuki et al. [113], who performed wear tests of Si and SiO₂ in a KOH solution to simulate CMP, a further discussion of which is found in Sect. 25.4.2; Chung and Kim [114] used Si and Si₂N₄ tips to wear the surfaces of Au, Cu, diamond-like carbon, and bare Si; Miyamoto et al. [94] used ion-implanted silicon; Thundat et al. [115] observed the layer by layer removal of atoms from the cleavage faces of lead pyrophosphate; and Kaneko et al. [116] tested the wear behavior of polycarbonate. Hamada and Kaneko [117] used a point-contact microscope to perform indent and wear tests on polycarbonate and epoxy surfaces; and Kaneko et al. [118] demonstrated the ability of the AFM to investigate microwear on silicon and C⁺ implanted silicon.

25.2.2 Friction/Lubricant Evaluation

Review of Lubrication Fundamentals

Wilson [119] identified four regimes of lubrication that can be distinguished by the unitless fluid film parameter:

$$\Lambda = \frac{h}{\sqrt{R_a^2 + R_b^2}} \quad (25.16)$$

where h is the film thickness, and R_a and R_b are the surface roughnesses of the surfaces in contact. For situations where Λ is greater than 10, the surfaces are said to be in thick film lubrication where there is no surface asperity contact. For Λ between 3 and 10, a thin film of lubricant separates the mating surfaces and prevents asperity contact. Both thick and thin film lubrication, also known in combination as hydrodynamic lubrication, display very low coefficients of friction and low wear rates. For this reason, most hydrodynamic and lubricated rolling element bearings are designed to operate in the thin or thick film lubrication regimes. When Λ is between 1 and 3, the surfaces are said to be in a mixed lubrication regime, in which the surfaces are separated by pockets of pressurized lubricant trapped between direct surface asperity contact. These direct contacts lead to an increased friction coefficient and higher wear rates than full-film lubrication. The transition between full-film and mixed lubrication occurs gradually as the surfaces become closer and asperities begin to contact the opposing surface under higher loads. At film parameters unit less than 1, the surfaces are in intimate contact with the only separation due to the adsorbed films on their surfaces. High friction and high wear rates are inevitable as the surfaces are forced together and asperity contact becomes severe. These regimes and their corresponding coefficients of friction are displayed in Fig. 25.12, commonly referred to as a Stribeck curve.

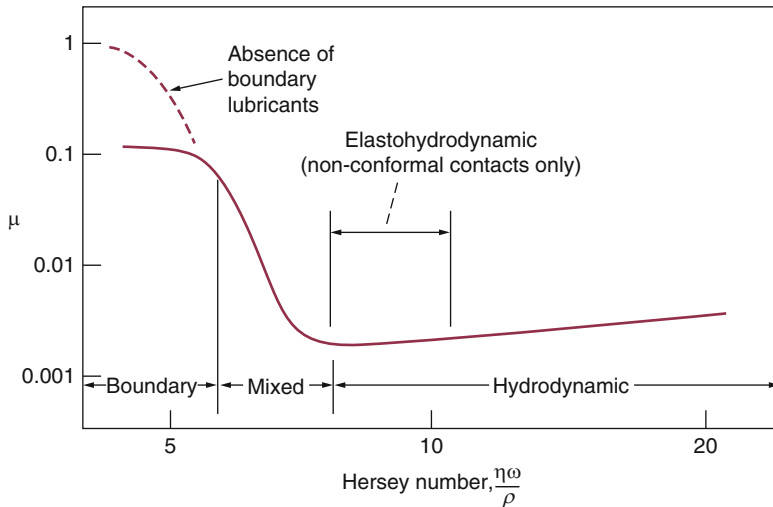


Fig. 25.12. The Stribeck curve showing coefficient of friction as a function of film parameter [120]

Boundary and mixed lubrication rely on boundary and extreme pressure lubricants to control friction and wear. As described by Ku [121] and summarized by Hamrock et al. [122], surface films develop through a number of different mechanisms:

1. Physisorption: Molecule chains are attracted and adhere to the material surface by weak van der Waals forces. The bonding energy is usually less than 40 kJ mol^{-1} [123] and the chains are easily removed under various conditions. This mechanism is illustrated in Fig. 25.13.
2. Chemisorption: Molecule chains form a chemical bond with the surface, providing bonding energies much larger than 40 kJ mol^{-1} [123]. Bonding can also occur between the chains chemisorbed onto the surface which is illustrated in Fig. 25.14, providing a closely packed boundary lubricant film. Because of their intimate bondage with the surface, these films are limited to a thickness of only one molecule.
3. Chemical reaction not involving the substrate: As a result of chemical reactions that occur within the lubricant itself, separate from the material surface, coatings of reaction products are plated onto the material surface.
4. Chemical reaction involving the substrate: As a result of a chemical reaction between the lubricant and the material surface, the chemical composition of the surface is changed and a thin film is created, depicted in Fig. 25.15.

Although Yoshizawa and Israelachvili [124], O'Shea et al. [125], and similar researchers have evaluated physisorbed and chemisorbed boundary lubricants by using friction force microscopy (FFM), chemical reactions involving and not involving the substrate require direct measurement of material properties

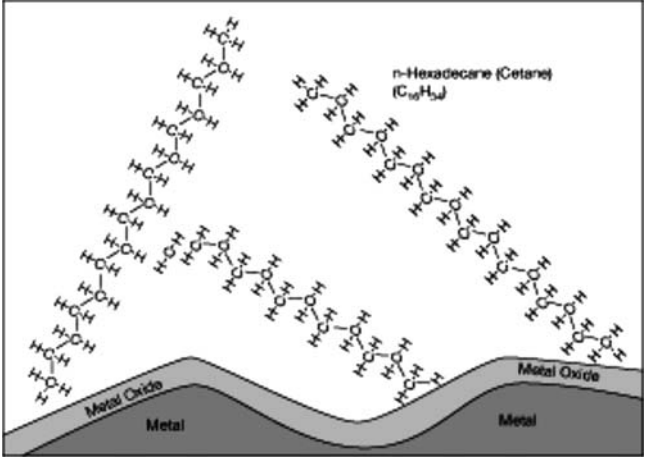


Fig. 25.13. Physisorbed lubricant on a metal surface [121]

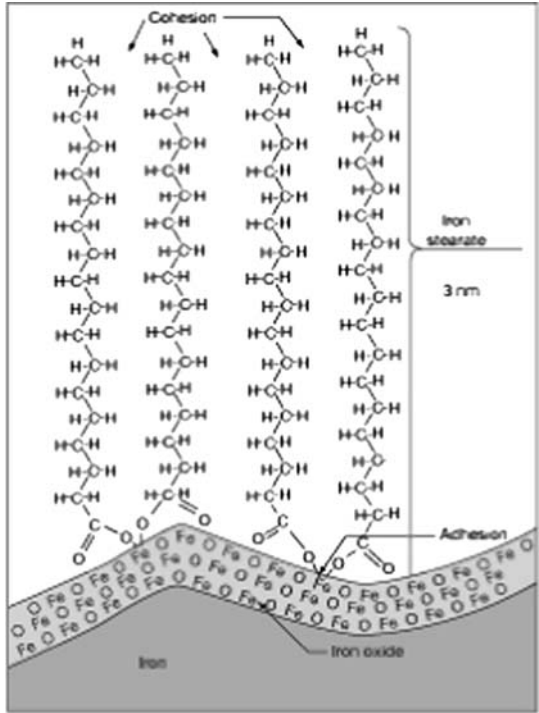


Fig. 25.14. Stearic acid chemisorbed on the surface of iron [121]

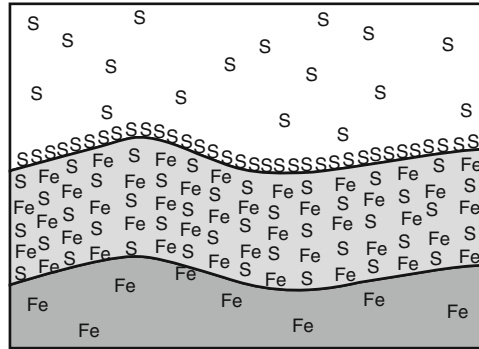


Fig. 25.15. Production of iron sulfate via chemical reaction involving the substrate [122]

on surface layers since the reactions involve more than weak bonding with the material surface. In order to test lubricants that induce reactions involving or not involving the substrate, scratch and wear testing as described above are more suitable.

Friction Force Microscopy

Friction Force Microscopy has been available since its development by Mate et al. in 1987 [126] and allows measurement of friction on the nanoscale. The FFM technique is conceptually simple: an AFM cantilever tip is lowered toward a surface until adhesion forces cause a jump-to-contact. The tip is then rastered across the surface, perpendicular to the cantilever axis, in a back-and-forth motion. Adhesion between the tip and the sample surface exerts a torque on the cantilever and causes it to twist (see Fig. 25.16). This twist registers as a voltage difference on the left–right (L and R) photodetectors seen in Fig. 25.16, and allows calculation of local friction values if the torsional cantilever stiffness is known.

Numerous researchers have used FFM to investigate the nanoscale friction properties of various materials, the most popular of which were initially mica [14, 127–130] and freshly cleaved highly ordered pyrolytic graphite [35, 126, 131, 132]. These materials had the benefit of achieving very smooth surfaces, allowing friction to be decoupled from surface roughness effects.

With the development of nanolubrication theories, FFM became the ideal method for investigating boundary lubrication at individual surface asperities. When the AFM tip is brought toward a specimen, hydrodynamic lubrication conditions occur. As the tip approaches the surface, the gap is reduced, that is the probe acts as a squeeze-film bearing [122], and eventually a boundary lubrication condition is reached. With a single asperity on an AFM probe, no partial lubrication or gradual transition between regimes is generally encountered; the normal load between asperity and surface is either transferred by

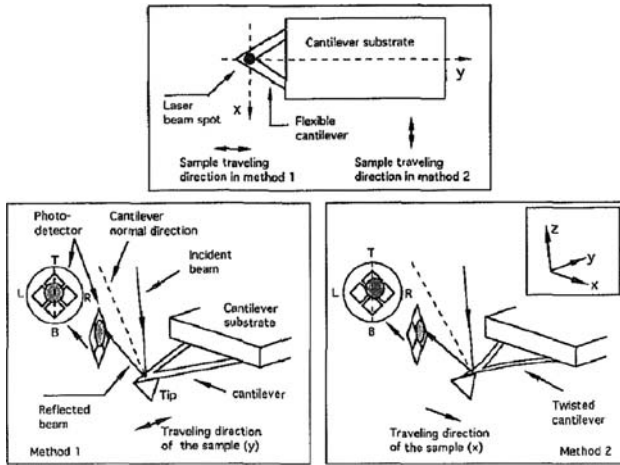


Fig. 25.16. Cantilever twisting and corresponding effect on photodetector signal [35]

the boundary layer or it is not. This offers opportunities and challenges for simulation of manufacturing processes since these commonly are performed in the mixed lubrication regime.

While the AFM probes directly simulate contacting asperities in boundary or mixed lubrication regimes, the behavior and rheology of the surrounding fluid cannot be directly investigated in an AFM. FFM has been used by many researchers to investigate and evaluate boundary films and their properties. As examples, Meyer et al. [111] and Bhushan et al. [133] focused on Langmuir-Blodgett films; Overney et al. [134] observed local anisotropic behavior in organic bilayers; Ahn et al. [135] investigated self-assembled monolayers; and O’Shea et al. [125] used FFM to evaluate the strength of multilayered boundary films.

25.3 Applications to Metal Forming

25.3.1 Evaluation of Lubricants

Lubrication plays a key role in most manufacturing processes; its serves to lower friction, reduce wear, and aid in heat removal. As discussed below, orange peel considerations require that the relevant lubrication regimes in manufacturing operations be confined to boundary or mixed lubrication. As discussed by Opalka et al. [90], most lubricants in manufacturing are formulated with extreme pressure and boundary additive blends, so that they form boundary films through which the load is transferred [136]. Lubricant blends are formulated for particular material and tooling combinations, and also will

include additives for nontribological considerations (such as biocides, brighteners, antioxidants). Boundary films can be formed by any of the mechanisms mentioned in Sect. 25.2.2.

Since the boundary lubrication and mixed regimes dominate, asperities on the tooling surface make physical contact with and often plow into the workpiece material. This frequent and harsh wear under heavy loads causes debris from the workpiece to be formed, which can subsequently become entrained in the lubricant. During sheet rolling, for example, these wear particles are pressed onto the surface of the workpiece or the tooling, causing either “smudge” on the resulting product [137] or a degradation of tooling surface. Not only do these asperity contacts cause smudge and tooling damage, they also increase the necessary process forces and decrease the overall efficiency.

25.3.2 Bulk and Sheet Forming

Bulk deformation processes involve such operations as forging, rolling, extrusion, drawing, and swaging and can be done hot or cold. Bulk deformation operations involve large strains, high strain rates, surface generation, and the exposure of the substrate to the surface [1]. A general introduction to these operations is contained in Kalpakjian and Schmid [1] and Schey [138]. Forging is the controlled plastic deformation of a material through compressive stresses and is done hot or cold, or in open dies or closed dies. Open die forging involves simple tooling and large plastic deformation, whereas closed die forging uses more complicated die shapes and confined plastic flow. Rolling produces constant cross sections by deforming a workpiece between two or more cylindrical or shaped rolls, and can be considered to be a continuous form of forging. In extrusion, a billet is forced through a die opening to produce a workpiece of constant cross section and finite length, and is often done hot. Drawing pulls material through a converging die to produce rods, tube, and wire, and is usually continuous and done cold. During swaging, the diameter of a rod or tube is reduced by two or four rotating and reciprocating dies.

Unconstrained surfaces can encounter roughening due to slip bands moving to the surface within a grain or by surface grain motion [139]. Grain-induced surface roughening is called “orange peel” and has been commonly observed and is objectionable for aesthetic reasons. Tong et al. [139] used the AFM and a specially fabricated in-situ tension test apparatus to study the occurrence of orange peel on Al-5%Mg and its evolution during increasing strain. They successfully identified slip bands within a grain, which surfaced during their experiments; these slip bands are clearly visible in the AFM traces seen in Fig. 25.17. They also documented orange peel associated with grain translation and rotation at the surface. Orange peel, or grain-scale roughening, has also been studied by many others [140–158].

As an example operation, flat rolling is depicted in Fig. 25.18. In rolling, as in other bulk deformation processes, the real area of contact in the absence of a lubricant is close to the apparent area of contact. This means that there is

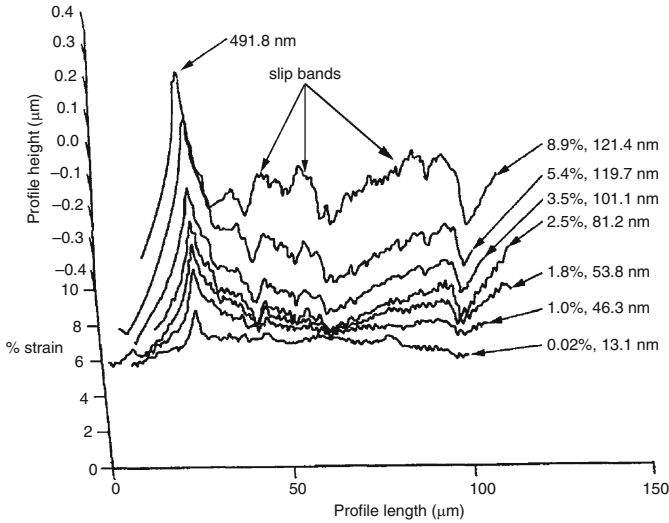


Fig. 25.17. In situ tensile test within an AFM of aluminum showing surface deformation under increasing tensile strain. Numbers shown are the percent strain and roughness. The appearance of slip bands has been identified [139]

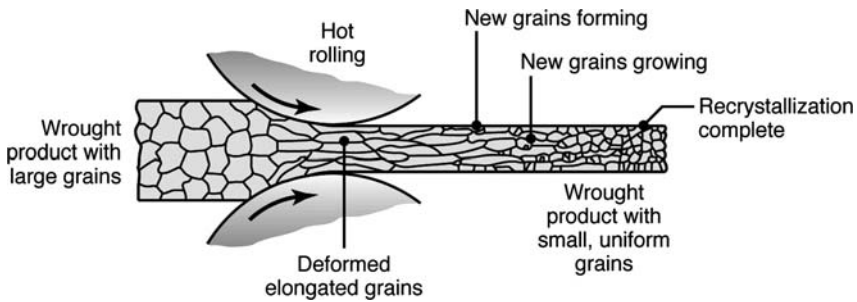


Fig. 25.18. Illustration of the flat rolling process [1]

intimate contact between the mating surfaces of the rollers and the workpiece. Principal sources of friction are adhesive and abrasive interactions between the roll and the workpiece. Work rolls are usually iron or steel, with titanium carbide used only occasionally in cluster mills. Abrasion is assisted by the fact that the workpiece surface is soft due to a substrate strain rate effect [159–161] and therefore, the roll asperities readily penetrate and plow through the workpiece. This has the effect of also impressing the roll surface roughness onto the workpiece, and allows manufacture of smooth surfaces if the roll roughness is properly maintained. This is done in practice by periodically grinding the roll surface; the ground topology presents sharp ridges that plow workpieces, so that asperity plowing discussed above is very pertinent to friction and wear

in rolling. Plouraboué and Boehm [162] found that surface features as small as 50 nm are transferred from roller to workpiece during cold rolling of aluminum.

Normally, tribologists are concerned with minimizing friction, but in rolling, friction is required to draw material through the work rolls. Excessive friction leads to high torque and rolling forces, however, so that a well-controlled level of friction is required.

One other complication in metal rolling should be briefly introduced. To entrain workpieces in the absence of front and back tension, the work roll must move faster than the workpiece. As the workpiece thickness is decreased, conservation of mass requires that its velocity increase. Generally, the workpiece velocity exceeds the roll velocity at the outlet. Somewhere between first contact and the outlet, there is a point at which the rolls and the workpiece are moving with the same velocity. This point is referred to as the neutral point or neutral plane, and is a location where friction is undefined.

Despite its importance, single-asperity simulations of plowing in bulk deformation was difficult to perform at the proper length scales until contact mode AFM and superior diamond-tipped stainless steel probes became available. Schmid et al. [163] investigated asperity plowing of three aluminum substrates that are relevant to industrial rolling operations: 99.99% pure aluminum, a 5182-O alloy, and a 7150-T6 aerospace alloy. Using two diamond-tipped cantilevers, one “blunt” and one “sharp,” they performed scratch tests and compared their results with those experienced during the actual rolling process. The “blunt” indenter has an irregular tip just as occurs on tooling, with a relevant tip radius of around 100 nm. The “sharp” indenter was used to correspond to the mathematical models of plowing discussed earlier. Their results show that the blunt indenter resulted in shallow plow tracks and uneven ridges surrounding the scratch. Dislodged material had a tendency of breaking off and forming wear particles (smudge). When a sharp indenter, with a tip radius of approximately 5 nm was used, quantitative material properties could be obtained and material behavior in plowing differentiated. The pure aluminum sample was plowed inside a single grain parallel to the surface slip lines and always had clean, continuous ridges, as seen in Fig. 25.19a. The 5182-O alloy resulted in ridge waviness in both the transverse and longitudinal directions, as shown in Fig. 25.19b. In machining, such waviness is generally attributed to unstable friction and flow of the material. The waviness of the ridge, along with the presence of hardened substrate material on the leading edges of the indenter, revealed the presence of a “built-up edge” on the leading indenter face. This phenomenon is due to adhesion between the surfaces and material transfer from substrate to tooling. The material comprising the built up edge is extremely hard and lowers the effective sharpness of the tooling. Duller tooling leads to higher forces and explains the pronounced ISE for this material. This result confirms practical experience in forming this material; 5182-O is generally associated with high friction, smudge, and galling.

The 7150-T6 alloy exhibited mixed behavior; with shallow plow tracks, the associated ridges were clean, consistent, and remained attached to the

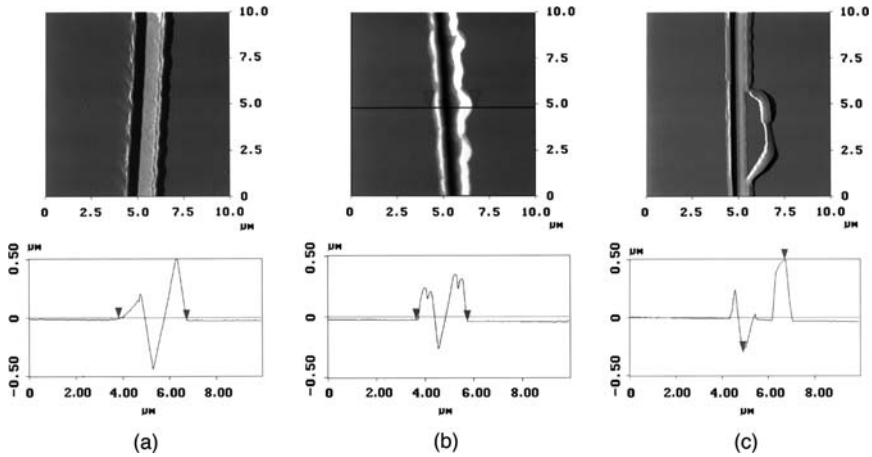


Fig. 25.19. Typical plow tracks observed during the plowing of (a) 99.99% Al, (b) a 5182 Al alloy, and (c) a 7150 aerospace alloy

substrate. However, when speed and load were increased, deeper cuts were made, and the ridge behavior changed dramatically. Undulating ridges were observed, some of which broke loose from the sample and caused difficulty during subsequent imaging, as seen in Fig. 25.19c. This lack of ductility is to be expected from a hard alloy such as 7150-T6.

Using the method of Hector and Schmid [85], the flow stresses for each material at various depths of cut were calculated and are shown in Fig. 25.20. The pronounced size effect mimics that found in nanoindentation tests [52] and is curve fit to the form of (25.5).

These results demonstrated the utility of contact mode studies in an AFM to investigate macroscale tribology effects in metalworking. The rapid development of the built-up edge when plowing the 5182-O alloy is attributable to strong adhesion between workpiece and asperity. The chip formation found while using the 7150-T6 alloy is attributed to its low fracture toughness. These results are consistent with industrial experience.

Opalka et al. [90] performed similar tests on 3004 aluminum alloys to investigate the mechanics of boundary lubricants used in metalworking. Their target material, 3004 aluminum, is of significant industrial importance since it comprises the body of most metal beverage containers sold in the United States of America. A number of 3004 alloy specimens were prepared and boundary films of selected additives (butyle stearate, stearic acid, and lauryl alcohol) were cast from solvent solutions onto the surface. Plowing experiments then took place in a similar fashion to previous work on unlubricated surfaces. A clear distinction could be made, as seen in Fig. 25.21b, c, where the ridges for the 3004 specimen lubricated with butyl stearate are partially detached but still connected to the material, whereas the ridges for a stearic

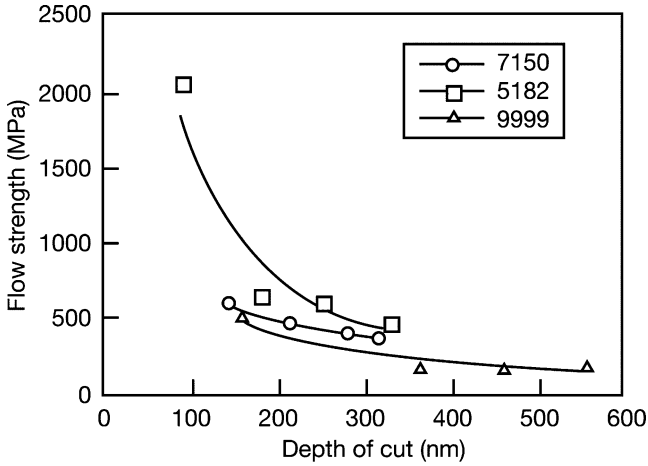


Fig. 25.20. Predicted flow strength as a function of penetration depth during plowing experiments [163]

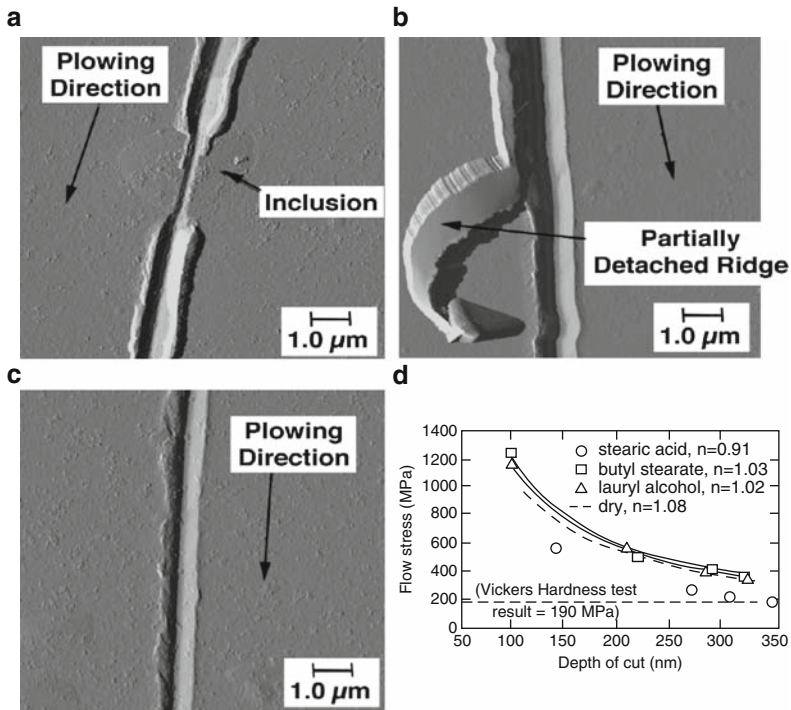


Fig. 25.21. Results from plowing with boundary lubricants. (a) Effect of inclusion in plowing 3004 alloy, (b) plowing in the presence of butyl stearate, (c) plowing in the presence of stearic acid, and (d) resultant surface flow stress predictions [77]

acid lubricated specimen are totally removed. These results demonstrated the effectiveness of the chemisorbed fatty acid boundary lubricants in lowering the surface flow strength of the material. Figure 25.21a demonstrates the effect that an inclusion has a plowing asperity. Note how drastically the plow track changes: much shallower penetration depth and lower ridges are clearly seen. These hard inclusions are thought to remove adhered aluminum from the tooling, thereby preserving the tooling surface finish and resulting in smoother containers.

A more provocative finding [90] was the ability to discern the mechanism through which a boundary lubricant acts. Physisorbed and chemisorbed lubricants are commonly encountered, but chemical reactions involving the substrate are also present in some manufacturing applications. Situations where surface layers are affected by chemical reactions are sometimes called chemomechanical effects, to which the interested reader is directed to the excellent work of Westwood and Lockwood [164]. Since the plowing experiments conducted on an AFM are not strongly affected by adhesion, a boundary lubricant that depends on a boundary film mechanism will not be readily discernable from dry plowing. On the contrary, if the surface layer mechanical properties are affected, plowing experiments will readily detect this mechanism. Thus, physisorption and chemisorption can be separated from chemical reactions involving the substrate, or other reactions that lead to a surface layer.

This difference is clearly visible when comparing the flow stresses obtained from the mathematical model for the three-boundary additives compared with dry plowing. As seen in Fig. 25.21d, the flow stress is totally unaffected by the presence of butyl stearate or lauryl alcohol, but is clearly reduced by the presence of stearic acid.

These experiments have demonstrated that these single asperity plowing experiments are useful for screening lubricant additives in terms of preventing smudge and also for determining the lubrication mechanism for a given additive. As such, contact mode AFM is a very valuable diagnostic tool and also a method of performing fundamental research on lubrication theory.

The previously mentioned plowing tests were all performed without regard for the crystallographic orientation with which plowing was performed. Hector et al. [63] investigated the dependence of surface mechanical properties on the grain orientation of the substrate material. The materials used were annealed aluminum alloys with columnar equiaxed grains 1.0–5.0 mm in diameter (they were visible with the naked eye). This assures that plowing experiments do not occur on multiple grains or grain boundaries, and is one reason for the well-behaved systems investigated to date. Another subtle point is that surface slip lines are visible on the surface of the puck (apparently from using pliers to remove the pucks from the adhesive in the fixture used in polishing). All plowing experiments conducted in the research described above were performed so that the plowing direction was parallel to visible surface slip lines.

It was desired to determine the effect of grain orientation on the plowing process.

The crystallographic orientations of selected surface grains on the pucks were first determined with a scanning electron microscope-based technique known as orientation imaging microscopy (OIM). During OIM, the sample volume is inclined to the incident electron beam by 70° and subsequently stepped across a crystalline surface. Any sample volume with an undisturbed lattice will give rise to a characteristic diffraction pattern that is a two-dimensional projection of a three-dimensional unit cell. From this diffraction pattern, the three angles relating the crystal coordinate system with the sample system (i.e., the crystallographic orientation) are calculated using automated pattern recognition procedures. The probe/sample interaction volume for aluminum determines the lateral resolution, which is typically $0.5\ \mu\text{m}$. In addition to providing quantitative information about crystallographic texture, OIM can also provide insight into grain size, misorientation across grain boundaries, and local strain.

Plowing experiments were then performed along the $[110]$, $[1\bar{1}0]$, and $[111]$ directions as well as in the (111) plane perpendicular to the $[111]$ direction. Two sets of plow tracks were generated in the four crystallographic directions, each set containing four plow tracks generated at progressively increasing normal forces.

Friction/plowing force results are shown in Fig. 25.22 and show little variation between values for each material except in the (111) and perpendicular to the (111) direction. In these directions, the Al-1.0wt%Mn produces the highest friction forces. This can be attributed to the fact that this alloy has the highest weight percent of additive used in these experiments. The authors

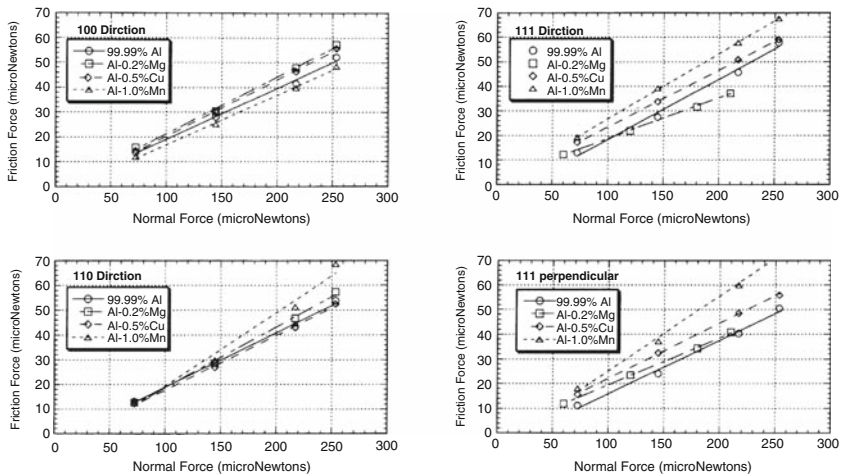


Fig. 25.22. Friction versus normal force for plowing in various directions [63]

also note that the chemical attraction between alloying atoms with the indenter face plays a role in friction measurements as well as in the density of atoms along the plowing direction. The high-purity alloy has the lowest flow stresses under all normal forces and directions, as expected. Low amounts of alloying elements result in substitutionally embedded atoms that act as obstacles to dislocation movement along glide planes, resulting in measurable increase in flow stress. It is determined that crystallographic orientation plays an important role in the measurement of surface mechanical properties and should be included in such experiments.

25.3.3 Powder Processing

Powder metallurgy (PM) is a common method of manufacturing a wide variety of parts, but is most commonly used for tribological components such as bearings, gears, sprockets, and cams because of its ability to self-lubricate under boundary lubrication conditions [165]. PM operations generally include the following steps:

1. A green part is formed, most commonly through compression, as seen in Fig. 25.23 or metal injection molding where the powder is mixed with thermoplastic and injected into a mold. In addition, simple shapes can be isostatically pressed, spray formed, or rolled.
2. The green part is then sintered, where it is raised to an elevated temperature approaching its melting point. The bond between powders is dramatically increased during sintering, resulting in a strong part.
3. Finishing operations such as machining, drilling, grinding, infiltration, etc., are performed as needed.

The AFM is a useful tool for investigating a number of fundamental mechanisms of PM processing. Of course, abrasion of metal particles against tooling

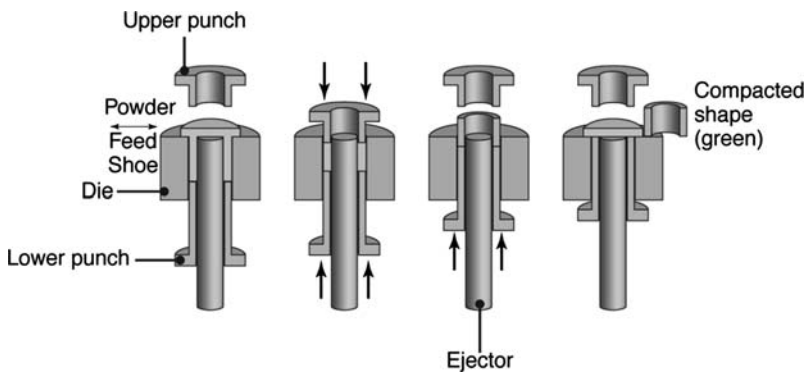


Fig. 25.23. Illustration of the powder compaction process in the formation of PM parts [1]

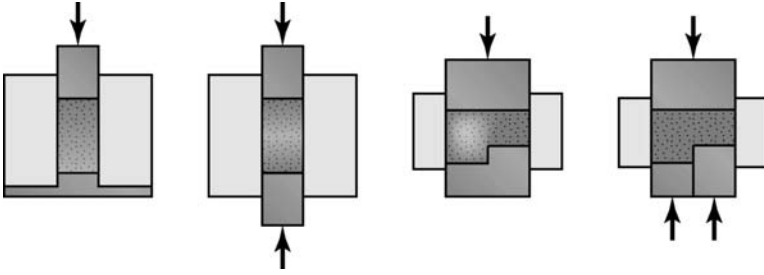


Fig. 25.24. Powder compaction and associated change in part density with various punch configurations [1]

in powder compaction or metal injection molding is analogous to scratch and wear testing discussed earlier and will not be elaborated here. The other main contribution of AFM research to understand PM involves understanding surface forces and the interaction of particles during compaction.

To prevent excessive warpage or fracture during sintering and cooling, it is desired to obtain a uniform density after compaction. For this reason, multipunch systems are often used as shown in Fig. 25.24. Similarly, models of compaction of metal powders have been developed to provide design guidance for PM parts.

Compaction modeling theories are either continuum based or founded on interparticle forces and contact models. Original efforts relied heavily on Hertz theory [166] and later contact models developed by Johnson et al. [167], Derjaguin et al. [168], Muller et al. [169], Hughes and White [170], and Maugis and Pollock [171]. These contact models were the basis of the original compaction simulations that have developed into numerous other categories: plasticity theory models [172, 173], microscopic models [174–183], continuum models [184, 185], numerical modeling [186–190], and discrete/finite element models [191–204].

When attempting to model the compaction of powders whose length scales are much smaller than those contacts originally investigated by Hertz, interparticular adhesion must be taken into account. The AFM has been used by many to determine the interparticulate forces. It is thought that these values can be integrated into existing models for more accurate compaction modeling.

The interaction between single particles was first performed by Ducker et al. [205, 206] and Butt [207], and is referred to as colloidal AFM. Current experiments either alter the tip of a pre-existing cantilever [208] or affix a sphere onto a tipless AFM cantilever. The sphere is then brought into close contact with a surface and the attraction–repulsion forces are recorded from the cantilever displacements on the AFM.

Results from these types of experiments are summarized by Kappl and Butt [209] and have provided insight into the adhesive forces acting at such

small length scales and the effect of applied load and loading time [210–222], humidity [210, 223, 224], surface roughness [211, 225–231], and particle coatings [232–237] on these measurements. Surface force data is helpful for static analysis of powders, but during compaction, powder particles slide, roll, and indent each other. To better predict the flow characteristics of powders during compaction, the AFM has been used to measure the tangential force evolution of small particles [214, 238–242].

25.4 Abrasive Machining Processes

Often it is necessary to impart better surface finish or achieve tighter tolerances than are typically possible through forming or machining operations. For these applications, abrasive machining is often a viable and economical finishing process. Examples of abrasive machining operations include grinding, wire brushing, honing, lapping, polishing, buffing, and deburring. Material removal occurs through chip formation due to the plowing of the material surface by hard abrasive grains. Since the size of the grains used can vary and their hardness is very high, these operations are capable of achieving a smooth surface finish and removing even the hardest surfaces. These processes are routinely used on heat-treated surfaces, weld beads, concrete, and even ceramics [1]. Abrasion can be combined with other mechanisms in electrochemical grinding and CMP, where abrasion plays a major role in material removal but is not the only active mechanism.

25.4.1 Grinding and Polishing

Grinding is a material removal process that involves abrasive wear of the workpiece by individual abrasive grains that have been bonded to tooling of various shapes and sizes, illustrated in Fig. 25.25. Abrasive grain–workpiece interactions, shown in Fig. 25.26, typically involve large negative rake angles (and associated high specific energies), high sliding speeds, and small depths of cut. Grinding usually involves mounted abrasives, commonly in vitrified (i.e., glassy) wheels. The mechanics and technology of grinding is reviewed by Malkin [243] and Shaw [244].

In the AFM, scratch and wear tests are directly applicable to grinding operations, although the indenter geometry is a triangular-based pyramid in the AFM and usually highly irregular in a grinding wheel. Regardless, the same mechanics of abrasive asperity interaction with a soft surface can be simulated, and the theory and experimental approaches described earlier translate directly to grinding and polishing operations. Although there have been many attempts to model the grinding and abrasive machining processes [245–254], it was not until recently that these processes have been directly simulated in an AFM.

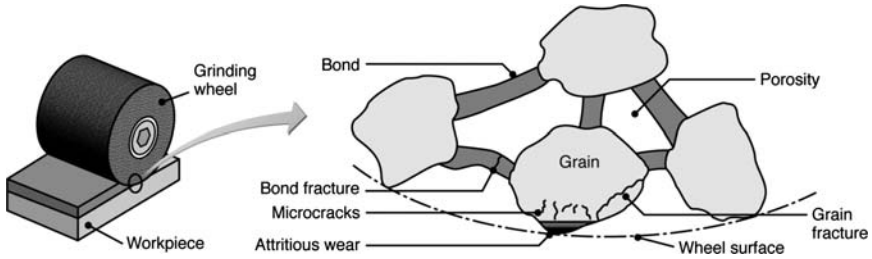


Fig. 25.25. Illustration of the grinding process showing the grinding wheel and the intergrain structure [1]

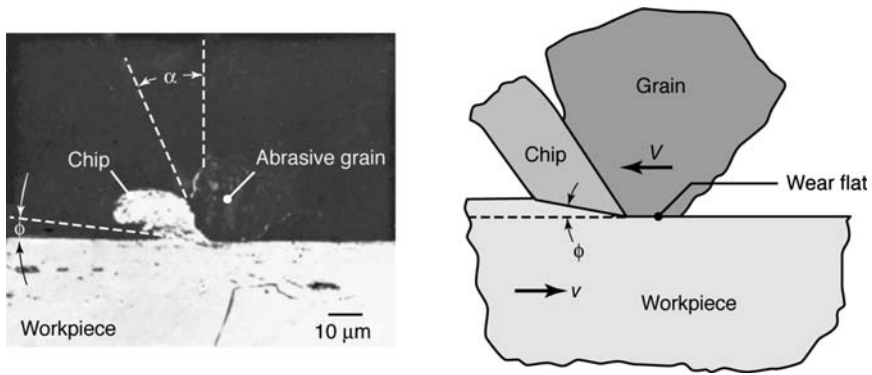


Fig. 25.26. Magnification of the interaction between workpiece and abrasive grain [1]

25.4.2 Chemical Mechanical Polishing

It is often necessary to produce an extremely smooth and flat surface, and in microelectronics and MEMS manufacture as well as hard drive production, CMP is a common approach. In the microelectronics industry, this process is often called chemical mechanical planarization because the process also removes variations in workpiece thickness across a wafer [255]. In CMP, a workpiece such as a silicon wafer is brought into contact with a polishing pad on which a slurry of erosive chemicals and abrasive particles is deposited, as illustrated in Fig. 25.27. The combination of chemical effects and abrasive mechanical wear results in a smooth surface finish. A review of CMP is contained in Liang and Craven [256].

The selection of the chemicals and particles to include in the slurry solution plays an important role in the final surface finish of the workpieces. The role of the abrasive particles is obvious to remove surface layers and to create a smooth finish. The role of the chemicals is to aide the abrasive wear process by softening the surface of the material, often by controlling the thickness or completely removing the oxide layer. Figure 25.28 shows that thicker oxide

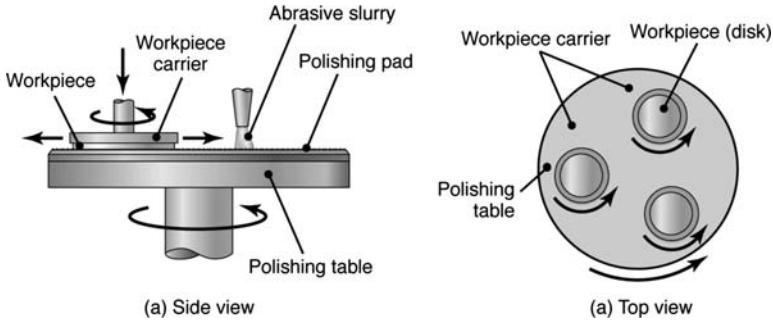


Fig. 25.27. (a) Side and (b) top views of the chemical mechanical polishing process [1]

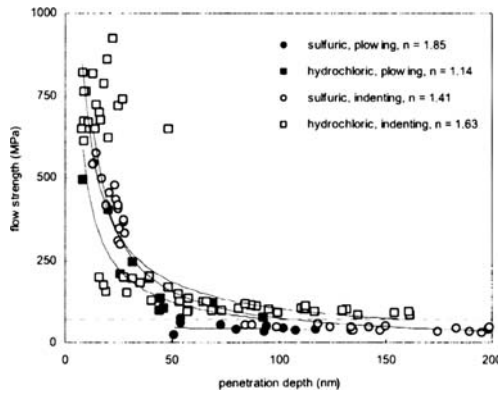


Fig. 25.28. Estimated flow strength as a function of penetration depth [257]

layers result in harder acting surfaces that lead to less effective CMP. Accordingly, chemicals that soften the surface more than others are more effective CMP solutions. The balance between the abrasive action of the particles and the chemical corrosion is an important process parameter. Enough surface material must be removed to expose new layers to the chemical solution, but removal of too much material compromises the final surface finish of the product. Attempts to model this process have been made [258–277], as well as simulate the process directly in an AFM [113,257,278,279]. Effective AFM modeling of the CMP process is complicated by the necessity of performing all experiments inside of a fluid cell.

Katsuki et al. [113] used an AFM to investigate the wear characteristics of silicon and silicon dioxide submerged in a chemically reactive KOH solution. Using a tapping mode silicon cantilever, they rastered the surface of each sample under KOH solutions of varying pH. Examination of the surfaces and probes revealed that the tip wore preferentially to the sample, seen in Fig. 25.29. Although these results accurately describe wear characteristics

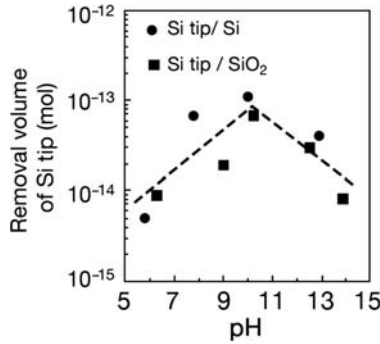


Fig. 25.29. Dependence of Si AFM tip material removal volume as a function of solution pH [113]

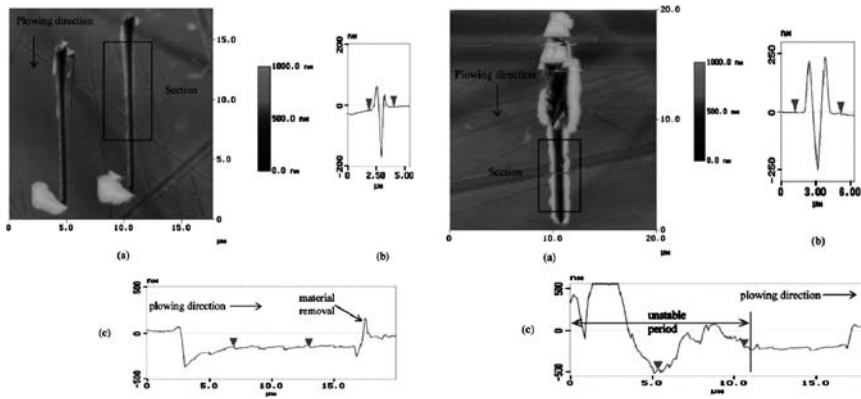


Fig. 25.30. (a) Transverse and (b) axial plowing of aluminum [279]

using an AFM, a more appropriate model for abrasive wear during manufacturing is one that involves a hard asperity plowing a flat, soft substrate material.

Liu et al. [279] simulated CMP using an AFM with aluminum and copper substrates. Indentation tests performed on a 99.99% aluminum sample produced results as expected: a linear indentation depth to normal force ratio and an ISE index of 1.49.

Scratch tests were also performed on the dry aluminum and copper surfaces. Transverse and axial plowing of the aluminum resulted in the scratches seen in Fig. 25.30a, b, respectively. It can be seen that performing transverse plowing allows for more reliable and cleaner results. The calculated flow stress from the indentation and plowing experiments for copper can be seen in Fig. 25.31 and provide very good agreement.

Post et al. [257] used an AFM to investigate the softening effects of various chemical solutions on the surface of a copper sample. Using both

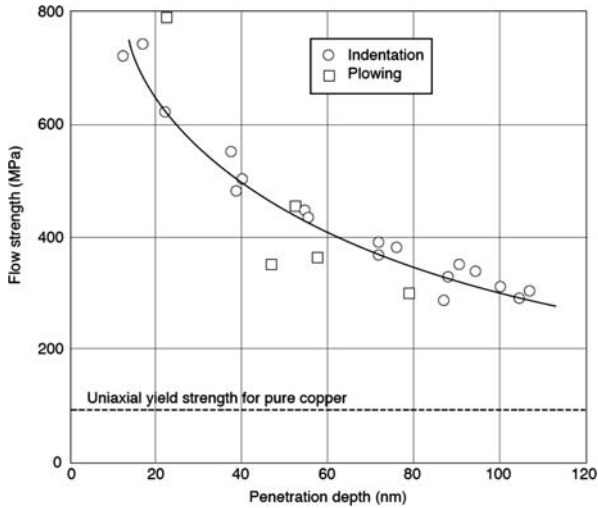


Fig. 25.31. Predicted flow stress of copper from indentation and plowing experiments [279]

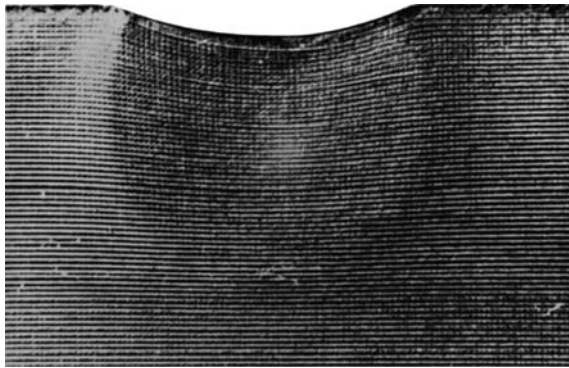


Fig. 25.32. Extent of plastic deformation into the substrate during indentation [280]

indentation and plowing techniques, hardness and flow stress values were calculated for differently treated materials. It was concluded that hardness-type tests are not well suited for the investigation of thin oxide films since most of the deformation extends well below the oxide film and into the substrate itself. As discussed earlier, plowing is preferred for such thin films because the deformation is concentrated near the surface and is projected in front of the indenting tip, as demonstrated experimentally in Fig. 25.32. It was found that sulfuric/BTA (1H-benzotriazole)-treated surfaces had thinner films than did hydrochloric/BTA-treated surfaces that led to a softer-acting surface. Typical penetration depths observed in CMP are about 50 nm which is larger than

the oxide film observed on the sulfuric/BTA-treated specimen. This means that the abrasive particles were able to penetrate the thin film and reach the substrate material more easily than with a thick oxide film caused by hydrochloric/BTA treatment. These experimental results are supported by industrial observations of CMP.

The CMP process has also been studied by Berdyyeva et al. [278] who performed wear tests on copper submerged in solutions of varying pH. Analysis of the changing wear depths allowed for the comparison of the solutions. It was found that a solution of pH 3 (the lowest tested) was the most effective and allowed for the most rapid material removal. A similar process was performed by Devecchio et al. [266] on 99.99% aluminum with similar results. They found that an area scanned in a corrosive solution for an hour exhibited faster material removal than either by the scanning or being in solution alone. These efforts show that in situ modeling of CMP using an AFM is able to differentiate between solutions and provide valuable information about the interaction between slurry and substrate.

25.4.3 Miscellaneous Applications

Nanolithography

Several approaches have been developed to perform nanoscale manufacturing with essential elements performed on AFMs. The probes used in AFM vary greatly in size, material, and capability. The diamond-tipped stainless steel cantilever has a tip radius around 10 nm; by contacting and plowing across a surface, it can produce grooves up to a few microns thick; the spacing between lines depends on the groove depth needed. However, the grooves produced in an AFM in contact mode can remove material and produce a mask for a reticle. When combined with extreme infrared lithography, submerged lithography, or x-ray lithography, this becomes a viable approach for the manufacture of structures with true nanoscale resolution [1, 281].

Nanoscale devices and nanoelectromechanical systems (NEMS) are an area of significant current research, but few (if any) commercially successful products have been produced. Most probes and sensors are microelectromechanical systems (MEMS), as are inkjet printer mechanisms and most medical devices. Increased miniaturization has occurred in packaging, but the commercial success of NEMS is still in the future.

A large number of researchers have investigated AFM-based lithography on an AFM, using a variety of materials, coatings, and substrates. A good overview of nanolithography and its potential is given by Tseng et al. [282], to which the interested reader is directed. The approaches used are generally to perform scratching to remove exposed resist [283–289], to expose the resist using heat provided by probes [290, 291], to expose resist through probe-delivered radiation [292–297], or to cause oxidation [298–304], all conducted in

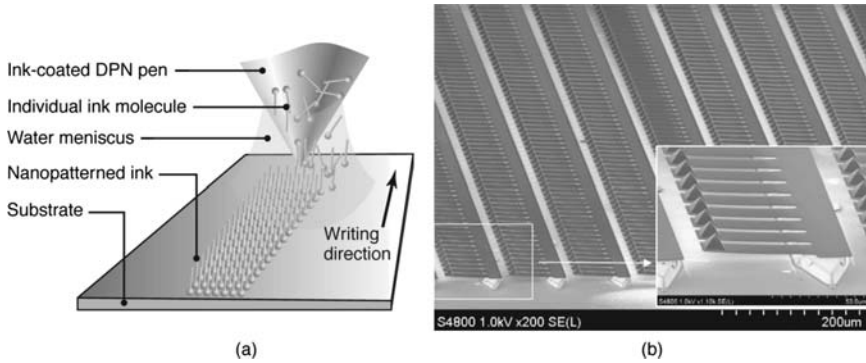


Fig. 25.33. (a) Illustration of the DPN process and (b) an SEM of a parallel DPN pen array [1]

nanolithography mode on the AFM. The techniques described in Sect. 25.2.1 are all relevant and applicable.

An innovative approach that should be briefly mentioned is dip pen nanolithography (DPN), illustrated in Fig. 25.33a. DPN is used on an AFM to transfer chemicals onto substrates. The process can produce lines as narrow as 10 nm and can be used with many parallel pens (Fig. 25.33b), typically produced of silicon nitride and containing as many as 55,000 pens in a 1 cm² area. In a topdown approach, DPN is used to produce a mask suitable for lithography. Because of the large number of cantilevers or pens, relatively large areas of a reticle can be processed at one time, making this a valuable lithography approach.

25.5 Polymer Processing

During the formation of a commercial beverage can, a deep-drawn aluminum blank is ironed multiple times to achieve thin side walls. The ironing process, illustrated in Fig. 25.34, usually involves a lubricant to reduce the friction between the die and cup, thus reducing the punch force needed. For aluminum cans, the use of a lubricant mandates an extensive cleaning process so that the food or beverage does not become contaminated with oils or lubricants. In order to prevent interaction between the contents of the container and the metal itself, usually a coat of polymer coating is sprayed on the interior of the can. It has been proposed that if a polymer-coated steel were used as a workpiece, the use of lubricants and subsequent cleaning operations would be unnecessary. The use of a polymer-coated steel in ironing has been evaluated by Jaworski et al. [305], who demonstrated that such an arrangement was feasible.

Huang et al. [306] conducted experiments on polymer-laminated steel sheet with a variety of tool geometries and temperatures. Their mathematical

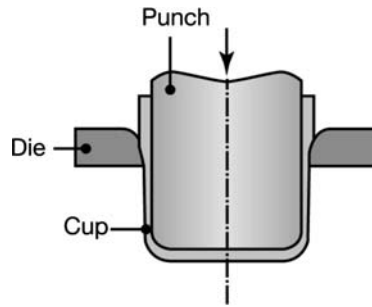


Fig. 25.34. Illustration of the ironing process [1]

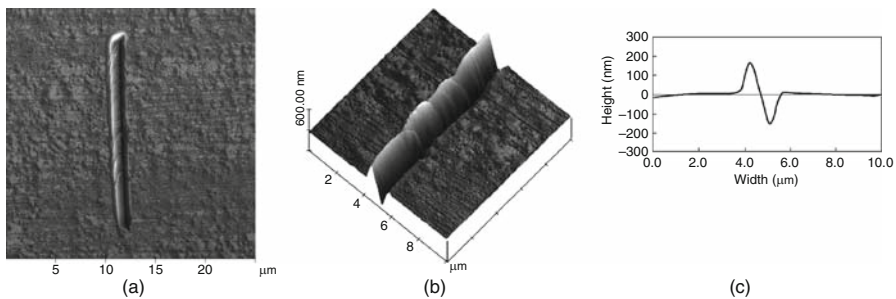


Fig. 25.35. Typical plowing experiment. (a) A 20- μm scratch obtained on PET coating, (b) 10 μm section detail, and (c) averaged ridge profile for the scratch in (b) [306]

models required the flow stress of the polymer solid lubricant. They used contact mode AFM to evaluate the flow stress of a polyethylene terephthalate (PET) lamination. Single asperity plowing experiments were performed in scratch mode with a sharp diamond indenter. The mathematical model discussed earlier suggested a friction factor of 0.72 and a shear strength of 20.28 MPa for the plow track in Fig. 25.35a.

The same diamond indenter was used for nanoindenting on the PET coating for comparison purposes. Note that shear strength of the polymer film can be approximated from hardness by the following approximation [52]:

$$H_s = 1.1\sigma_0 = 1.1\sqrt{3}\tau_0 = 1.91\tau_0 \quad (25.17)$$

where H_s is the hardness, σ_0 is the yield strength, and τ_0 is the shear strength of the polymer. Figure 25.36b shows the comparison of the results between nanoplowing and nanoindentation experiments. Indentation showed a slightly stronger scale effect than in plowing experiments. However, the shear strength of the PET coating measured by the two methods was close, suggesting that the either asperity plowing or nanoindentation measurement of shear strength gives reasonable results. Huang et al. [306] found that these

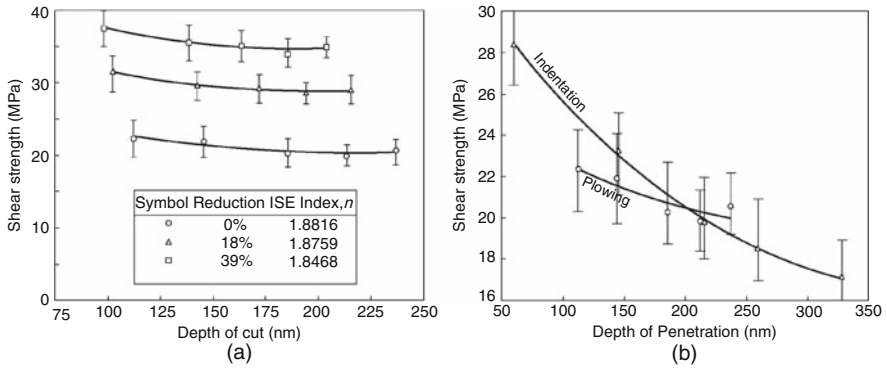


Fig. 25.36. Predicted shear strength as a function of penetration depth in plowing experiments. (a) Plowing experiments for ironed workpieces, and (b) comparison of plowing and indentation experiments [306]

measurements were valuable when combined with the pressure-dependent behavior of shear strength of polymers. The shear strength of the polymer coating is proportional to the contact pressure [307], or

$$\tau = \tau_0 + \omega p \tag{25.18}$$

where τ_0 is the shear strength at ambient pressure, τ is the shear strength at pressure p , and ω is a piezo-strengthening coefficient. The shear strength measured in nanoplowing tests actually yields the value of τ_0 in (25.18). The pressure p was measured in their ironing experiments, and the piezo-strengthening coefficient was taken from the literature [307]. This resulted in a flow stress measurement that corresponded with ironing behavior.

25.6 Conclusions

1. The AFM has become a standard research tool for tribologists, and has allowed fundamental research into lubrication, friction, and wear mechanisms to take place. This is also the case for manufacturing research where the AFM has demonstrated its value in a number of ways.
2. Imaging with an AFM is widespread and an essential quality assurance technology for the microelectronics and bearings industries, among others.
3. Hardness testing performed in an AFM is useful for determining mechanical properties and inferring performance of tooling and molds. The AFM allows determination of hardness at length scales that are pertinent to manufacturing processes such as rolling, forging, and extrusion.
4. Scratch testing and models of plowing have been demonstrated to directly simulate abrasive friction and wear, and AFM-based experiments in friction force and lithography modes have proven valuable. These approaches have

allowed investigation of lubrication mechanisms, quantification of surface layers, and have helped explain observations by production personnel that were difficult to directly explain otherwise.

5. The ISE is an important phenomenon for manufacturing operations, as macroscale material properties are not relevant to such processes as abrasive machining or asperity-scale interactions common in bulk forming. The AFM allows experiments to be conducted at proper length scales, and therefore inherently incorporate the proper size effect.
6. Wear testing involving rastering of a contact probe back and forth over a line or area allows direct simulation of processes such as grinding, polishing, lapping, abrasive flow machining, and ultrasonic machining. When used in conjunction with a fluid cell, CMP can be simulated at the proper lengths scales in an AFM.
7. The AFM has been directly applied in lithography operations using a number of strategies. DPN is an innovative tool for reproducing many replicas of a given pattern on a resist.

References

1. S. Kalpakjian, S.R. Schmid, *Manufacturing Processes for Engineering Materials* (Pearson, Up Saddle River, 2007)
2. S.R. Schmid, W.R.D. Wilson, in *Modern Tribology Handbook* (CRC, New York, 2000)
3. D. Tabor, *The Hardness of Metals* (Oxford University Press, Oxford, 1951)
4. W. Johnson, R. Sowerby, R. Venter, *Plane Strain Slip Line Fields for Metal Deformation Process Source Book & Bibliography* (Pergamon, Oxford, 1982)
5. V.E. Lysaght, *Indentation Hardness Testing* (Reinhold, New York, 1949)
6. E.S. Berkovich, *Ind. Diamond Rev.* **11**, 129–132 (1951)
7. B.W. Mott, *Micro-indentation Hardness Testing* (Buttersworth, London, 1956)
8. H. O'Neill, *Hardness Measurement of Metals and Alloys* (Chapman & Hall, London, 1967)
9. D. Tabor, *Rev. Phys. Technol.* **1**, 145–179 (1970)
10. J.H. Westbrook, H. Conrad, *The Science of Hardness Testing and Its Research Applications* (American Society for Metals, Metals Park, Ohio 1973)
11. K.L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1987)
12. B. Bhushan, B.K. Gupta, *Handbook of Tribology: Materials, coatings, and surface treatments* (McGraw-Hill, New York, 1991)
13. F Mohs, *Grundriss der Mineralogie*. (In der Arnoldischen Buchhandlung, Dresden 1822)
14. O. Marti, J. Colchero, J. Mlynek, *Nanotechnology* **1**, 141–144 (1990)
15. J.P. Cleveland, S. Manne, D. Bocek, P.K. Hansma, *Rev. Sci. Instrum.* **64**, 403–405 (1993)
16. T.J. Senden, W.A. Ducker, *Langmuir* **10**, 1003–1004 (1994)
17. J.E. Sader, *Rev. Sci. Instrum.* **66**, 4583 (1995)
18. J.E. Sader, *J. Appl. Phys.* **84**, 64–76 (1998)

19. J.E. Sader, J.W.M. Chon, P. Mulvaney, *Rev. Sci. Instrum.* **70**, 3967 (1999)
20. J.M. Neumeister, W.A. Ducker, *Rev. Sci. Instrum.* **65**, 2527–2531 (1994)
21. J.L. Hazel, V.V. Tsukruk, *J. Tribol.* **120**, 814–819 (1998)
22. J.L. Hazel, V.V. Tsukruk, *Thin Solid Film.* **339**, 249–257 (1999)
23. R.W. Stark, T. Drobek, W.M. Heckl, *Ultramicroscopy* **86**, 207–215 (2001)
24. H. Butt, P. Siedle, K. Seifert, K. Fendler, K. Seeger, E. Bamberg, A.L. Weisenhorn, K. Goldie, A. Engle, *J. Microsc.* **169**, 75–84 (1993)
25. E.L. Florin, V.T. Moy, H.E. Gaub, *Science* **264**, 415–417 (1994)
26. A. Torii, M. Sasaki, K. Hane, S. Okuma, *Measurement Sci. Technol.* **7**, 179–184 (1996)
27. C.T. Gibson, G.S. Watson, S. Myhra, *Nanotechnology* **7**, 259–262 (1996)
28. M. Tortonese, M. Kirk, *Proc. SPIE.* **3009**, 53–60 (1997)
29. J.D. Holbery, V.L. Eden, M. Sarikaya, R.M. Fisher, *Rev. Sci. Instrum.* **71**, 3769–3776 (2000)
30. P.J. Cumpson, P. Zhdan, J. Hedley, *Ultramicroscopy* **100**, 241–251 (2004)
31. S.B. Aksu, J.A. Turner, *Rev. Sci. Instrum.* **78**, 043704 (2007)
32. C.T. Gibson, B.L. Weeks, J.R.I. Lee, C. Abell, T. Rayment, *Rev. Sci. Instrum.* **72**, 2340–2343 (2001)
33. G.Y. Jing, J. Ma, D.P. Yu, *J. Electron Microsc.* **56**, 21–25 (2007)
34. J.L. Hutter, J. Bechhoefer, *Rev. Sci. Instrum.* **64**, 1868–1873 (1993)
35. J. Ruan, B. Bhushan, Atomic-scale friction measurements using friction force microscopy part I: General principles and new measurement techniques *Journal of Tribology* **116**, 378 (1994)
36. N.A. Burnham, X. Chen, C.S. Hodges, G.A. Matei, E.J. Thorenson, C.J. Roberts, M.C. Davies, S.J.B. Tendler, *Nanotechnology* **14**, 1–6 (2003)
37. C.T. Gibson, G.S. Watson, S. Myhra, *Scanning* **19**, 564–581 (1997)
38. C.T. Gibson, D.A. Smith, C.J. Roberts, *Nanotechnology* **16**, 234–238 (2005)
39. C. Huang, S.R. Schmid, *Wear* **252**, 704–710 (2002)
40. Veeco Dimension 3100 Manual (2000)
41. E.M. Onitsch, *Mikroskopie* **2**, 131 (1947)
42. D. Maugis, Desalos- G. Andarelli, A. Heurtel, R. Courtel, *Tribol. Trans.* **21**, 1–19 (1978)
43. M.D. Pashley, J.B. Pethica, D. Tabor, *Wear* **100**, 7 (1984)
44. G.M. Pharr, W.C. Oliver, *J. Mater. Res.* **4**, 94 (1989)
45. N. Gane, F.P. Bowden, *J. Appl. Phys.* **39**, 1432 (2003)
46. E.R. Marshall, M.C. Shaw, *Trans. ASME.* **74**, (1952)
47. W.R. Backer, E.R. Marshall, M.C. Shaw, *Trans. ASME.* **74**, (1952)
48. M. Merchant, *J. Appl. Phys.* **16**, 267–275 (1945)
49. K. Karmarsch, *Wien Polytechn. Inst. Jahrb.* **12**, 61 (1834)
50. A.A. Griffith, *Philos. Trans. R Soc. Lond.* **221**, 163 (1921)
51. E. Meyer, *Zeitschrift Vereines Deutscher Ing.* **52**, 82–85 (1908)
52. C.J. McHargue, in *Surface Mechanical Properties Using Nanoindentation*, ed. by B. Bhushan. *Micro/Nanotribology and its Applications*. pp. 467–492 (1997)
53. D.R. Tate, *Trans. ASME, Kluwer Academic Publishers: Boston*, **35**, 374 (1945)
54. B.R. Lawn, V.R. Howes, *J. Mater. Sci.* **16**, 2745–2752 (1981)
55. B.J. Briscoe, L. Fiori, E. Pelillo, *J. Phys. D Appl. Phys.* **31**, 2395–2405 (1998)
56. N.A. Stilwell, D. Tabor, *Proc. Phys. Soc.* **78**, 169–179 (1961)
57. M.R. VanLandingham, S.H. McKnight, G.R. Palmese, J.R. Elings, X. Huang, T.A. Bogetti, R.F. Eduljee, J.W., J. Gillespie, *J. Adhes.* **64**, 31–59 (1997)

58. M. Atkinson, H. Shi, *Mater. Sci. Technol.* **5**, 613 (1989)
59. H. Shi, M. Atkinson, *J. Mater. Sci.* **25**, 2111 (1990)
60. H. Li, A. Ghosh, Y.H. Han, R.C. Bradt, *J. Mater. Res.* **8**, 1028–1032 (1993)
61. M. Mata, J. Alcalá, *J. Mech. Phys. Solids* **52**, 145–165 (2004)
62. L.E. Farmer, P.L.B. Oxley, *J. Mech. Phys. Solids* **19**, 369 (1971)
63. Hector, L.G. Jr., S.M. Opalka, H. Weiland, S.R. Schmid, *Mater. Res. Soc. Symp. Proc.* **522**, 399–408 (1998)
64. B. Bhushan, *Nanomechanical Properties of Solid Surfaces and Thin Films*, ed. by B. Bhushan, *Handbook of Micro/Nano Tribology* (CRC, Boca Raton, 1995) pp. 321–396
65. M.F. Doerner, W.D. Nix, *J. Mater. Res.* **1**, 601–609 (1986)
66. J.G. Swadener, E.P. George, G.M. Pharr, *J. Mech. Phys. Solids* **50**, 681–694 (2002)
67. N. Gane, J.M. Cox, *Philos. Mag.* **22**, 881 (1970)
68. C. Hays, E.G. Kendall, *Metallography* **6**, 275 (1973)
69. K. Hirao, M. Tomozawa, *J. Am. Ceram. Soc.* **70**, 497–502 (1987)
70. S.J. Bull, T.F. Page, E.H. Yoffe, *Philos. Mag. Lett.* **59**, 281 (1989)
71. M. Atkinson, *J. Mater. Res.* **10**, 2908 (1995)
72. H. Li, R.C. Bradt, *J. Mater. Sci.* **28**, 917 (1993)
73. J.Y. Shu, N.A. Fleck, *Int. J. Solids Struct.* **35**, 1363–1383 (1998)
74. Y. Huang, F. Zhang, K.C. Hwang, W.D. Nix, G.M. Pharr, G. Feng, *J. Mech. Phys. Solids* **54**, 1668–1686 (2006)
75. B. Bhushan, V.N. Koinkar, *Appl. Phys. Lett.* **64**, 1653–1655 (1994)
76. L. Qian, M. Li, Z. Zhou, H. Yang, X. Shi, *Surf. Coating Technol.* **195**, 264–271 (2005)
77. S.R. Schmid, Hector, L.G. Jr., in *Macroscale Insight from Nanoscale Testing. Nanotribology: Critical Assessment of Research Needs* (Kluwer, Boston, 2002)
78. M. Devathire, F. Delamare, E. Felder, *Wear* **66**, 51–64 (1984)
79. A. Azarkhin, O. Richmond, *Wear* **157**, 408–419 (1992)
80. A. Azarkhin, O. Richmond, *J. Appl. Mech.* **56**, 10–14 (1989)
81. A. Azarkhin, O. Richmond, *J. Tribol.* **112**, 324–329 (1990)
82. A. Azarkhin, O. Richmond, *J. Appl. Mech.* **58**, 493–499 (1991)
83. A. Azarkhin, O. Richmond, M. Devenpeck, *Wear* **192**, 157–164 (1996)
84. A. Azarkhin, M.L. Devenpeck, *Wear* **206**, 147–155 (1997)
85. Hector, L.G. Jr., S.R. Schmid, *Wear* **215**, 247–256 (1998)
86. D.C. Drucker, W. Prager, H.J. Greenberg, *Quart. Appl. Math.* **9**, 381 (1952)
87. K. Komvopoulos, N. Saka, N.P. Suh, *J. Tribol.* **107**, 452 (1985)
88. A.A. Polycarpou, I. Etsion, *J. Tribol.* **120**, 296–304 (1998)
89. B. Bhushan, in *Micro/nanotribology and micro/nanomechanics of magnetic storage devices and ME004DS*, ed. by B. Bhushan *Handbook of Micro/nano Tribology*, (CRC, Boca Raton, 1995) pp. 443–504
90. S.M. Opalka, G., Louis J. Hector, S.R. Schmid, R.A. Reich, J.M. Epp, *J. Tribol.* **121**, 384–393 (1999)
91. V.N. Koinkar, B. Bhushan, *J. Vacuum. Sci. Technol. A* **14**, 2378–2391 (1996)
92. B. Bhushan, S. Sundararajan, *Acta Mater.* **46**, 3793 (1998)
93. T. Miyamoto, R. Kaneko, S. Miyake, *J. Vacuum Sci. Technol. B* **9**, 1336–1339 (1991)
94. T. Miyamoto, S. Miyake, R. Kaneko, *Wear* **162–164**, 733–738 (1993)
95. B. Bhushan, V.N. Koinkar, J. Ruan, *J. Eng. Tribol.* **208**, 17–29 (1994)

96. B. Bhushan, V.N. Koinkar, *J. Appl. Phys.* **75**, 5741–5746 (1994)
97. B. Bhushan, V.N. Koinkar, *Tribol. Trans.* **38**, 119–127 (1995)
98. B. Bhushan, V.N. Koinkar, *Wear* **180**, 9–16 (1995)
99. B. Bhushan, V.N. Koinkar, *Wear* **183**, 360–370 (1995)
100. B. Bhushan, J.N. Israelachvili, U. Landman, *Nature* **374**, 607–616 (1995)
101. R. Kaneko, T. Miyamoto, E. Hamada, *Microwear*. 183–221 (1995)
102. V.N. Koinkar, B. Bhushan, *J. Mater. Res.* **12**, 3219 (1997)
103. Y.C. Chiou, K. Kato, *J. JSLE Int. Ed.* **9**, 11–16 (1988)
104. B. Bhushan, *Appl. Mech. Rev.* **49**, 275–298 (1996)
105. K. Hokkirigawa, K. Kato, *Tribol. Int.* **21**, 51–57 (1988)
106. K. Schiffmann, *Wear* **216**, 27–34 (1998)
107. A. Tago, T. Masuda, Y. Taketa, Plated magnetic disk. *Rev. Electr. Commun. Lab.* **25**, 1315–1324 (1977)
108. F.K. King, *IEEE Transact. Magnet.* **17**, 1376–1378 (1981).
109. A.J. Pidduck, G.C. Smith, *Wear* **212**, 254–264 (1997)
110. T. Kasai, B. Bhushan, G. Kulik, L. Barbieri, P. Hoffmann, *J. Vacuum Sci. Technol. B Microelectron. Nanometer Struc.* **23**, 995–1003 (2005)
111. E. Meyer, R. Overney, D. Brodbeck, L. Howald, L.R. Uthi, J. Frommer, G.H.J. Untherodt, *Phys. Rev. Lett.* **69**, 1777–1780 (1992)
112. W. Maw, F. Stevens, S.C. Langford, J.T. Dickinson, *J. Appl. Phys.* **92**, 5103–5109 (2002)
113. F. Katsuki, A. Saguchi, W. Takahashi, J. Watanabe, *Jpn. J. Appl. Phys.* **41**, 4919–4923 (2002)
114. K. Chung, D. Kim, *Tribol. Lett.* **15**, 135–145 (2003)
115. T. Thundat, B.C. Sales, B.C. Chakoumakos, L.A. Boatner, D.P. Allison, R.J. Warmack, *Surf. Sci.* **293**, L863–L869 (1993)
116. R. Kaneko, E. Hamada, Y. Andon, *Jpn. J. Tribol.* **38**, 63–73 (1993)
117. E. Hamada, R. Kaneko, *J. Phys. D Appl. Phys.* **25**, A53–A56 (1992)
118. R. Kaneko, T. Miyamoto, Y. Andoh, E. Hamada, *Thin Solid Film.* **273**, 105–111 (1996)
119. W.R.D. Wilson, in *Friction and lubrication in sheet metal forming*, ed. by D.P. Koistinen and N.M. Wang, *Mechanics of sheet metal forming* (Plenum Press: New York, 1978) pp. 157–177
120. B.J. Hamrock, D. Dowson, *Ball Bearing Lubrication - The Elastohydrodynamics of Elliptical Contacts* (Wiley, New York, 1981)
121. P.M. Ku, *Interdisciplinary approach to the lubrication of concentrated contacts*. NASA (1970)
122. B.J. Hamrock, S.R. Schmid, B.O. Jacobson, *Fundamentals of Fluid Film Lubrication* (Marcel Dekker, New York, 2004)
123. D.H. Buckley, *Surface Effects in Adhesion, Friction, Wear, and Lubrication* (Elsevier: New York, 1981)
124. H. Yoshizawa, J.N. Israelachvili, *J. Phys. Chem.* **97**, 11300–11313 (1993)
125. S.J. O'Shea, M.E. Welland, T. Rayment, *Appl. Phys. Lett.* **61**, 2240–2242 (1992)
126. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, *Phys. Rev. Lett.* **59**, 1942–1945 (1987)
127. A. Bailey, *J. Appl. Phys.* **32**, 1407 (2004)
128. R. Erlandsson, G. Hadziioannou, C.M. Mate, M. McClelland, S. Chiang, *J. Chem. Phys.* **89**, 5190 (1988)

129. M. Hirano, K. Shinjo, R. Kaneko, Y. Murata, *Phys. Rev. Lett.* **67**, 2642 (1991)
130. Y. Hoshi, T. Kawagishi, H. Kawakatsu, *Jpn. J. Appl. Phys.* **39**, 3804 (2000)
131. C.M. Mate, *Wear* **168**, 17 (1993)
132. N. Sasaki, M. Tsukada, S. Fujisawa, Y. Sugawara, S. Morita, K. Kobayashi, *Phys. Rev. B* **57**, 3785–3786 (1998)
133. B. Bhushan, A.V. Kulkarni, V.N. Koinkar, M. Boehm, L. Odoni, C. Martelet, M. Belin, *Langmuir* **11**, 3189 (1995)
134. R.M. Overney, H. Takano, M. Fujihira, W. Paulus, H. Ringsdorf, *Phys. Rev. Lett.* **72**, 3546–3549 (1994)
135. H. Ahn, P.D. Cuong, S. Park, Y. Kim, J. Lim, *Wear* **255**, 819–825 (2003)
136. J.A. Schey, *Tribology in Metalworking* (American Society for Metals, Metals Park, 1983)
137. R.A. Reich, J.M. Epp, D.E. Gantzer, *Tribol. Transact.* **39**, 23 (1996)
138. J.A. Schey, *Introduction to Manufacturing Processes* (McGraw Hill, New York 1987)
139. W. Tong, Hector, L.G. Jr., H. Weiland, L.F. Wieserman, *Scripta Mater.* **36**, 1339 (1997)
140. K. Mietzner, *Stahl Eisen* **81**, 950 (1961)
141. H. Takechi, H. Kato, T. Sunami, T. Nakayama, *Trans. JIM* **8**, 233 (1967)
142. H.C. Chao, *ASM Trans. Q* **60**, 37 (1967)
143. K. Osakada, M. Oyane, *Jpn. Soc. Mech. Eng.* **14**, 171 (1971)
144. M. Fukuda, K. Yamaguchi, N. Takakura, Y. Sakano, *J. Jpn. Soc. Technol. Plast.* **15**, 994 (1974)
145. D.V. Wilson, W.T. Roberts, P.M.B. Rodrigues, *Metallurg. Mater. Transact. A* **12**, 1603–1611 (1981)
146. C.W. Snyder, B.G. Orr, D. Kessler, L.M. Sander, *Phys. Rev. Lett.* **66**, 3032–3035 (1991)
147. C.A. Bronkhorst, S.R. Kalidindi, L. Anand, *Philos. Transact. Phys. Sci. Eng.* **341**, 443–477 (1992)
148. K. Bethke, M. Holscher, K. Lucke, *Mater. Sci. Forum* **157**, 1137–1144 (1994)
149. T. Mizuno, H. Mulki, *Wear* **198**, 176–184 (1996)
150. M. Jain, D.J. Lloyd, S.R. Macewen, *Int. J. Mech. Sci.* **38**, 219–232 (1996)
151. R. Mahmudi, M. Mehdizadeh, *J. Mater. Proc. Technol.* **80**, 707 (1998)
152. R. Becker, *Acta Mater.* **46**, 1385–1401 (1998)
153. P.S. Lee, H.R. Piehler, B.L. Adams, G. Jarvis, H. Hampel, A.D. Rollett, *J. Mater. Proc. Technol.* **80–81**, 315–319 (1998)
154. N.J. Wittridge, R.D. Knutsen, *Mater. Sci. Eng. A* **269**, 205–216 (1999)
155. O. Engler, E. Brünger, *Mater. Sci. Forum* **396–402**, 345–350 (2002)
156. D. Raabe, M. Sachtleber, H. Weiland, G. Scheele, Z. Zhao, *Acta Mater.* **51**, 1539–1560 (2003)
157. Z. Zhao, R. Radovitzky, A. Cuitiño, *Acta Mater.* **52**, 5791–5804 (2004)
158. M. Sachtleber, D. Raabe, H. Weiland, *J. Mater. Proc. Technol.* **148**, 68–76 (2004)
159. S. Sheu, W.R.D. Wilson, *Manuf. Eng. Trans.* **11**, 172–178 (1983)
160. W.R.D. Wilson, S. Sheu, *Int. J. Mech. Sci.* **30**, 475 (1988)
161. W.R.D. Wilson, *J. Eng. Mater. Technol.* **113**, 60 (1991)
162. F. Plouraboué, M. Boehm, *Tribol. Int.* **32**, 45–57 (1999)
163. S.R. Schmid, G. Louis, J. Hector, *Wear* **215**, 257–266 (1998)

164. A.R.C. Westwood, F.E. Lockwood, in *Chemomechanical effects in lubrication*, ed. by J.M. Georges, Microscopic Aspects of Adhesion and Lubrication (Elsevier, Amsterdam 1982)
165. T.A. Harris, M.N. Kotzalas, *Essential Concepts of Bearing Technology* (CRC, Boca Raton 2007)
166. H. Hertz, *J Reine Angew. Math* **92**, 156 (1881)
167. K.L. Johnson, K. Kendall, A.D. Roberts, *Proc. R Soc. Lond. Ser. A Math. Phys. Sci.* **324**, 301–313 (1971)
168. B.V. Derjaguin, V.M. Muller, Y.P.J. Toporov, *Interface Sci.* **53**, 314 (1975)
169. V.M. Muller, V.S. Yushmanov, B.V. Derjaguin, *J. Colloid Interface Sci.* **77**, 91 (1980)
170. B.D. Hughes, L.R. White, *Q J. Mech. Appl. Math.* **32**, 445 (1979)
171. D. Maugis, H.M. Pollock, *Acta Metallurgica* **32**, 1323 (1984)
172. S. Shima, M. Oyane, *Int. J. Mech. Sci.* **18**, 285 (1976)
173. G. Coccoz, M. Bellet, R. Lecot, L. Ackermann, H. Haggblad, Cold compaction of iron powder: experiments and simulation. *Powder Metallurgy World Congress'94*. 709–712 (1994)
174. J. Christoffersen, M.M. Mehrabadi, Nemat- S. Nasser, *J. Appl. Mech.* **48**, 339 (1981)
175. E. Arzt, *Acta Metallurgica* **30**, 1883–1890 (1982)
176. C.D. Turner, in *Powder Metallurgy World Congress'94*. 713–716 (1994)
177. S. Tamura, T. Mitsuno, T. Aizawa, in *Powder Metallurgy World Congress'94*. 753–756 (1994)
178. A.R. Akisanya, A.C.F. Cocks, N.A. Fleck, in *Powder Metallurgy World Congress'94*. 757–760 (1994)
179. N.A. Fleck, *J. Mech. Phys. Solid.* **43**, 1409–1431 (1995)
180. F. Emeriault, B. Cambou, *Int. J. Solid. Struct.* **33**, 2591–2607 (1996)
181. P.L. Larsson, S. Biwa, B. Storåkers, *Acta Materialia* **44**, 3655–3666 (1996)
182. B. Storåkers, N.A. Fleck, R.M. McMeeking, *J. Mech. Phys. Solid.* **47**, 785–815 (1999)
183. R.J. Henderson, H.W. Chandler, A.R. Akisanya, C.M. Chandler, S.A. Nixon, *J. Mech. Phys. Solid.* **49**, 739–759 (2001)
184. M. Abouaf, Modélisation de la compactacion de poudres métalliques frittées. This is a Ph.D thesis from Grenoble University in France (1985)
185. G.G. Weber, S.B. Brown, *Adv. Powder Metall.* **1**, 105–188 (1992)
186. H. Riedel, D. Meyer, J. Svoboda, H. Zipse, *Int. J. Refract. Metal. Hard Mater.* **12**, 55 (1993)
187. D.V. Tran, R.W. Lewis, D.T. Gethin, A.K. Ariffin, *Powder Metall.* **36**, 257 (1993)
188. O. Coube, H. Riedel, *Powder Metall.* **43**, 123–131 (2000)
189. T. Kraft, O. Coube, H. Riedel, *Numerical simulation of pressing and sintering in the ceramic and hard metal industry*, ed. by A. Zavaliangos and A. Laptev, in Recent developments in computer modeling of powder metallurgy process. Published by IOS Press in Burke, VA (2001)
190. T. Kraft, H. Riedel, *Powder Metall.* **45**, 227–231 (2002)
191. P.A. Cundall, O.D.L. Strack, A Discrete numerical model for granular assemblies. *Gsotechnique* **29**, 47–65 (1979)
192. K.J. Trasorras, T.M. Krauss, B.L. Ferguson, in *Modeling of powder compaction using the finite element method*. Advances in Powder Metallurgy Conference, pp. 85–104 (1989)

193. J.P. Bandstra, Otto, W.L. Jr., T.R. Massa, *Finite Element Simulation of Cold Die Compaction Leading to Sintered Part Dimension Control*. This was published in the Advances in Powder Metallurgy Proceedings 1990 Powder Metallurgy Conference Exhibition, pp. 181–196 (1990)
194. J.L. Chenot, F. Bay, L. Fourment, *Int. J. Numer. Method. Eng.* **30**, 1649 (1990)
195. D.T. Gethin, R.W. Lewis, D.V. Tran, J.G. Ashoka, *Adv. Powder Metall. Particulate Materr.* **2**, 11–27 (1992)
196. J. Lian, S. Shima, *Int. J. Numeric. Method. Eng.* **37**, 763 (1994)
197. D.T. Gethin, R.W. Lewis, in *Powder Metallurgy World Congress'94*. 689–692 (1994)
198. S. Shima, H. Kotera, Y. Ujie, *Mater. Sci. Res. Int.* **1**, 163–168 (1995)
199. A.R. Khoei, R.W. Lewis, *Finite Elem. Anal. Design.* **30**, 335 (1998)
200. A.K. Ariffin, D.T. Gethin, R.W. Lewis, *Powder Metall.* **41**, 189 (1998)
201. A.R. Khoei, R.W. Lewis, *Int. J. Numeric. Method. Eng.* **45**, 801 (1999)
202. R.S. Ransing, D.T. Gethin, A.R. Khoei, P. Mosbah, R.W. Lewis, *Mater. Design.* **21**, 263–269 (2000)
203. D.T. Gethin, R.S. Ransing, R.W. Lewis, M. Dutko, A.J.L. Crook, *Comput. Struc.* **79**, 1287–1294 (2001)
204. C.L. Martin, D. Bouvard, S. Shima, *J. Mech. Phys. Solid.* **51**, 667–693 (2003)
205. W.A. Ducker, T.J. Senden, R.M. Pashley, *Nature* **353**, 239–241 (1991)
206. W.A. Ducker, T.J. Senden, *Langmuir* **8**, 1831–1836 (1992)
207. H. Butt, *Biophys. J.* **60**, 1438–1444 (1991)
208. G. Hüttl, D. Beyer, E. Müller, *Surf. Interface Anal.* **25**, 543–547 (1997)
209. M. Kappl, H. Butt, The colloidal probe technique and its application to adhesion force measurements. *Part. Part. Syst. Character.* **19**, 129–143 (2002)
210. D.M. Schaefer, M. Carpenter, R. Reifenberger, L.P. Demejo, D.S. Rimai, *J. Adhes. Sci. Technol.* **8**, 197–210 (1994)
211. D.M. Schaefer, M. Carpenter, B. Gady, R. Reifenberger, L.P. Demejo, D.S. imai, *J. Adhes. Sci. Technol.* **9**, 1049–1062 (1995)
212. S. Biggs, G. Spinks, *J. Adhes. Sci. Technol.* **12**, 461–478 (1998)
213. B. Gady, R. Reifenberger, D.M. Schaefer, R.C. Bowen, D.S. Rimai, L.P. Demejo, W. Vreeland, *J. Adhes.* **67**, 19–36 (1998)
214. L.O. Heim, J. Blum, M. Preuss, H.J. Butt, *Phys. Rev. Lett.* **83**, 3328–3331 (1999)
215. M. Reitsma, V.S.J. Craig, S. Biggs, *J. Adhes.* **74**, 125–142 (2000)
216. M. Reitsma, V.S.J. Craig, S. Biggs, *Int. J. Adhes. Adhesive.* **20**, 445–448 (2000)
217. I.U. Vakarelski, K. Ishimura, K. Higashitani, *J. Colloid. Interface Sci.* **227**, 111–118 (2000)
218. S. Ecke, R. Raiteri, E. Bonaccorso, C. Reiner, H.J. Deiseroth, H.J. Butt, *Rev. Sci. Instrum.* **72**, 4164–4170 (2001)
219. I.U. Vakarelski, K. Higashitani, *J. Colloid. Interface Sci.* **242**, 110–120 (2001)
220. I.U. Vakarelski, A. Toritani, M. Nakayama, K. Higashitani, *Langmuir* **17**, 4739–4745 (2001)
221. G. Toikka, G.M. Spinks, H.R. Brown, *Langmuir* **17**, 6207–6212 (2001)
222. L.O. Heim, S. Ecke, M. Preuss, H.J. Butt, *J. Adhes. Sci. Technol.* **16**, 829–844 (2002)
223. M. Fuji, K. Machida, T. Takei, T. Watanabe, M. Chikazawa, *J. Phys. Chem.* **102**, 8782–8787 (1998)
224. M. Fuji, K. Machida, T. Takei, T. Watanabe, M. Chikazawa, *Langmuir* **15**, 4584–4589 (1999)

225. G. Toikka, R.A. Hayes, J. Ralston, J. Colloid. Interface Sci. **180**, 329–338 (1996)
226. R.W. Bowen, T.A. Doneva, J. Colloid. Interface Sci. **229**, 544–549 (2000)
227. K. Cooper, A. Gupta, S. Beaudoin, J. Colloid. Interface Sci. **228**, 213–219 (2000)
228. K. Cooper, N. Ohler, A. Gupta, S. Beaudoin, J. Colloid. Interface Sci. **222**, 63–74 (2000)
229. Y.I. Rabinovich, J.J. Adler, A. Ata, R.K. Singh, B.M. Moudgil, J. Colloid. Interface Sci. **232**, 17–24 (2000)
230. Y.I. Rabinovich, J.J. Adler, A. Ata, R.K. Singh, B.M. Moudgil, J. Colloid. Interface Sci. **232**, 10–16 (2000)
231. K. Cooper, A. Gupta, S. Beaudoin, J. Colloid. Interface Sci. **234**, 284–292 (2001)
232. G.J.C. Braithwaite, A. Howe, P.F. Luckham, Langmuir **12**, 4224–4237 (1996)
233. G.J.C. Braithwaite, P.F. Luckham, J. Chem. Soc. Faraday Trans. **93**, 1409–1415 (1997)
234. M. Giesbers, J.M. Kleijn, G.J. Fleer, C. Stuart, Forces between polymer-covered surfaces: a colloidal probe study. Colloid Surf. A Physicochem. Eng. Asp. **142**, 343–353 (1998)
235. K. Feldman, T. Tervoort, P. Smith, N.D. Spencer, Langmuir **14**, 372–378 (1998)
236. H.G. Pedersen, L. Bergström, J. Am. Ceram. Soc. **52**, 1137–1145 (1999)
237. G.J.C. Braithwaite, P.F. Luckham, A.M. Howe, J. Colloid. Interface Sci. **213**, 525–545 (1999)
238. R.G. Cain, N.W. Page, S. Biggs, Phys. Rev. E **62**, 8369–8379 (2000)
239. A. Feiler, I. Larson, P. Jenkins, P. Attard, Langmuir **16**, 10269–10277 (2000)
240. S. Ecke, R. Raiteri, E. Bonaccurson, C. Reiner, H.J. Deiseroth, H.J. Butt, Rev. Sci. Instrum. **72**, 4164–4170 (2001)
241. S. Ecke, H.J. Butt, J. Colloid. Interface Sci. **244**, 432–435 (2001)
242. S. Zauscher, D.J. Klingenberg, Friction between cellulose surfaces measured with colloidal probe microscopy. Colloid Surf. A Physicochem. Eng. Asp. **178**, 213–229 (2001)
243. S. Malkin, *Grinding Technology: Theory & Applications of Machining with Abrasives* (SME, Dearborn, 1989)
244. M.C. Shaw, *Principles of Abrasive Processing* (Clarendon, Oxford, 1996)
245. S.S. Law, S.M. Wu, J. Eng. Industry. **95**, 972–978 (1973)
246. L.C. Li, J.Z. Fu, Ann CIRP **29**, 245–249 (1980)
247. K. Steffens, Ann CIRP **32**, 255–259 (1983)
248. M. Younis, M.M. Sadek, El- T. Wardani, J. Eng. Ind **109**, 306–313 (1987)
249. M. Hashish, J. Eng. Mater. Technol. **111**, 154 (1989)
250. H.K. Toenshoff, J. Peters, I. Inasaki, T. Paul, Ann CIRP **41**, 677 (1992)
251. N. Chiu, S. Malkin, Ann CIRP **42**, 383–387 (1993)
252. C. Xun, W.B. Rowe, Int. J. Mach Tool. Manuf. **36**, 883–896 (1996)
253. G. Warnecke, U. Zitt, Ann CIRP **47**, 265–270 (1998)
254. A.M. Hoogstrate, B. Karpuschewski, van C.A. Luttervelt, H.J.J. Kals, CIRP Ann. Manuf. Technol. **51**, 263–266 (2002)
255. G. Nanz, L.E. Camilletti, IEEE Trans. Semicond. Manuf. **8**, 382–389 (1995)
256. H. Liang, D.R. Craven, *Tribology in Chemical-Mechanical Planarization* (2005)
257. M.J. Post, S.R. Schmid, T.C. Ovaert, M.P. Laurent, in *ICTMP International Conference* (2007)

258. L.M. Cook, *J. Non-Crystal. Solid.* **120**, 152 (1990)
259. P.A. Burke, in *VLSI Multilevel Interconnection Conference, 1991, Proceedings, Eighth International IEEE*, pp. 379–384 (1991)
260. J. Warnock, *J. Electrochem. Soc.* **138**, 2398–2402 (1991)
261. S.R. Runnels, P. Renteln, *Dielect. Sci. Technol.* **6**, 110–121 (1993)
262. T. Yu, C.C. Yu, M. Orlowski, *International Electron Devices Meeting Technical Digest.* (1993)
263. S. Sivaram, R. Tolles, H. Bath, E. Lee, *Mater. Res. Soc. Symp. Proc.* **260**, 53 (1993)
264. S.R. Runnels, *J. Electrochem. Soc.* **141**, 1900 (1994)
265. S.R. Runnels, L.M. Eyman, *J. Electrochem. Soc.* **141**, 1698 (1994)
266. D. Devecchio, P. Schmutz, G.S. Frankel, *Electrochem. Solid-State Lett.* **3**, 90 (2000)
267. C. Cho, S. Park, Y. Ahn, *Thin Solid Film.* **389**, 254–260 (2001)
268. J. Luo, D. Dornfeld, *IEEE Trans. Semiconductor Manuf.* **14**, 112 (2001)
269. G. Ahmadi, X. Xia, *J. Electrochem. Soc.* **148**, G99 (2001)
270. G. Fu, A. Chandra, S. Guha, G. Subhash, *IEEE Trans. Semiconductor Manuf.* **14**, 406–417 (2001)
271. Y. Zhao, L. Chang, *Wear* **252**, 220 (2002)
272. E. Paul, *J. Electrochem. Soc.* **149**, G305 (2002)
273. Y. Zhao, L. Chang, S.H. Kim, *Wear* **254**, 332–339 (2003)
274. J. Seok, C.P. Sukam, A.T. Kim, J.A. Tichy, T.S. Cale, *Wear* **254**, 307–320 (2003)
275. K. Qin, B. Moudgil, C. Park, *Thin Solid Film.* **446**, 277–286 (2004)
276. Y.Y. Lin, S.P. Lo, *Int. J. Adv. Manuf. Technol.* **23**, 644–650 (2004)
277. W. Che, Y. Guo, A. Chandra, A. Bastawros, *J. Manuf. Sci. Eng.* **127**, 545–554 (2005)
278. T.K. Berdyeva, S.B. Emery, I.Y. Sokolov, *Electrochem. Solid-State Lett.* **6**, G91–G94 (2003)
279. J. Liu, M. Post, S.R. Schmid, T.C. Ovaert, *Nanoscale Simulation of Chemical Mechanical Polishing* (2006)
280. M.C. Shaw, G.J. DeSalvo, *Trans. ASME* **92**, 480 (1970)
281. M. Wendel, B. Irmer, J. Cortes, R. Kaiser, H. Lorenz, J.P. Kotthaus, A. Lorke, *Superlattice. Microstruct.* **20**, 349–356 (1996)
282. A.A. Tseng, A. Notargiacomo, T.P. Chen, *J. Vacuum Sci. Technol. B* **23**, 877–894 (2005)
283. R. Magno, B.R. Bennett, *Appl. Phys. Lett.* **70**, 1855–1857 (1997)
284. S. Hu, A. Hamidi, S. Altmeyer, T. Koster, B. Spangenberg, H. Kurz, *J. Vacuum Sci. Technol. B Microelectron. Nanometer Struct.* **16**, 2822–2824 (1998)
285. A. Notargiacomo, V. Foglietti, E. Cianci, G. Capellini, M. Adami, P. Faraci, F. Evangelisti, C. Nicolini, *Nanotechnology* **10**, 458 (1999)
286. H. Sugihara, A. Takahara, T. Kajiyama, *J. Vacuum Sci. Technol. B Microelectron. Nanometer Struct.* **19**, 593–595 (2001)
287. U. Kunze, *Superlattice. Microstruct.* **31**, 3–17 (2002)
288. L. Santinacci, T. Djenizian, H. Hildebrand, S. Ecoffey, H. Mokdad, T. Campanella, P. Schmuki, *Electrochimica Acta* **48**, 3123–3130 (2003)
289. H.D.F. Filho, M.H.P. Maurício, C.R. Ponciano, R. Prioli, *Mater. Sci. Eng. B* **112**, 194–199 (2004)
290. H.J. Mamin, *Appl. Phys. Lett.* **69**, 433–435 (1996)

291. P. Vettiger, G. Cross, M. Despont, U. Drechsler, U. Durig, B. Gotsmann, W. Haberle, M.A. Lantz, H.E. Rothuizen, R. Stutz, G.K. Binnig, *IEEE Trans. Nanotechnol.* **1**, 39–55 (2002)
292. A. Majumdar, P.I. Oden, J.P. Carrejo, L.A. Nagahara, J.J. Graham, J. Alexander, *Appl. Phys. Lett.* **61**, 2293–2295 (1992)
293. S.W. Park, H.T. Soh, C.F. Quate, S.I. Park, *Appl. Phys. Lett.* **67**, 2415–2417 (1995)
294. K. Wilder, C.F. Quate, D. Adderton, R. Bernstein, V. Elings, *Appl. Phys. Lett.* **73**, 2527–2529 (1998)
295. M. Ishibashi, S. Heike, H. Kajiyama, Y. Wada, T. Hashizume, *Appl. Phys. Lett.* **72**, 1581–1583 (1998)
296. T. Shiokawa, Y. Aoyagi, M. Shigeno, S. Namba, *Appl. Phys. Lett.* **72**, 2481–2483 (1998)
297. P. Davidsson, A. Lindell, T. Makela, M. Paalanen, J. Pekola, *Microelectron. Eng.* **45**, 1 (1999)
298. P.M. Campbell, E.S. Snow, P.J. McMarr, *Appl. Phys. Lett.* **66**, 1388–1390 (1995)
299. D. Wang, L. Tsau, K.L. Wang, P. Chow, *Appl. Phys. Lett.* **67**, 1295–1297 (1995)
300. E.S. Snow, D. Park, P.M. Campbell, *Appl. Phys. Lett.* **69**, 269–271 (1996)
301. R. Martel, T. Schmidt, H.R. Shea, T. Hertel, P. Avouris, *Appl. Phys. Lett.* **73**, 2447–2449 (1998)
302. A. Boisen, K. Birkelund, O. Hansen, F. Grey, J. Vacuum Sci. Technol. B *Microelectron. Nanometer Struct.* **16**, 2977–2981 (1998)
303. U.F. Keyser, H.W. Schumacher, U. Zeitler, R.J. Haug, K. Eberl, *Appl. Phys. Lett.* **76**, 457–459 (2000)
304. K. Matsumoto, Y. Gotoh, T. Maeda, J.A. Dagata, J.S. Harris, *Appl. Phys. Lett.* **76**, 239–241 (2000)
305. J. Jaworski, S.R. Schmid, J.E. Wang, *J. Manuf. Sci. Eng.* **121**, 232–238 (1999)
306. C. Huang, S.R. Schmid, J.E. Wang, *J. Manuf. Sci. Eng.* **123**, 225–230 (2001)
307. D.T. Clark, W.J. Feast, *Polymer Surfaces* (Wiley, New York, 1978)

Scanning Probe Microscopy as a Tool Applied to Agriculture

Fabio Lima Leite, Alexandra Manzoli, Paulo Sérgio de Paula Herrmann Jr, Osvaldo Novais Oliveira Jr, and Luiz Henrique Capparelli Mattoso

Summary. The control of materials properties and processes at the molecular level inherent in nanotechnology has been exploited in many areas of science and technology, including agriculture where nanotech methods are used in release of herbicides and monitoring of food quality and environmental impact. Atomic force microscopy (AFM) and related techniques are among the most employed nanotech methods, particularly with the possibility of direct measurements of intermolecular interactions. This chapter presents a brief review of the applications of AFM in agriculture that may be categorized into four main topics, namely thin films, research on nanomaterials and nanostructures, biological systems and natural fibers, and soils science. Examples of recent applications will be provided to give the reader a sense of the power of the technique and potential contributions to agriculture.

Key words: Agriculture, Atomic force microscopy, Atomic force spectroscopy, Nanoscience, Nanotechnology.

26.1 Applications of Nanotechnology in Agriculture

The control of materials and processes at the molecular length scale makes nanotechnology a powerful tool for developing many areas, from biology to materials science [1] and agriculture [2]. Of particular importance are the self-assembly methods that allow fabrication of highly controlled structures in nanowires [3,4], self-assembled molecules and particles [5,6], and 3D architectures. Some of the novel materials created with nanotechnology may be used in devices and instrumentation that may be useful for food safety, environment protection, and biosecurity [7,8]. In addition, controlling the processes of crop production, food processing, and storage may increase the efficiency of agricultural methods.

Various examples of applications to agriculture have been provided in recent years. For instance, the concept of controlled release, which is now pervasive in drug release, has been exploited in pesticides and fertilizers, with obvious advantages for the environment [9,10]. It can also be used to deliver

hormones and vaccines to the livestock [11], or DNA for genetic engineering of plants [12]. With regard to food control and diagnosis, the applications stem from sensors and biosensors made of carbon nanotubes (CNTs) to detect diseases in plants and animals [13–16] to biochips and intelligent tags for analysis of gene expression and intracellular histopathology [17, 18]. Natural products from agriculture have also found uses when in the form of nanostructures. Indeed, electrospun cellulose nanofibers [19] can be used in air filtration, protective clothing, biodegradable nanocomposites, and biodegradable cellulose mats to absorb fertilizers and pesticides.

Of the many experimental tools of nanotechnology, the family of surface probe microscopy techniques [20] is certainly among the most used. These techniques stemmed from the scanning tunneling microscope [21] and then the atomic force microscope (AFM) [22]. The use of AFM for agriculture-related applications is described in Sect. 26.2.

26.2 Applications of AFM in Agriculture

26.2.1 Introduction

Agriculture and agribusiness have benefited from the use of microscopy techniques, such as transmission/scanning electron microscopy (SEM/TEM), confocal laser scanning microscopy and AFM, for investigating fruits and vegetables [23–25]. Further contributions have also come from knowledge on sensors, development of new drugs, and DNA chip technology. AFM, in particular, is advantageous for providing resolution at the nanoscale and allowing direct measurements of forces between the tip and the sample, in the so-called atomic force spectroscopy (AFS) [20]. The interest in using AFM for agriculture has grown steadily since 1991, as shown in Fig. 26.1. In that year, AFM was introduced as a tool for monitoring colloidal forces (*times cited: 877*) [26]. Applications of AFM in agriculture can be divided into four main categories: (1) thin films, (2) fundamental research on nanomaterials and nanostructures, (3) biological systems and natural fibers, and (4) soils science.

26.2.2 Some Examples and Results of Agricultural Research

Here we shall present illustrative examples of possible uses of AFM in agricultural research, according to the four main areas mentioned in the previous paragraph. Obviously, this survey covers only part of the material published in the literature, and emphasis will be placed in work produced by our research groups.

Nanostructured Films

The connection between nanostructured films and agriculture is established by the use of thin films in sensors and biosensors applied to agriculture, in

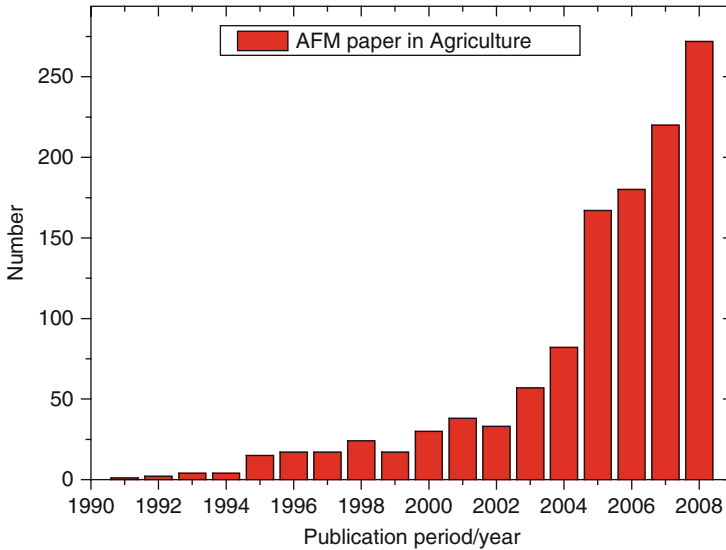


Fig. 26.1. Publications involving AFM applied to agriculture, according to the ISI web of science

addition to cases where such films are adsorbed on agricultural products for specific purposes, as is the case of fruits with retarded ripening owing to a protecting thin film [27]. These thin films may be obtained with techniques such as the Langmuir-Blodgett (LB) [28] or the layer-by-layer (LbL) method [29], which offer precise control of thickness and molecular architecture. Among the many materials used in thin films for agriculture, conducting polymers should be singled out due to the extensive research in the last years. In the following, a short description will be given of results obtained with AFM for these polymers.

Lobo et al. [30] used an AFM to monitor the growth of LbL films from poly (*o*-methoxyaniline) (POMA) alternated with polyvinyl sulfonic acid (PVS) deposited on various types of substrate. The samples were characterized with AFM operating in the contact, friction, and tapping modes, which then allowed surface properties to be studied, including roughness, mean grain size, grain boundaries, and film thickness (using the tip as a scraping tool). The authors showed that surface roughness increased with the number of bilayers until a constant value was reached, and depended strongly on the substrate onto which the films were deposited. For example, Fig. 26.2 illustrates the large difference in film morphology as the substrate changed from hydrophilic to hydrophobic glass. For the other substrates investigated (silicon oxide, gold, ITO, and mica), film morphology displayed intermediate behavior with the grain size depending particularly on the wettability of the substrate.

While film morphology does depend on the substrate, there is also a strong dependence on the material used to produce the films. Of particular relevance

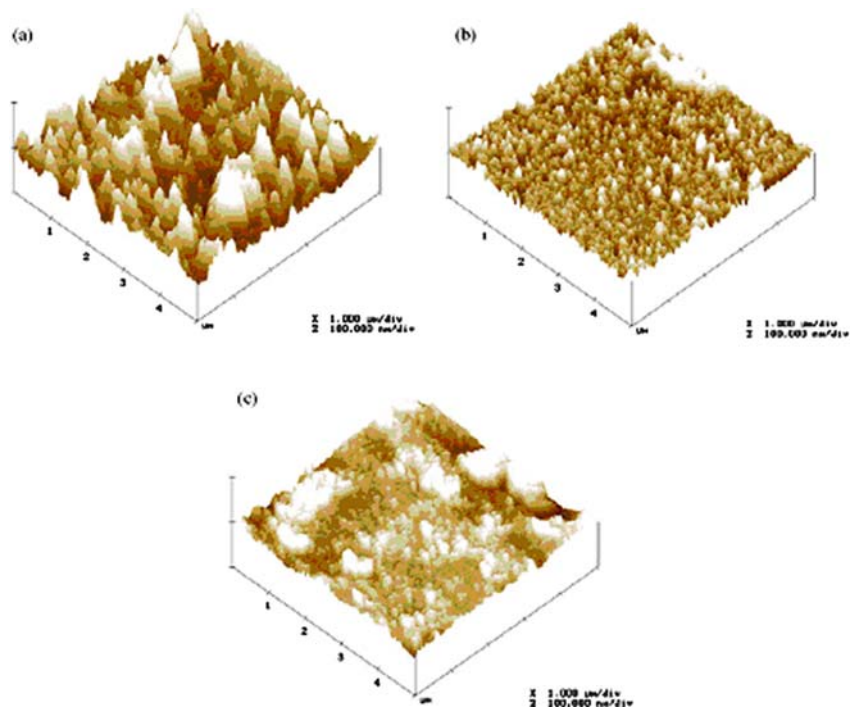


Fig. 26.2. Tapping mode 3D height images corresponding to ten bilayers on several substrates: (a) hydrophobic, (b) hydrophilic glass, and (c) gold (reprint from [30])

are those molecules that tend to assemble themselves into planar structures. This was observed by Paterno and Mattoso, who reported very flat films obtained with sulfonated lignin (SL) [31]. In LbL films of poly(*o*-ethoxyaniline) (POEA) alternated with SL, they found that the surface roughness varied with the number of deposited layers, being higher for POEA than for SL layers, as a result of the grain morphology of POEA layers and the smoother, planar surface of SL layers. Therefore, lignins are potentially useful for applications requiring smooth films.

The semiconducting polymer POEA has also been investigated using AFM with regard to fractal dimension and relative surface coverage [32]. Studying the kinetics of polymer adsorption was useful to optimize the film fabrication conditions. In fact, the combination of AFM and optical absorption spectroscopy allowed one to demonstrate that LbL films could be produced within very short time intervals, which has important implications for possible use in real devices [33]. From the point of view of fundamental adsorption processes, morphological features could be correlated with film growth models. For AFM images indicated a globular morphology in films containing one or more layers of POEA, which is typical of a nucleation adsorption process, where the first polymer chains are anchored to the substrate and serve as nuclei for further

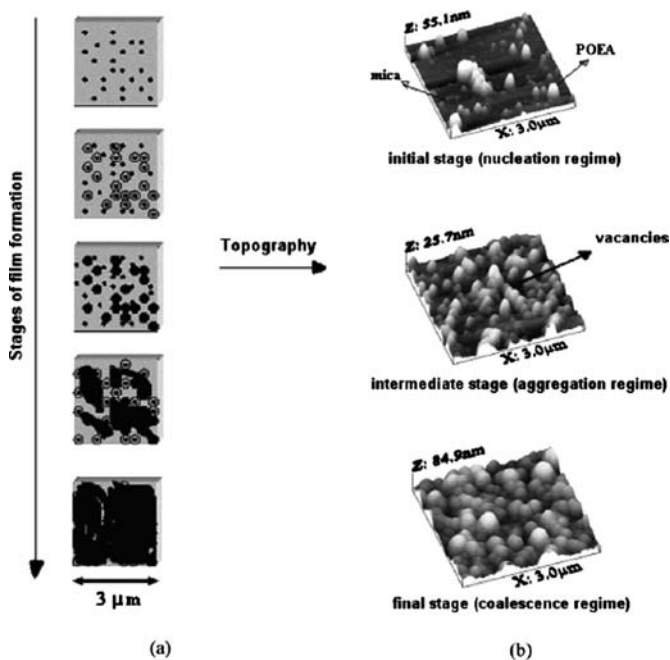


Fig. 26.3. (a) Typical model of progressive nucleation, growth, and coalescence; (b) AFM images of POEA on muscovite mica for initial, intermediate, and final stages of adsorption (reprint from [34])

polymer adsorption. Figure 26.3a depicts a possible model for film growth, which includes nucleation, aggregation, and coalescence processes – with the corresponding AFM images for a POEA film in Fig. 26.3b [34].

The AFM images in Fig. 26.3b show only some snapshots of the growth process. In the upper image, corresponding to the initial stage of growth, a number of scattered domains are seen, which is consistent with a nucleation regime. For immersion times higher than 60 s and up to ca. 180 s, aggregation is observed, as can be seen in the image in the middle of Fig. 26.3b, where the substrate is practically fully covered. Coalescence occurs for longer periods of immersion, which is denoted by a smaller number of larger domains. These morphological features of nanostructured films obtained with AFM have been correlated with the chain conformations in solution [35], for which a dummy atom model (DAM) was used. The AFM images represented an “off print” of the solution conformation when molecules of emeraldine base POEA (POEA-EB) were adsorbed on the substrate (*see* Fig. 26.4). The size of the globules obtained in the films is comparable with aggregates of polymer molecules in solution. For nanostructured films from polyanilines, several studies have shown that adsorption mechanisms and film properties are entirely altered

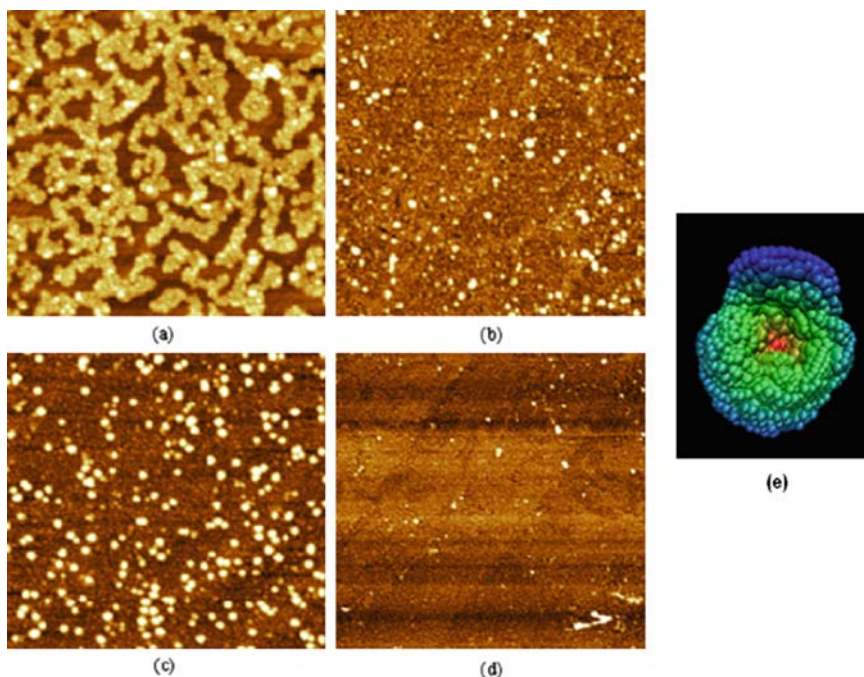


Fig. 26.4. AFM images ($10 \times 10 \mu\text{m}$) of POEA-EB (pH 10.0) for immersion time of (a) 1 s, (b) 3 s, (c) 5 s, (d) 180 s, and (e) average dummy atom model for POEA-EB (adapted from [35])

by a mere change in pH [36]. Indeed, this dependence has been exploited in sensors and other applications [37, 38].

Another important issue is related to the dependence of the amount of POEA adsorbed on the type of acid used for doping. Paterno and Mattoso [39] used UV-Vis spectra and AFM images to infer that the amount of POEA in the films increased according to the type of acid in the following order: $\text{HCl} < \text{CSA} < \text{TSA} < \text{HNO}_3 < \text{H}_2\text{SO}_4$. They verified that organic, bulkier dopant acids led to a larger amount of polymer adsorbed in comparison to films fabricated with HCl. This is explained by the fact that the counterions from these acids are less solvated by water. Therefore, they are more strongly associated with the protonated nitrogen atoms from POEA, leading to enhanced charge screening. Intra- and intermolecular electrostatic repulsions between doped POEA chains are thus minimized, allowing the approximation of a greater amount of polymeric segments and inducing the polymer to assume a more compact molecular conformation.

In addition to nanostructured films produced with the LbL method, LB films have also been studied aiming at applications in agriculture-related topics. For example, LB films of polyaniline and ruthenium complex [40] exhibited electrical and electrochemical properties that differed completely

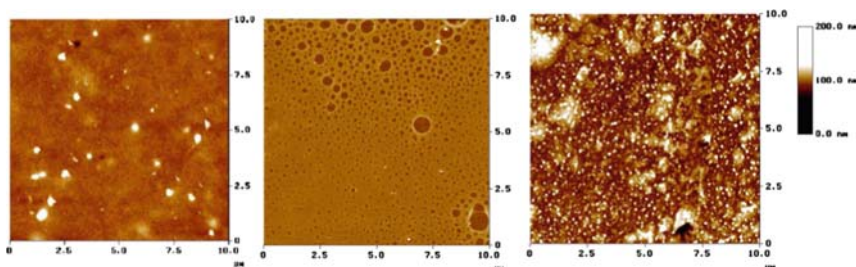


Fig. 26.5. AFM images of OC_1OC_6 – PPV films processed via (a) casting, (b) spin coating and (c) LB technique (reprint from [41])

from those obtained for cast or spin-coated films of the same materials, which was attributed to the intimate contact between the molecules of the two components. Marletta et al. [41] investigated the effects from the film-forming techniques on the optical properties, absorption and emission of poly(2-methoxy-5-hexyloxy-p-phenylenevinylene) OC_1OC_6 – PPV films. Three techniques were used to produce the films, namely casting, spin coating, and LB. As Fig. 26.5 shows, film morphology for OC_1OC_6 – PPV varied strongly with the film processing technique. Furthermore, the morphological feature that affected most the luminescence properties was the thickness, with the thick cast film displaying a large structural disorder. On the contrary, the organization that could be induced by the LB technique appears neither to lead to smooth films nor to affect the emission properties in comparison with the cast films.

Using AFM, Riul Jr. et al. [32] obtained information on the surface roughness and on the presence of different phases (topography and lateral force, respectively) of composite LB films containing different weight percentages of 16-mer polyaniline and cadmium stearate. The results pointed to a phase separation of 16-mer and cadmium stearate in the composite, with cadmium stearate being distributed within the polymer matrix. Gaffo et al. [42] used AFM to investigate the degree of mixing and homogeneity in LB films of ytterbium bisphthalocyanine ($YbPc_2$) and stearic acid (SA). The nanoscale morphology inferred from AFM for LB films on mica was similar to the micrometer scale morphology obtained with micro-Raman images. A mixed LB film with 75/25% $YbPc_2$ /SA presented smaller, more homogeneous distribution of aggregates compared with a mixed LB film of 25/75% $YbPc_2$ /SA, whose results were confirmed with the height profiles and with the root-mean-square roughness. AFM was also used to investigate the morphology of pure and mixed LB films of ruthenium bi- and triphosphines complexes with SA [43] and of LB films made from lignins [44, 45]. With the latter work, it was possible to demonstrate that lignins from distinct sources led to very smooth LB films, in agreement with the low-roughness LbL films of SL already mentioned here. It seems therefore that the small roughness observed in nanostructured films

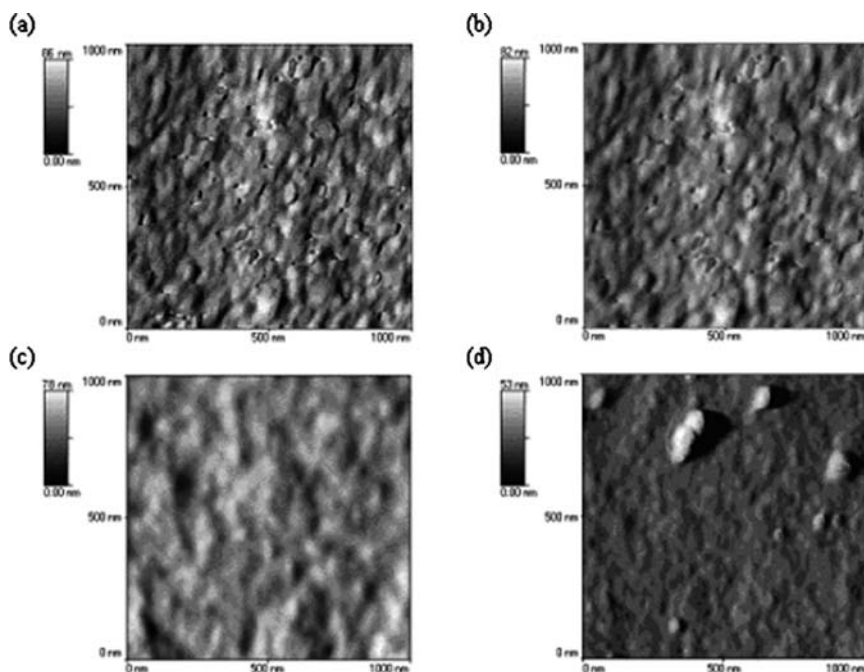


Fig. 26.6. AFM images for (a) as-prepared hydrophilic fiber, (b) fiber coated with 5-min immersion in the surfactant, (c) fiber coated after 10-min immersion, and (d) fiber coated after 12-min immersion. The decreased roughness indicates lysozyme deposition. The ridges seen after 12 min indicate excess of deposited enzyme (reprint from [46])

of lignins should be ascribed to the way these molecules assemble themselves, rather than to the method of film fabrication.

Protein films that do not affect the permeability of the medium have been studied for bioactive membranes and reactors. One example is the work of Assis et al. [46], where lysozyme film was formed via self-assembly for bioactive membranes. The surface roughness measured with an AFM was an indirect indicator of film formation, as depicted in Fig. 26.6. The roughness decreased with increasing adsorption time, which was attributed to the coating of fibers by a homogeneous lysozyme film. Lysozyme was also immobilized on a glass surface for remediation of the microorganism *Escherichia coli* JM 109 into fresh water and saline solutions [47]. The authors verified that the homogeneous enzymatic film was strongly coupled with the glass surface.

Structures Containing Nanoparticles and Nanofibers

The widespread use of nanoparticles and other nanostructures has also reached agriculture. In addition to their use in fabricating sensors and biosensors, as

will be commented on in Sect. 26.2.2, the nanoparticles are useful for pesticides, fertilizers, food additives, cosmetics, veterinary medicine, and packaging [48]. Such applications normally exploit their large surface area that typically results in greater chemical reactivity, biological and catalytic activity than larger particles of the same chemical composition. Furthermore, nanomaterials have greater access to a human body (known as bioavailability), yielding enhanced uptake into individual cells, tissues, and organs.

AFM measurements were employed by Medeiros et al. [49] to investigate the morphology of blends of PANI and poly(methacrylic acid) (PMMA), prepared using a two-stage core-shell particle synthesis. In the first step, the acrylic latex was obtained, after which PANI was deposited and seemed to destroy the regularity of the primary acrylic latex particles. These blended particles may be useful for electrorheological fluids. Goulet et al. [50] fabricated LbL films of dendrimers and metallic nanoparticles, which were employed as substrates for surface-enhanced Raman scattering (SERS) of 2-naphthalenethiol. SERS involves large increases in Raman scattering cross sections of molecules adsorbed at the surface of nanometric-scale metallic particles, which is useful for analytical applications, including single molecule detection [51].

In Sect. 26.2.2, thin films of polyanilines that may be used for agriculture-related research were mentioned. These polymers are also amenable to produce structures such as nanotubes [52], nanoparticles [53], and nanofibers [54]. AFM was used to study LbL films of POEA, poly(sodium-4-styrene sulfonate), and positively charged maghemite ($\gamma - \text{Fe}_2\text{O}_3$) nanoparticles, where these nanocomposites exhibited conducting and magnetic properties [53]. Such magnetic films can be used in magnetic and spintronic devices, magnetic recording materials, sensors, electromagnetic interference shielding, and capacitors. AFM images showed the typical globular morphology of LbL films of polyanilines. The size of the globules increased with the amount of maghemite added to the POEA solution.

Nanoparticles and nanofibers (diameter 40–80 nm) of polyaniline were synthesized by electrodeposition using a potentiodynamic electrochemical method [54]. The type, amount, and dimensions of nanostructures depended on the electropolymerization conditions such as monomer concentration, scan rate, and the range of electrical potential used. These nanostructures may open new possibilities for the preparation of sensors for liquids (electronic tongue) and gases (electronic nose). Figure 26.7 displays SEM and AFM images of PANI prepared using different experimental conditions, where nanoparticles assemble into interconnected agglomerates and then form branched network-like nanofibers.

Sensors and Biosensors

Food control, environment monitoring, and water quality are topics akin to agriculture and agribusiness, for which the development of novel sensors and

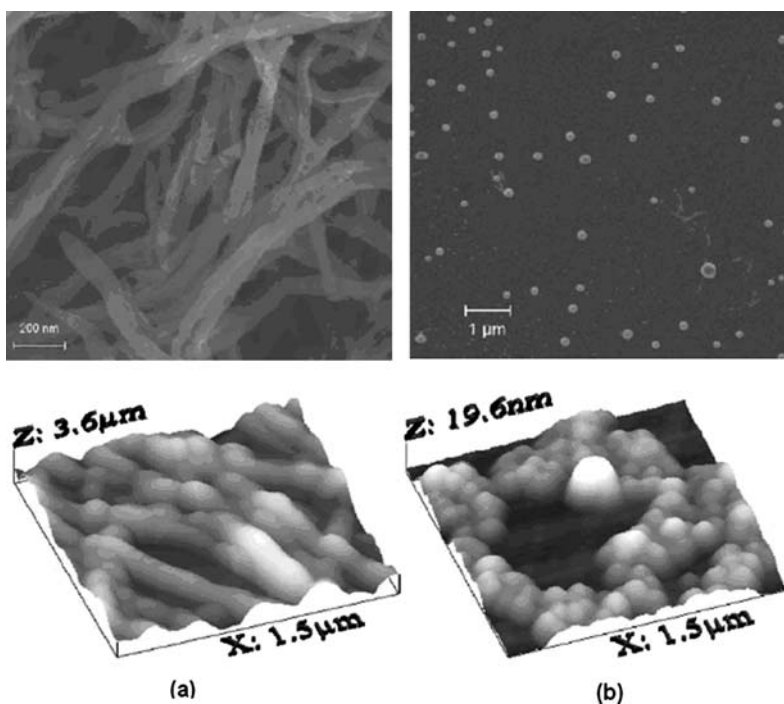


Fig. 26.7. SEM and AFM images of PANI nanofibers (a) and nanoparticles (b) (reprint from [54])

biosensors is crucial. These may be based on a variety of detection methods, including optical measurements, in addition to electrical and electrochemical techniques. In a number of cases, a high sensitivity and selectivity are only achieved if the sensing material is in the form of a nanostructured film, which then requires characterization of features at the nanoscale. Of particular relevance is film morphology, which has been shown to be determinant for the sensitivity [55]. Therefore, AFM has been used to characterize the films that make up the sensing units in several works. For the sensing ability in liquids, for instance, one has to investigate how minute changes in the liquid affect the film properties.

Nanostructured films of chitosan (CS), a polysaccharide obtained from the desacetylation of chitin, have been fabricated for various applications, including bactericide and antimicrobial agents [56]. CS is soluble in acidic media and can be used to produce LbL films employed as sensing units in taste sensors [57] and as templates for immobilization of enzymes for biosensing [58]. The main interest in CS in agriculture, in the form of thin films, is to exploit its biocompatibility and ability to complex heavy metal ions [59], in addition to its strong interaction with cholesterol and other biomolecules [60]. Borato et al. [61] compared the morphology of CS and CS-PMMA nanoparticles in

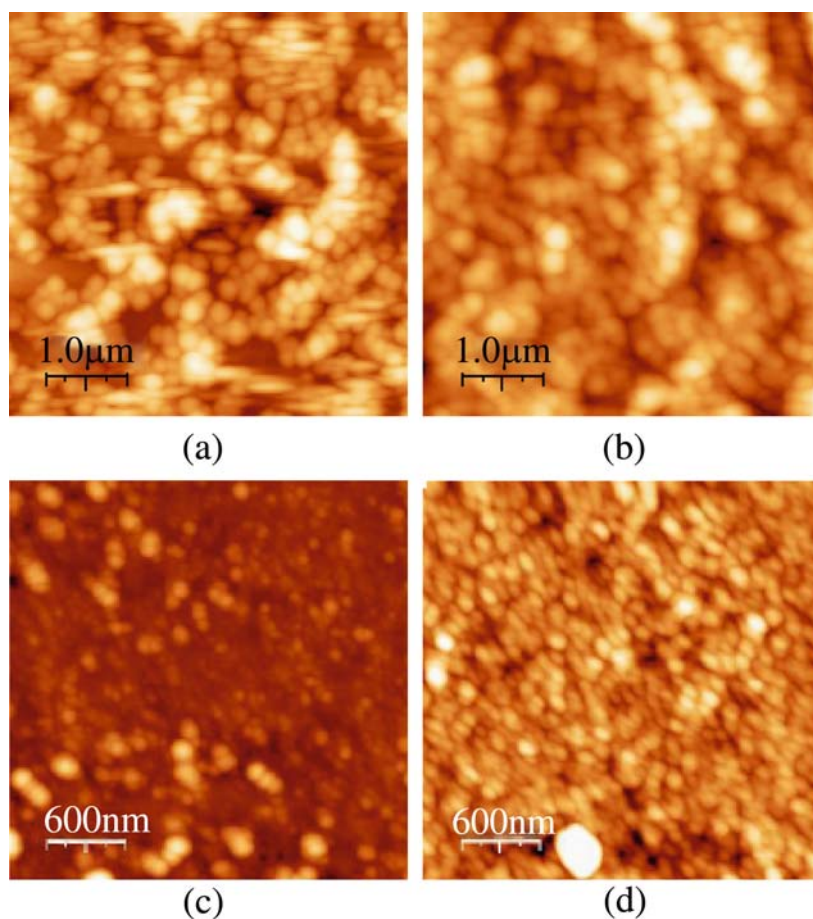


Fig. 26.8. AFM images of (a) doped POEA alternated with CS-PMMA nanoparticles, (b) POEA-EB alternated with CS-PMMA nanoparticles, (c) POEA-ES alternated with chitosan, and (d) POEA-EB alternated with chitosan (reprint from [61])

LbL films with alternated layers of POEA in contact with heavy metals, as illustrated in Fig. 26.8. The nanostructured films were deposited on chromium electrodes to form a sensor array (electronic tongue) based on impedance spectroscopy. This system was used to detect copper ions in aqueous solutions down to the micromolar level.

Siqueira et al. [62] reported on electroactive films made with polyamidoamine (PAMAM) dendrimers, multiwalled CNTs, and alternated with nickel tetrasulfonated metallophthalocyanines (NiTsPc), and the films were used in sensors to detect dopamine in the presence of the interferent ascorbic acid. The incorporation of CNTs in LbL films enhanced the NiTsPc redox

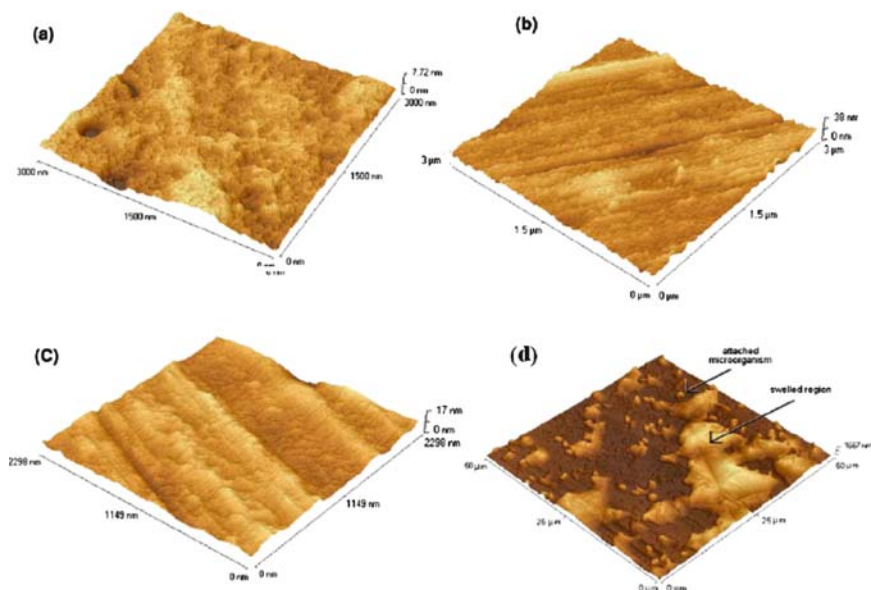


Fig. 26.9. AFM images of ten-layer films for (a) lysozyme, (b) trypsin, (c) pepsin, and (d) enzymatic (lysozyme) surface after interaction with microorganisms (*Escherichia coli*) in solution. The arrows indicate attachment of microorganisms (adapted from [63])

process and improved its catalytic activity for dopamine detection, thus leading to modified electrodes with high electroactivity, sensitivity, and selectivity. The growth and formation of PAMAM/NiTsPc and PAMAM-NT/NiTsPc LbL films were analyzed with AFM images.

Biosensors were produced by Morais et al. [63] with protease films, whose morphology, wettability, and bacteria attachment properties were studied. Figure 26.9 shows a set of AFM images for surfaces with ten deposited layers. For lysozyme films (Fig. 26.9a), regions can be observed with lack of deposition forming circular wells, probably owing to defects on the glass surface or unfavorable sites for interaction. For trypsin (Fig. 26.9b) and pepsin (Fig. 26.9c), the films seem to be compact and structurally ordered. Figure 26.9d shows the appearance of lysozyme film surface after contact with microorganisms. For the latter, two important points may be noted: swollen regions in the film and bacteria attached to the surface (identified with arrows on the lysozyme surface).

Perinotto et al. [64] used an electronic tongue, depicted in Fig. 26.10a, to detect ethanol in diluted aqueous solutions. The enzyme alcohol dehydrogenase (ADH) was immobilized in LbL films onto Au-interdigitated electrodes (IDEs), in conjunction with layers of PAMAM dendrimers (Fig. 26.10b). Brugnollo et al. used a similar electronic tongue (see Fig. 26.10a) with LbL films of POEA and poly(sodium-4-styrene sulfonate) to detect and distinguish four chemicals in the solution at 20 mM, including NaCl, HCl, and caffeine [65].

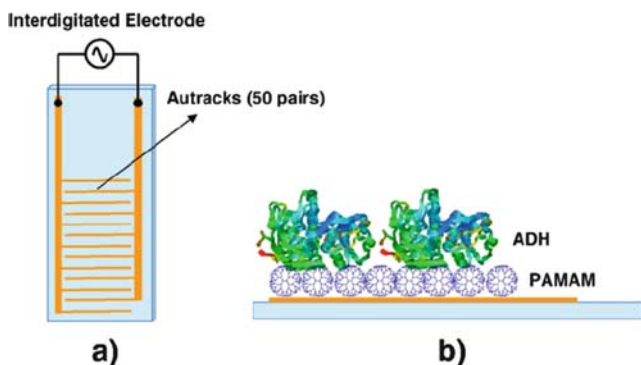


Fig. 26.10. Schematic view of (a) Au-interdigitated electrode (*IDE*) (*front view*) and (b) idealized architecture for a PAMAM/ADH film deposited onto an IDE (*side view*) (reprint from [64])

The substances were used to mimic the four basic tastes recognized by humans, namely sweet, salty, sour, and bitter. AFM images of the sensing layers showed that different acids yielded distinct film morphologies, with globules varying between 100 and 160 nm in diameter. Films containing greater amounts of POEA exhibited greater surface roughness.

Immobilized humic acid in smooth LbL films, as observed with AFM, were used to detect pesticides [66], where pentachlorophenol could be determined in solutions at concentrations as low as 10^{-9} mol/Liter. Consolin et al. [67] showed that the interaction with humic substances (HSs) induced doping of POEA via protonation, which was exploited in sensors made with POEA LbL films that could detect and distinguish HSs in aqueous solutions. The detection of HS is motivated by the need of controlling the quality of water and soil. Significantly, the interaction with HS caused the morphology of POEA films to be altered. The size of the globules typical of a POEA film [34] increased significantly when the film was immersed into buffer solutions containing fulvic and humic acids.

Similar results were obtained in the interaction between POEA adsorbed onto solid substrates and HS and Cu^{2+} ions [68] that doped POEA and changed film morphology. This interaction was exploited in a sensor array made with films of POEA, SL, and HS, which could detect small concentrations of HS and Cu^{2+} in water. Figure 26.11 shows AFM images of POEA films, as deposited and after being immersed for 20 min into solutions containing HS and Hs/Cu^{+2} ions and then dried.

Direct Measurement of Interaction Forces

In many applications, it is crucial to understand the interactions between the molecules involved, which is clearly demonstrated for the case of sensors and biosensors [69, 70], whose performance is governed by surface interactions

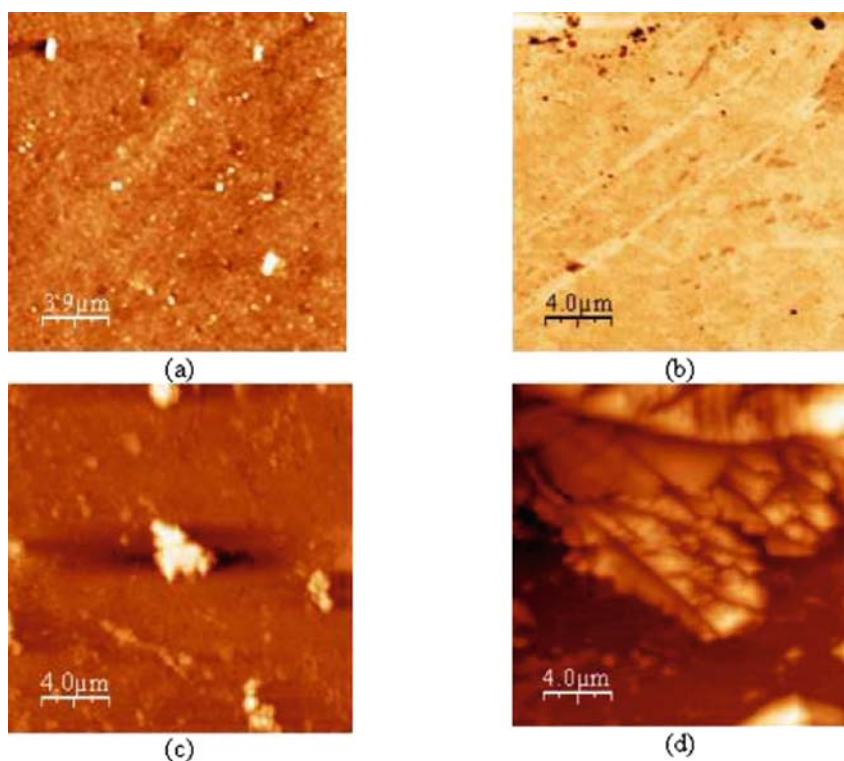


Fig. 26.11. AFM images for (a) as-deposited POEA films, (b) POEA films after being immersed into solutions containing HS, (c) HS/Cu²⁺, and (d) Cu²⁺ (reprint from [68])

[71]. In this context, AFM with the so-called AFS is perhaps the most important tool currently available, as direct force measurements can be done and interaction maps may be obtained.

Leite et al. [72] studied nonspecific interactions using AFS to measure interaction forces between an AFM tip and nanostructured films of POEA in pure water and CuSO₄ solutions, which mimic the experimental situation in taste sensors. The presence of impurities, for example, heavy metals, in the liquid cell, deliberately introduced or not, affected the force curves, possibly through adsorption and desorption of molecules. These changes could be captured in the force curves shown in Fig. 26.12 [73]. The measurements performed in pH 1 showed a larger double layer force than for pH 3. However, in pH 1, the interaction map is more inhomogeneous, with several repulsive regions, which is consistent with previous results [74]. The interaction forces measured with AFS, especially in liquid cells, may help understand the mechanism behind the high sensitivity of taste sensors, which has been attributed to

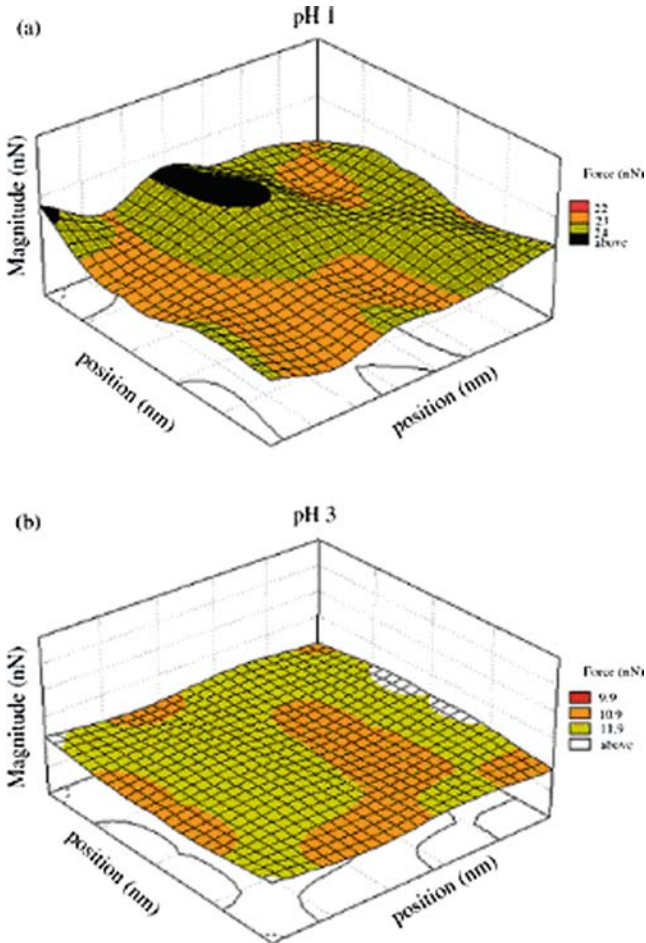


Fig. 26.12. Interaction maps for (a) POEA-ES (pH = 3) and (b) POEA-ES (pH = 1) (reprint from [73])

interfacial phenomena between the liquid under analysis and the film-forming molecules of the sensing unit.

That impurities in the liquid cell affect the force curves was also observed by Borato et al., [75], who employed AFS to investigate water ageing. As shown schematically in Fig. 26.13, the curve for fresh water displays a minimum with the distance between the tip and a flat mica surface that indicates predominance of attractive van der Waals interactions. For longer times, that is, after the water was left ageing in the cell, repulsive double-layer forces dominate until the force curve is practically purely repulsive (for t''). It seems that ageing of the water is accompanied by a change to lower pH values, which then increases the charge of the AFM tip, leading to an increase in the repulsive, double-layer force.

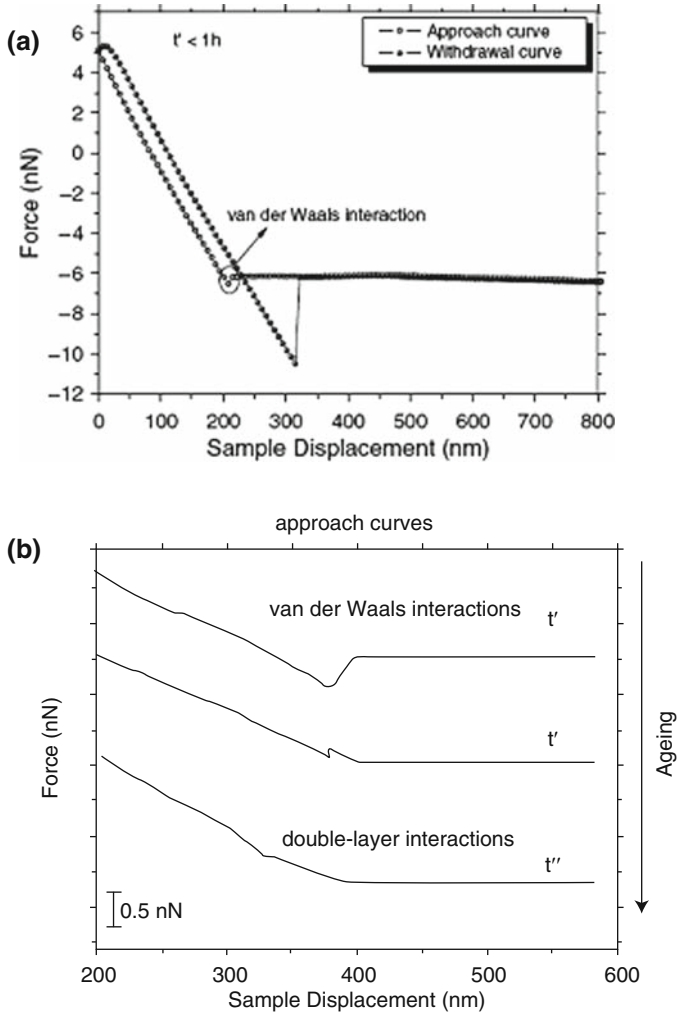


Fig. 26.13. Typical force curve for an AFM tip and mica sample immersed in water and (b) schema (approach curves) showing water ageing for various periods (increasing in the direction $t < t' < t''$). The difference between each of two times is 2 h (reprint from [75])

Force curves have also been used to probe theoretical models to explain charge transport in conducting polymers [74]. The existence of conducting islands was confirmed in force curve measurements, with the islands being denoted by regions with repulsive forces due to the double-layer forces. Figure 26.14a shows an adhesion map with islands of repulsive forces with diameter varying from 100 to 470 Å (average size = 306 ± 109 Å) in a matrix with attractive forces. These islands should be made of protonated and

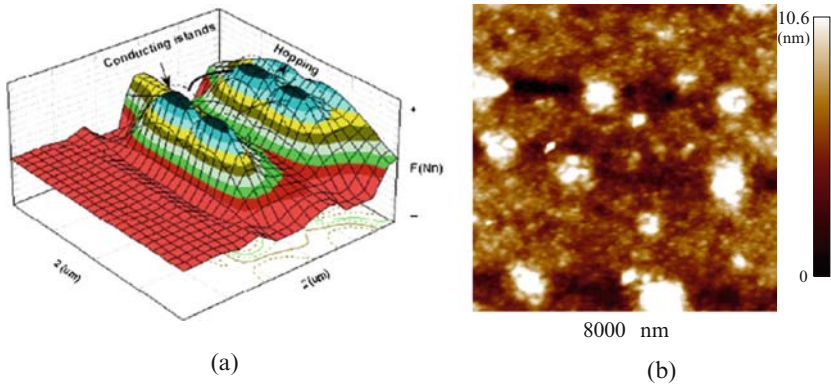


Fig. 26.14. (a) Map of forces obtained with atomic force spectroscopy showing regions of repulsive (conducting islands) and attractive interactions on POEA films in solution (pH = 3) (reprint from [74]). (b) AFM images of conducting islands in PANI films

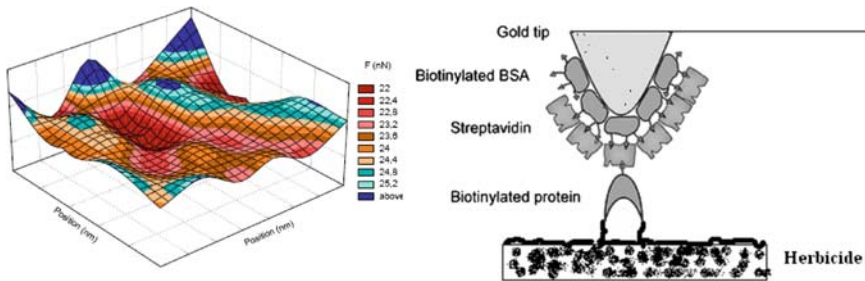


Fig. 26.15. (a) Adhesion map illustrating the variability of adhesion forces onto a POEA surface in contact with a pesticide (reprint from [77]) and (b) scheme for chemical modification of tips and sample substrates (adapted from [78])

crystalline POEA. These findings can provide insights into the mechanisms responsible for conduction in thin films and detection of trace amounts of analytes in a liquid, for example, pollutants in waters. Figure 26.14b displays results with conducting islands embedded in an amorphous matrix.

The surfaces of POEA LbL films had their roughness increased on interaction with heterocyclic herbicides (atrazine, imazaquin, and metribuzin), whereas the roughness decreased with interaction with paraquat [76]. Such interactions were further examined using force curve measurements and adhesion maps. Figure 26.15a shows an adhesion map for a POEA film in contact with atrazine for 2 h, where regions with larger forces were probably those most contaminated by atrazine. This was inferred from the fact that the tip was functionalized with the protein bovine serum albumin, thus possessing a stronger, specific interaction with atrazine. In fact, the use of functionalized

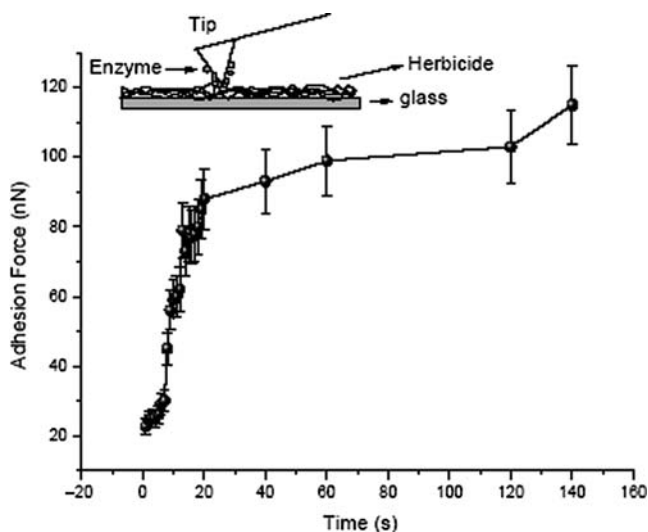


Fig. 26.16. Adhesion force versus adsorption times (self-assembly)

tips has led to the so-called chemical force microscopy (CFM), in which the force of interaction can be estimated from the excess force required to pull the tip free from the surface, as represented in Fig. 26.15b.

The concept of CFM is used to develop nanobiosensors, for example, to detect herbicides and other environmental pollutants. Figure 26.16 shows the time evolution of the adhesion force as adsorption takes place of acetyl-Coa carboxylase (ACCase) onto the surface of a tip AFM. The receptors (ACCase) were covalently anchored to the cantilever (tip surface functionalization). Enzyme inhibitors bind to enzymes and decrease their activity. Since blocking an enzyme activity can kill a pathogen or correct a metabolic imbalance, many drugs are enzyme inhibitors. With CFM, one could distinguish between nonspecific adhesion and specific interactions – brought about by the herbicide – as shown in Fig. 26.16. This force curve was acquired with a thin film of diclofop-methyl in contact with ACCase for 5 h. These results are in accordance with those obtained with an electronic tongue as the sensing device.

Natural Fibers and Soil Science

Plant fibers have been employed in a wide range of manufacturing applications, including textiles, paper, automobile, and chemical industries, and as reinforcement in polymer-based composites to create materials with improved performance [79]. Sisal fibers are among the most widely used natural fibers, as in reinforcing composite materials with polypropylene [80] and polyethylene [81]. Their properties vary strongly with the chemical treatment, and this has

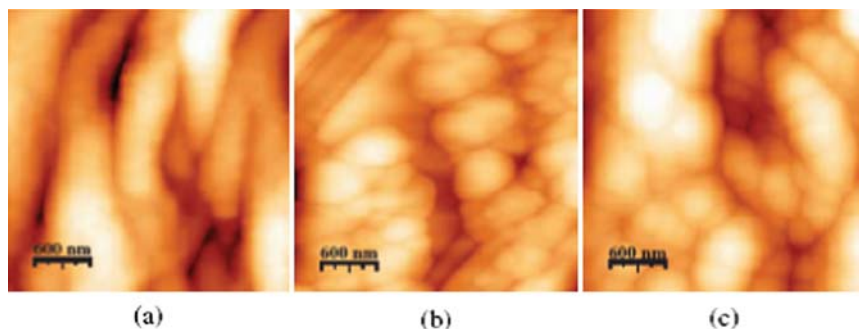


Fig. 26.17. 2D AFM images (scan size: $3\ \mu\text{m}$). (a) Treated with 40% NaOH; (b) benzylated for 0.5 h, and (c) for 2 h (reprint from [82])

been investigated through changes in morphology and adhesion [82]. Indeed, Fig. 26.17 shows that treatment with NaOH and via benzylations reactions caused the fiber morphology to change significantly. The fibrillar structure in the sample treated with 40% NaOH, in Fig. 26.17a, changes to a globular morphology after incorporation of benzyl groups for 0.5 h (Fig. 26.17b). With 2 h of benzylation, only the globular morphology was observed in Fig. 26.17c. Furthermore, the adhesion force between the sisal fiber and the AFM tip increased significantly with benzylation since the tip has a hydrophobic character, which demonstrated the importance of benzylation to decrease the polar component of the fibers. Hence, force curve measurements are promising to evaluate the surface energy of natural fibers, which is the main parameter governing the wettability and adhesion properties.

AFM has been used to study the morphology of cross-linked thermosets used as matrices in high-performance, low-cost composites. Mosiewicki et al. [83] analyzed the relationship between the morphology of styrene (St)/vinylester and St/unsaturated polyester resins and the St content. At low St concentration or high nanogel content, the nanogels are closely packed together, so that no individual nanoparticles are observed. At high St concentrations, the nanogels are connected by longer St–St chains, and the reduced coalescence leads to a microstructure formed by smaller, further separated particles.

Another AFM technique used for characterizing natural fibers is the dynamic AFS or “damping oscillation” [84], which consists of a vibrating microcantilever after pull-off force, with a nanoscale tip that interacts with a sample surface via short- and long-range intermolecular forces. Figure 26.18 shows the spectrum with amplitude and frequency of the damped oscillations, whose features may be correlated with the physicochemical properties of sisal fibers. The adhesion force and the surface energy cause the oscillations amplitude, the frequency delay, and curve profile to change significantly. This method may be useful for assessing the state of the cantilever and supply information on the environment where the experiment is made, for example, viscosity values.

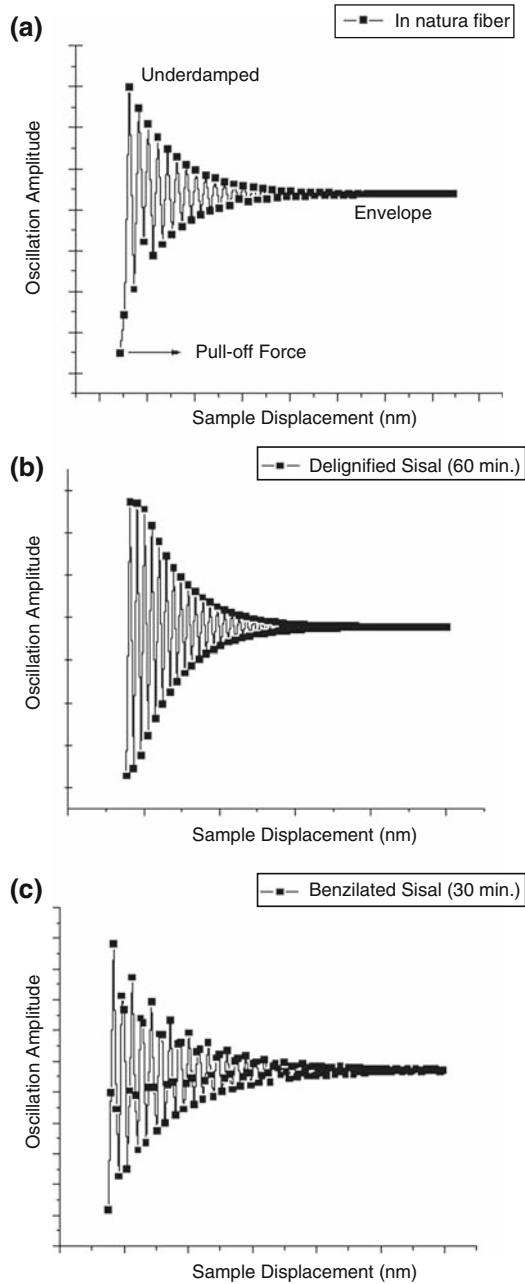


Fig. 26.18. Damping oscillation profile of sisal fiber (a) in natura, (b) delignified (60 min), and (c) benzilated (30 min) (reprint from [84])

Cellulose nanofibrils [85] in LbL films with POEA were studied by Medeiros et al. [86], who observed a globular morphology for plain POEA films and a typical fibrillar morphology for the films containing cellulose nanofibrils, as shown in Fig. 26.19. The fibers were found to have an average diameter (d) of 38 ± 13 nm, length (l) of 360 ± 60 nm, and aspect ratio (l/d) of 10 ± 4 , consistent with the literature [87].

The particle size distribution in soils may be determinant for many soil properties, and this may be investigated with AFM. For an oxisol soil from Brazil, Vaz et al. [88] obtained both particle size and thickness distribution with 2D images, such as those shown in Fig. 26.20.

Leite et al. [89] used AFM to measure the adhesion force between an AFM tip and mica and quartz surfaces, in air and water. The adhesion (pull-off force) depended on both substrate morphology and the environment, and the adhesion maps could be used to identify regions contaminated with organic compounds, as indicated in Fig. 26.21. The force curves obtained on the soil mineral particles (mica and quartz) showed clearly that the adhesion force is sensitive to both the surface roughness and the environment conditions (*see* Fig. 26.21b and c). The adhesion maps made from hundreds of measurements at different points allowed to obtain information on heterogeneities in sample topography and on organic contamination. A small long-range repulsive force

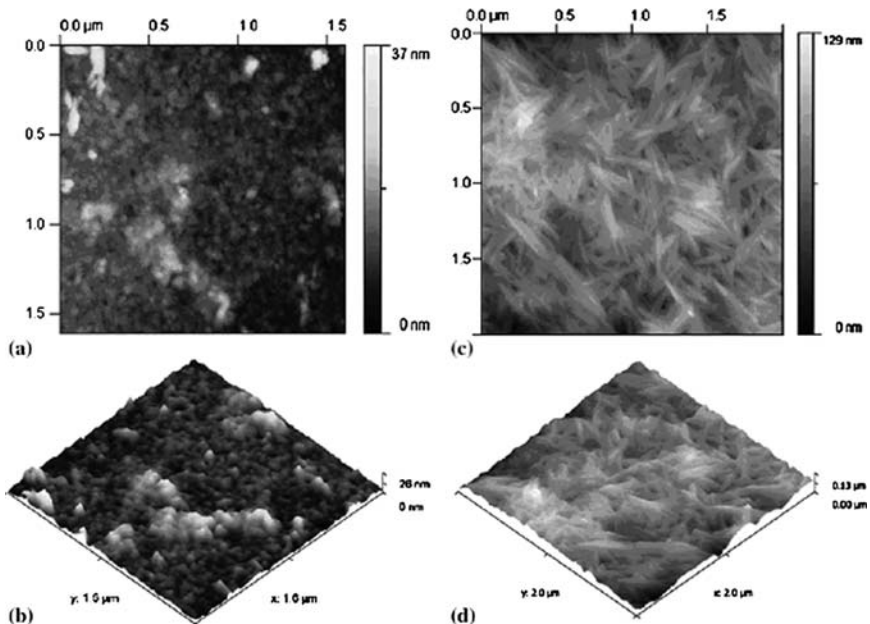


Fig. 26.19. 2D and 3D images of self-assembled thin films showing the top layer (tenth immersion of POEA or tenth bilayer of POEA-CnF) of POEA (a,b) and POEA-CnF (c,d) (reprint from [86])

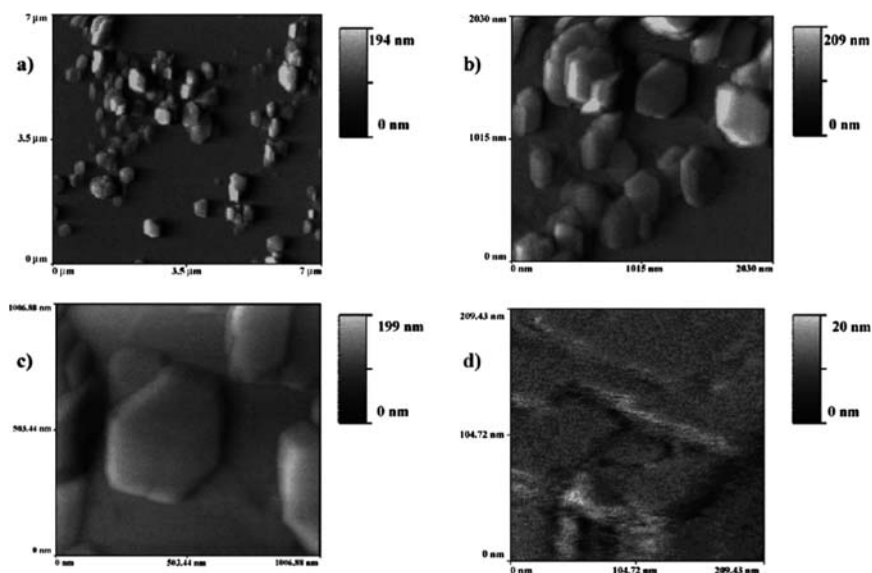


Fig. 26.20. Sequence of AFM images of kaolinite particles (nr 4 Oneal Pit) deposited on a mica substrate, in four different scan sizes (reprint from [88])

is also observed in water due to the overlapping electrical double layers formed on both the tip and sample surfaces.

The influence from the thermal treatment on the morphology of nanostructured hydroxyapatite was studied with AFM [90], where the particle size increased with the heating temperature. Aggregates were observed on the surface, as illustrated in Fig. 26.22, regardless of the type of substrate and deposition conditions. Some aggregates were less than 2,300 nm in size for samples heated at 1,150°C, which also appeared in Brazilian Latosols (Oxisols) samples [91].

Other Applications

The tip of an AFM may be used to produce changes in the sample, for instance, in processes akin to nanolithography [92] with a mechanical “writing” or by applying a voltage to the tip. In addition, furrows may be created on the sample by pressing the tip in a controlled way that was used for the direct measurement of ultrathin polymer films from the height profile provided in the AFM tapping mode [93]. For scrapping the polymer surface, AFM is used in the contact mode. Significantly, the thickness of LbL films POMA/PVS adsorbed onto glass and silicon substrates [92] was comparable to those obtained with profilometry. Figure 26.23 illustrates the furrows produced with an AFM tip on a POMA/PVS LbL film on hydrophilic glass and silicon oxide [30].

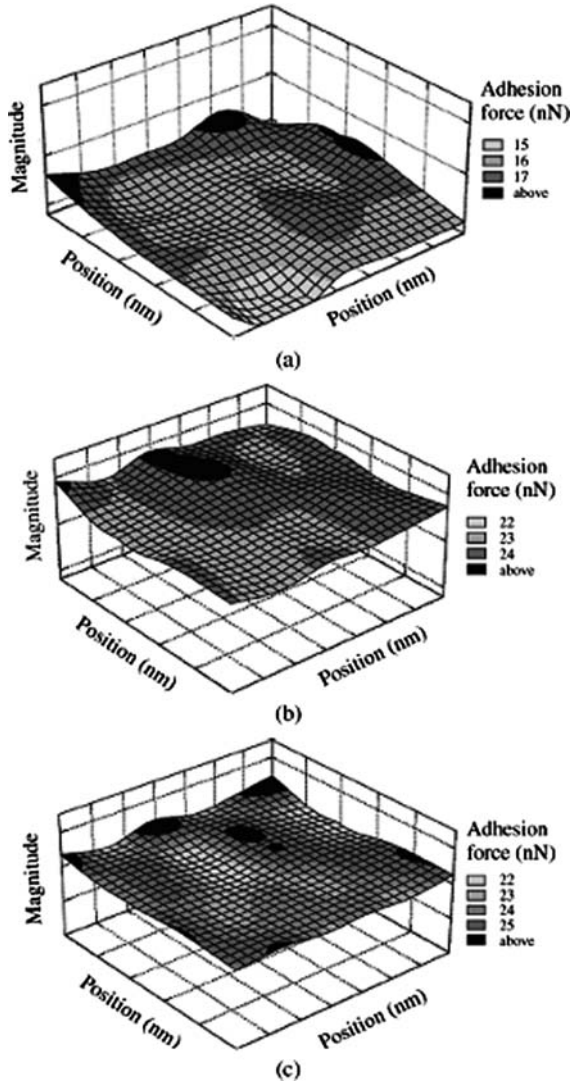


Fig. 26.21. Adhesion maps onto mica: (a) region A, (b) region B, and (c) region C. Each adhesion map corresponds to a scanning area of $1\mu\text{m}^2$. The black regions correspond to values up to 30% above the average adhesion (reprint from [89]).

Magnetic properties are also possible to study with a surface probe microscopy method [94,95], for which the tip is coated with a magnetic material. The magnetic force between tip and sample induces a change in the resonant frequency of the oscillating tip-cantilever system, which is captured with images over the sample. This technique is promising for investigations of soils with ferromagnetic properties. Indeed, Fig. 26.24 shows the topographical

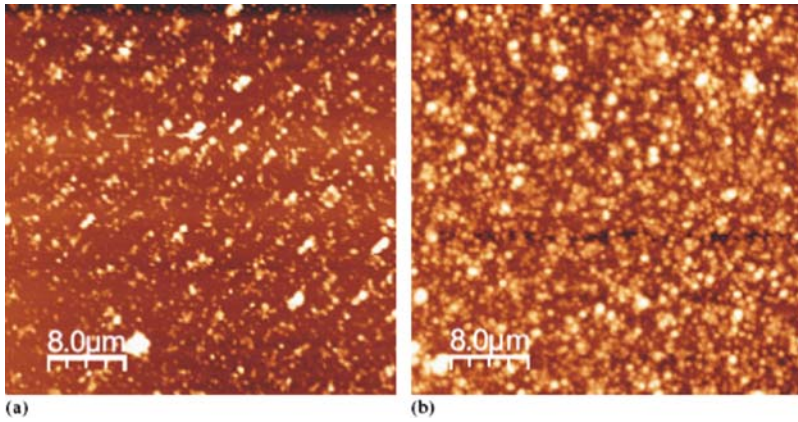


Fig. 26.22. AFM images of Oxisol Clays (Hapludox). Parent rock: (a) sandstone and (b) shist (reprint from [91])

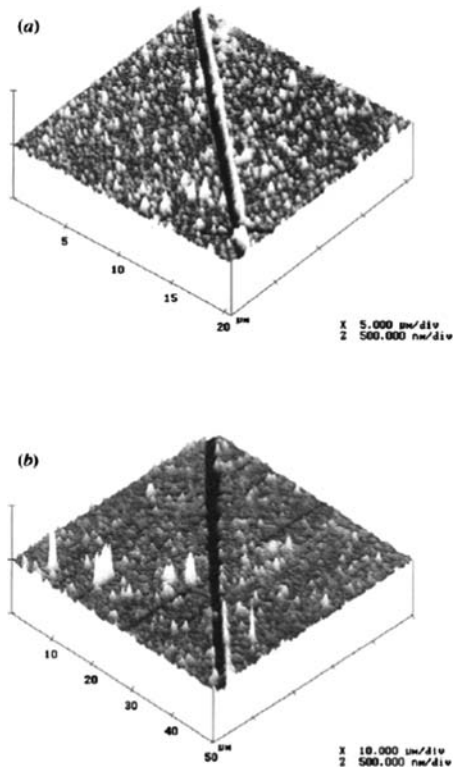


Fig. 26.23. AFM images of two furrows made on: (a) 10-bilayer POMA/PVS film on glass and (b) 14-bilayer POMA/PVS film on silicon (reprint from [93])

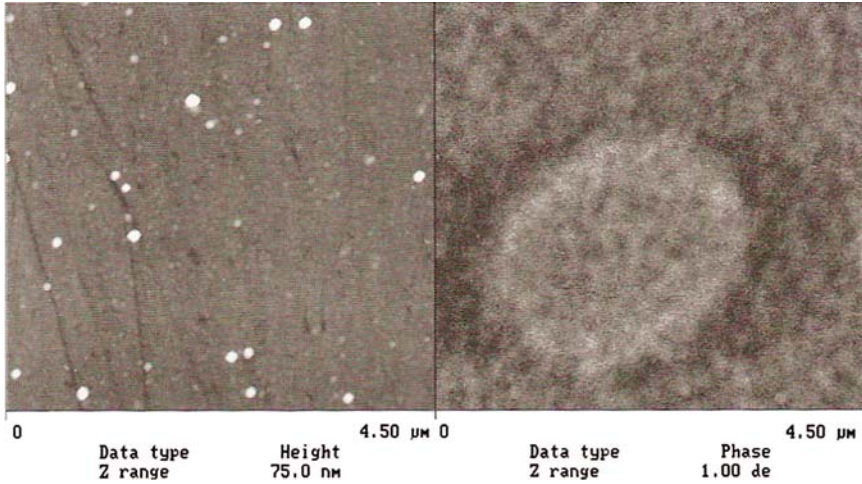


Fig. 26.24. Topography (*left*) and MFM (*right*) images from a standard sample used to calibrate the detectors. The specimen is made from glass as the basis substrate, a first layer of platinum (*Pt*) (1.2 nm thick) and above it *N* layers of *Pt* and $\text{Co}_{50}\text{Ni}_{50}$ (reprinted from [96])

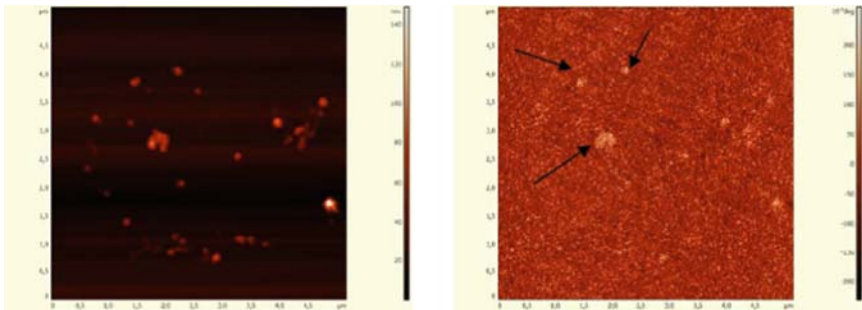


Fig. 26.25. AFM image (semicontact mode regime) (*left*) and MFM image (second pass regime) (*right*) of a mixture of maghemite-hematite nanoparticles. (reprinted from [97])

image obtained by AFM and the phase image obtained by magnetic force microscopy (MFM), displaying a typical pattern of a magnetic domain.

MFM has been also exploited in nanomagnetic studies to clarify the dependence of magnetic properties on structural features of maghemite-hematite nanoparticles [97] (see Fig. 26.25). With regard to biological systems, MFM has been used in investigating interactions between streptavidin and magnetotactic bacterium [98, 99].

26.3 Conclusions and Perspectives

The few examples presented in this chapter may have served to illustrate the potential of AFM and related methods to agriculture and agribusiness. When combined with other nanotech methods, AFM may be useful for many further applications, as in characterizing bioactive and/or engineered thin films and in designing sensors and biosensors for monitoring food quality and impacts to the environment. In this context, the advent of CFM has brought new perspectives, with the unprecedented measurements of intermolecular forces. Indeed, it is now possible to functionalize the AFM tip that will interact with the material of interest on the sample.

Also important has been the introduction of novel concepts and systems through nanotechnology, such as monopropulsion, nanomeasurement, nanosensing, nanoprobes, and nanoservodevices. It is predicted that within 5–15 years, many practical results of nanotechnology and nanosciences will be seen in agriculture and agribusiness, which will only be possible with strengthened partnerships among universities, research institutes, and industries [100]. This requires large investments in human resources, with multidisciplinary and interinstitutional character, in addition to financial resources for infrastructure and equipment.

References

1. P. Liu, Y.W. Zhang, C. Lu, *Phys. Rev. B* **68**, 195314 (2003).
2. M.S.B. Brandão, *Self-Assembly and Self-Organization of Surfactants in Agriculture*, vol. 4 ed. by A.T. Hubbard. *Encyclopedia of Surface and Colloid Science*, 1st edn. (CRC, 2002).
3. L.J. Lauhon, M.S. Gudiksen, D. Wang, C.M. Lieber, *Nature* **420**, 57–61 (2002).
4. A.M. Morales, C.M. Lieber, *Science* **279**, 208–211 (1998).
5. T. Ohgi, H.Y. Sheng, H. Nejo, *App. Surf. Sci.* **130–132**, 919–924 (1998).
6. I. Sokolov, Y.Y. Kievsky, J.M. Kaszpurenko, *Small* **3**, 419–423 (2007).
7. C. Moraru, C. Panchapakesan, Q. Huang, P. Takhistov, S. Liu, J. Kokini, *Nanotechnology: A New Frontier in Food Science*, vol. 57, number 12 (Institute of Food Technologists, December 2003).
8. <http://www.tahan.com/charlie/nanosociety/course201/>
9. R.C. Barry, Y.H. Lin, J. Wang, G.D. Liu, C.A. Timchalk, *J. Exp. Sci. Environ. Epidemiol.* **19**, 1–18 (2009).
10. S.Q. Liu, L. Yuan, X.L. Yue, Z.Z. Zheng, Z.Y. Tang, *Adv. Powder Technol.* **19**, 419–441 (2008).
11. S. Stolnik, K. Shakesheff, *Biotechnol. Lett.* **31**, 1–11 (2009).
12. E. Sivamani, R.K. DeLong, R.D. Qu, *Plant Cell Rep.* **28**, 213–221 (2009).
13. A. Kulamarva, P.M.V. Raja, J. Bhatena, H. Chen, S. Talapatra, P.M. Ajayan, O. Nalamasu, S. Prakash, *Nanotechnology* **20**, 025612 (2009).
14. X.F. Xie, B. Goodell, D.J. Zhang, D.C. Nagle, Y.H. Qian, M.L. Peterson, J. Jellison, *Bioresour. Tech.* **100**, 1797–1802 (2009).

15. A. Otten, S. Herminghaus, *Langmuir* **20**, 2405–2408 (2004).
16. Y. Shimizu, T. Sasaki, T. Kodaira, K. Kawaguchi, K. Terashima, N. Koshikazi, *Chem. Phys. Lett.* **370**, 774–780 (2003).
17. S. Chan, X. Su, M. Yamakawa Detecting, identifying and sequencing of biomolecules using controlled alignment of nano-barcodes encoding specific information for scanning probe microscopy, useful in the fields of molecular biology. Patent Number(s): WO2004038037-A2; US2004126820-A1; AU2003278852-A1; EP1546983-A2; JP2006501485-W; CN1682237-A; AU2003278852-A8; US2006281119-A1; US7476786-B2.
18. S.R. Nicewarner-Pena, R.G. Freeman, B.D. Reiss, L. He, D.J. Pena, I.D. Walton, R. Cromer, C.D. Keating, M.J. Natan, *Science* **294**, 137–141 (2001).
19. E.S. Medeiros, L.H.C. Mattoso, E.N. Ito, K.S. Gregorski, G.H. Robertson, R.D. Offeman, D.F. Wood, W.J. Orts, S.H. Imam, *J. Biobased Mat. Bionenerg.* **2**, 231–242 (2008).
20. F.L. Leite, P.S.P. Herrmann, *J. Adhes. Sci. Technol.* **19**, 365–405 (2005).
21. G. Binnig, H. Rohrer, Ch. Gerber, E. Weibel, *Appl. Phys. Lett.* **40**, 178–180 (1982).
22. G. Binnig, C.F. Quate, C. Gerber, *Phys. Rev. Lett.* **56**, 930–933 (1986).
23. S.Y. He, G.P. Feng, H.S. Yang, Y. Wu, Y.F. Li, *Postharvest Biol. Technol.* **33**, 263–273 (2004).
24. E.A. Veraverbeke, P. Verboven, P.V. Oostveldt, B.M. Nicolai, *Postharvest Biol. Technol.* **30**, 75–88 (2003).
25. H. Yang, Y. Wang, S. Lai, H. An, Y. Li, F. Chen, *J Food Sci* **72**, R65–R75 (2007).
26. W.A. Ducker, T.J. Senden, R.M. Pashley, *Nature* **353**, 239–241 (1991).
27. O.B.G. Assis, L.A. Forato, de D. Britto, *Higiene Alimentar* **22**, 99–106 (2008).
28. G.G. Roberts, *Langmuir-Blodgett Films* (Plenum, New York, 1990).
29. G. Decher, J.D. Hong, J. Schmitt, *Thin Solid Films* **210**, 831–835 (1992).
30. R.F.M. Lobo, M.A. Pereira-da-Silva, M. Raposo, R.M. Faria, O.N. Oliveira Jr, *Nanotechnology* **14**, 101–108 (2003).
31. L.G. Paterno, L.H.C. Mattoso, *Polymer* **42**, 5239–5245 (2001).
32. A. Riul Jr, A. Dhanabalan, M.A. Cotta, P.S.P. Herrmann, L.H.C. Mattoso, A.G. MacDiarmid, O.N. Oliveira Jr, *Synth. Met.* **101**, 830–831 (1999).
33. D. Volpati, A.E. Job, R.F. Aroca, C.J.L. Constantino, *J. Phys. Chem. B* **12**, 3894–3902 (2008).
34. F.L. Leite, L.G. Paterno, C.E. Borato, P.S.P. Herrmann, O.N. Oliveira Jr, L.H.C. Mattoso, *Polymer* **46**, 12503–12510 (2005).
35. F.L. Leite, M. Oliveira Neto, L.G. Paterno, M.R.M. Ballesterio, I. Polikarpov, Y.P. Mascarenhas, P.S.P. Herrmann, L.H.C. Mattoso, O.N. Oliveira Jr, *J. Coll. Interface Sci.* **316**, 376–387 (2007).
36. E.C. Venancio, L.G. Paterno, C.E. Borato, A. Firmino, L.H.C. Mattoso, *J. Braz. Chem. Soc.* **16**, 558–564 (2005).
37. K. Xu, L.H. Zhu, H.Q. Tang, *Electrochim Acta* **52**, 723–727 (2006).
38. D.H. Zhang, Y.Y. Wang, *Mat. Sci. Eng. B* **134**, 9–19 (2006).
39. L.G. Paterno, L.H.C. Mattoso, *J. Appl. Polym. Sci.* **83**, 1309–1316 (2002).
40. K. Wohnrath, J.R. Garcia, F.C. Nart, A.A. Batista, O.N. Oliveira Jr, *Thin Solid Films* **402**, 272–279 (2002).
41. A. Marletta, C.A. Olivati, M. Ferreira, M.L. Veiga, D.T. Balogh, R.M. Faria, O.N. Oliveira Jr, *Braz. J. Phys.* **36**, 496–498 (2006).

42. L. Gaffo, C.J.L. Constantino, W.C. Moreira, R.F. Aroca, O.N. Oliveira Jr, *Langmuir* **18**, 3561–3566 (2002).
43. K. Wonhrath, S.V. Mello, M.A. Pereira-da-Silva, O.N. Oliveira Jr, *Synth. Met.* **121**, 1425–1426 (2001).
44. D. Pasquini, D.T. Balogh, P.A. Antunes, C.J.L. Constantino, A.A.S. Curvelo, R.F. Aroca, O.N. Oliveira Jr, *Langmuir* **18**, 6593–6596 (2002).
45. D. Pasquini, D.T. Balogh, O.N. Oliveira Jr, A.A.S. Curvelo, *Coll. Surface A: Physicochem. Eng. Aspec.* **252**, 193–200 (2005).
46. O.B.G. Assis, D.C. Vieira, R. Bernardes-Filho, *Braz. J. Chem. Engineer.* **2000** **17**, 245–249 (2000).
47. C.R. Alves, M.G.R. Pimenta, R.H.S.F. Vieira, R.F. Furtado, M.I.F. Guedes, R.C.B. Silva, O.B.G. Assis, *Elect. J. Biotechnol.* **10**, 160–165 (2007).
48. A. Navrotsky, *J. Nanopart. Res.* **2**, 321–323 (2000).
49. D.W.O. Medeiros, Trindade C.G. Neto, D.E.S. Santos, F.J. Pavinatto, D.S. Santos, O.N. Oliveira Jr, A.E. Job, J.Á. Giacometti, T.N.C. Dantas, M.R. Pereira, J.L.C. Fonseca, *J. Dispers. Sci. Technol.* **26**, 267–273 (2005).
50. P.J.G. Goulet, D.S. dos Santos Jr, R.A. Alvarez-Puebla, O.N. Oliveira Jr, R.F. Aroca, *Langmuir* **21**, 5576–5581 (2005).
51. Y.C. Cao, R. Jin, C.A. Mirkin, *Science* **297**, 1536–1540 (2002).
52. J. Stejskal, I. Sapurina, M. Trchova, E.N. Konyushenko, P. Holler, *Polymer* **47**, 8253–8262 (2006).
53. L.G. Paterno, F.J. Fonseca, G.B. Alcantara, M.A.G. Soler, P.C. Morais, J.P. Sinnecker, M.A. Novak, E.C.D. Lima, F.L. Leite, L.H.C. Mattoso, *Thin Sol. Films* **517**, 1753–1758 (2009).
54. M.G. Xavier, E.C. Venancio, E.C. Pereira, F.L. Leite, E.R. Leite, A.G. MacDiarmid, L.H.C. Mattoso, *J. Nanosci. Nanotechnol.* **8**, 2169–2172 (2009).
55. P.A. Antunes, C.M. Santana, R.F. Aroca, O.N. Oliveira Jr, C.J.L. Constantino, Riul A. Jr, *Synth. Met.* **148**, 21–24 (2005).
56. L. Qi, Z. Xu, X.Q. Jiang, X. Hu, *Carbohydr. Res.* **339**, 2693–2700 (2004).
57. R.A.A. Muzzarelli, *Cell. Mol. Biol. Life Sci.* **53**, 131–140 (1997).
58. C.A. Constantine, S.V. Mello, A. Dupont, X. Cao, dos Santos D.S. Jr, O.N. Oliveira Jr, F.T. Strixino, E.C. Pereira, T. Cheng, J.J. Defrank, R.M. Leblanc, *J. Am. Chem. Soc.* **125**, 1805–1809 (2003).
59. K. Kurita, Y. Koyama, A. Tanaguchi, *J. Appl. Polym. Sci.* **31**, 1169–1176 (1986).
60. M.N.V.R. Kumar, *React. Funct. Polym.* **46**, 1–27 (2000).
61. C.E. Borato, F.L. Leite, L.H.C. Mattoso, R.C. Goy, S.P. Campana Filho, C.L. de Vasconcelos, C.G. da Trindade Neto, M.R. Pereira, J.L.C. Fonseca, O.N. Oliveira Jr, *IEEE Trans. Dielec. Electrical. Ins.* **13**, 1101–1109 (2006).
62. J.R. Siqueira, L.H.S. Gasparatto, O.N. Oliveira Jr, V. Zucolotto, *J. Phys. Chem. C* **112**, 9050–9055 (2008).
63. L.C. de Morais, R. Bernardes-Filho, O.B.G. Assis, *World J. Microbiol. Biotechnol.* **25**, 123–129 (2009).
64. A.C. Perinotto, L. Caseli, C.O. Hayasaka, A. Riul Jr, O.N. Oliveira Jr, V. Zucolotto, *Thin Sol. Films* **516**, 9002–9005 (2008).
65. E.D. Brugnollo, L.G. Paterno, F.L. Leite, F.J. Fonseca, C.J.L. Constantino, P.A. Antunes, L.H.C. Mattoso, *Thin Sol. Films* **516**, 3274–3281 (2008).
66. F.N. Crespilho, V. Zucolotto, J.R. Siqueira Jr, C.J.L. Constantino, F.C. Nart, O.N. Oliveira Jr, *Environ. Sci. Technol.* **39**, 5385–5389 (2005).

67. E.C. Venancio, N. Consolin Filho, C.J.L. Constantino, L. Martin-Neto, L.H.C. Mattoso, *J. Braz. Chem. Soc.* **16**, 24–30 (2005).
68. F.L. Leite, A. Firmino, C.E. Borato, O.N. Oliveira Jr, L.H.C. Mattoso, W.T.L. da Silva, *Synth. Met.* (2008), in press. DOI: 10.1016/j.synthmet.2009.07.058
69. M. Ferreira, A. Riul Jr, K. Wohnrath, F.J. Fonseca, O.N. Oliveira Jr, L.H.C. Mattoso, *Anal. Chem.* **75**, 953–955 (2003).
70. E.-L. Florin, M. Rief, H. Lehmann, M. Ludwig, C. Dornmair, V.T. Moy, H.E. Gaub, *Biosen. Bioelectron.* **10**, 895–901 (1995).
71. A. Riul Jr, D.S. dos Santos Jr, K. Wohnrath, R. Di Tomazo, A.A.C.P.L.E. Carvalho, F.J. Fonseca, O.N. Oliveira Jr, D.M. Taylor, L.H.C. Mattoso, *Langmuir* **18**, 239–245 (2002).
72. F.L. Leite, C.E. Borato, W.T.L. da Silva, P.S.P. Herrmann, O.N. Oliveira Jr, L.H.C. Mattoso, *Microscopy Microanal.* **13**, 304–312 (2007).
73. F.L. Leite, W.F. Alves, M. Oliveira Neto, I. Polikarpov, P.S.P. Herrmann, L.H.C. Mattoso, O.N. Oliveira Jr, *Micron* **39**, 1119–1125 (2008).
74. F.L. Leite, W.F. Alves, M. Mir, Y.P. Mascarenhas, P.S.P. Herrmann, L.H.C. Mattoso, O.N. Oliveira Jr, *Appl. Phys. A* **93**, 537–542 (2008).
75. C.E. Borato, F.L. Leite, O.N. Oliveira Jr, L.H.C. Mattoso, *Sensor Lett.* **4**, 155–159 (2006).
76. N. Consolin Filho, F.L. Leite, E.R. Carvalho, E.C. Venâncio, C.M.P. Vaz, L.H.C. Mattoso, *J. Braz. Chem. Soc.* **18**, 577–584 (2007).
77. F.L. Leite, L.H.C. Mattoso, O.N. Oliveira, P.S.P. Herrmann, *The Atomic Force Spectroscopy as a Tool to Investigate Surface Forces: Basic Principles and Applications*, ed. by A. Méndez-Vilas and J. Diaz, *Modern Research and Educational Topics in Microscopy* (Formatex, 2007), 747–757.
78. P. Hinterdorfer, Y.F. Dufrière, O.N. Oliveira Jr, *Nat. Meth.* **3**, 347–355 (2006).
79. L.Y. Mwaikambo, M.P. Ansell, *J. Appl. Polym. Sci.* **84**, 2222–2234 (2002).
80. M.A. Martins, L.H.C. Mattoso, *J. Appl. Polym. Sci.* **91**, 670–677 (2004).
81. K. Joseph, S. Thomas, C. Pavithran, *Polymer* **37**, 5139–5149 (1996).
82. F.L. Leite, P.S.P. Herrmann, A.L. Da Róz, F.C. Ferreira, A.A.S. Curvelo, L.H.C. Mattoso, *J. Nanosci. Nanotechnol.* **6**, 2354–2361 (2006).
83. M.A. Mosiewicki, W.F. Schroeder, F.L. Leite, P.S.P. Herrmann, A.A.S. Curvelo, M.I. Aranguren, J. Borrajo, *J. Mater. Sci.* **41**, 6154–6158 (2006).
84. F.L. Leite, E.C. Ziemath, P.S.P. Herrmann, *Scanning* **30**, 271 (2008).
85. N.E. Marcovich, M.L. Auid, N.E. Bellesi, S.R. Nutt, M.I. Aranguren *J. Mater. Res.* **21**, 870–881 (2006).
86. E.S. Medeiros, L.H.C. Mattoso, R. Bernardes-Filho, D.F. Wood, W.J. Orts, *Coll. Polym. Sci.* **286**, 1265–1272 (2008).
87. N.L.G. Rodriguez, W. Thielemans, A. Dufresne, *Cellulose*, **13**, 261–270 (2006).
88. C.M.P. Vaz, P.S.P. Herrmann, S. Crestana, *Powd. Technol.* **126**, 51–58 (2002).
89. F.L. Leite, Riul A. Jr, P.S.P. Herrmann, *J. Adhes. Sci. Technol.* **17**, 2141 (2003).
90. F.L. Leite, M. Mir, A.M. Rossi, E.L. Moreira, Y.P. Mascarenhas, P.S.P. Herrmann, *Scanning* **30**, 265 (2008).
91. F.L. Leite, P.S.P. Herrmann, Y.P. Mascarenhas, M.E. Alves, *Scanning* **30**, 275–276 (2008).
92. R.M. Nyffenegger, R.M. Penner, *Chem. Rev.* **97**, 1195–1230 (1997).
93. R.F.M. Lobo, M.A. Pereira-da-Silva, M. Raposo, R.M. Faria, O.N. Oliveira Jr, *Nanotechnology* **10**, 389–393 (1999).

94. Y. Martin, H.K. Wickramasinghe, *Appl. Phys. Lett.* **50**, 1455 (1987).
95. J.J. Saenz, N. Garcia, P. Grutter, E. Meyer, H. Heinzelmann, R. Wiesendanger, L. Rosenthaler, H.R. Hidber, H.-J. Guntherodt, *J. Appl. Phys.* **62**, 4293–4295 (1987).
96. P.S.P. Herrmann, L.A. Colnago, L.H.C. Mattoso, P.E. Cruvinel, J. Frommer, *Nanotechnology and Nanosciences, Perspectives and Potential Applications in Agricultural Research* (chapter 3), ed. by P.E. Cruvinel, S. Mascarenhas, *Advanced Studies in Agricultural Instrumentation* (Rima Editora, Brazil, 2002).
97. A. Hendrych, R. Kubinek, A.V. Zhukov, *The Magnetic Force Microscopy and its Capability for Nanomagnetic Studies-The Short Compendium*, ed. by A. Méndez, J. Diaz, *Modern Research and Educational Topics in Microscopy* (©Formatex, 2007), 805–811.
98. R.B. Proksch, B.M. Moskowitz, E.D. Dahlberg, T. Schaeffer, *Appl. Phys. Lett.* **66**, 2582–2584 (1995).
99. Y. Amemiya, T. Tanaka, B. Yozaand, T. Matsunaga, *Biotechnology* **120**, 308–314 (2005).
100. M. Gross, *Travel to Nanoworld, Miniature, Machinery in Nature and Technology* (Plenum Trade, New York, 1999), 1–254.

Index

- Abrasion, 648, 658, 670, 672, 674, 675
- Added mass method, 296, 304, 317
- Adhesion, 216
 - force, 13, 14, 263, 278, 280, 663–664
 - hysteresis, 241, 242
 - wear, 648, 651, 652, 666, 670–671, 674, 680
- Adhesive contact, 240, 241
- Admittance A^* , 628
- Affinity, 544
- AFM applications, 724, 727
 - biological macromolecules, 725, 726
 - nanometer scale imaging, 724, 725
 - technical advancement, 726, 727
- AFM with force transducer (AFM-FT), 202, 219
- Agriculture, 915, 917, 920, 923, 924, 940
- Alumina, 218
- Amontons' law, 263, 268, 283, 318
- Amplitude-modulation (AM) mode, 11, 12
- Analyzer azimuth, 77
- Angle-resolved photoemission spectroscopy (ARPES), 519–522, 525, 540, 561
- Antibody-antigen, 327, 329, 346
- Antiferromagnetic, 517, 518, 525, 526, 537, 548, 551
- Antiferromagnetism, 511
- Apex atoms, 559–561
- Apical oxygen, 526, 542, 559
- Archaeal histones, 742
- Archard formula, 649–650, 663, 665
- Area function, 210
- Asymptotic homogenisation theory, 460
- Atomic force microscopy (AFM), 3, 129, 182, 273, 274, 276, 280, 284, 285, 288, 290, 297, 300–302, 306, 318, 724
 - cantilever, 797
 - complex shear modulus (G), 369
 - displacement resolution, 402–404
 - displacement sensitivity, 402–404
 - epi-fluorescence and phase images, 806–808
 - force mode, 366, 368
 - force resolution, 404
 - force sensitivity, 404
 - force spectroscopy, 395–414
 - imaging, 366
 - instrumentation, 367
 - operating mode, 228–230
 - photodiode, 406
 - piezoscanner, 406
 - principle, 364
 - spring stiffness, 231–233, 240
 - topography image, 799–801, 808
- Atomic force spectroscopy (AFS), 916, 928, 933
- Atoms, 507
 - fermionic, 565
 - ultra-cold, 565
- Axial tensile loading, 574
- Back focal plane (BFP) interferometry, 29–31
- Background suppression, 75

- Bacterial nucleoids, 741
- Bandwidth, 402–404, 406, 414
- Barrier layer capacitance elements (BLCEs), 633
- BCS theory, 518, 551, 553, 558, 562, 564
- Bead-spring model, 173, 174
- Beam-bending, 576
- Beam path, 264, 272, 273, 277, 288, 307
- Bending stiffness, 298, 299
- Berkovich tip, 204, 206
- Bethe-ansatz, 528, 529, 539
- Biological applications of CNT, 160
 - AFM tips, 163
 - imaging experiments, 160
 - nanoscale injection, 163
- Bloch, 520
- Block copolymers, 834, 839–853
- Bogoliubov transformation, 531
- Boson, 511, 528, 547
- Bosonization, 548, 553
- Brownian motion, 30–38
 - correlation functions, 31–33, 36, 37
 - hydrodynamic damping/mobility, 31, 35
 - Langevin equation, 31, 36
 - power spectrum density, 31, 32
- Building block, 512, 543, 547, 726
- Built up edge (BUE), 889

- CaCu₃Ti₄O₁₂ (CCTO), 613, 614
- C-AFM, 613, 629, 636
- Calibration grating/grid, 272, 278, 284, 286
- Cantilever, 262
 - bending, 152
 - dynamics, 580
 - frequency shift, 91
 - modification, 730, 731
 - neutral axis, 266
 - shear centre, 267, 318
 - stiffness, 870
- Capacitive footprints, 111
- Capacitive forces, 92
- Carbon nanotube (CNT), 24, 33–38, 49, 50, 487, 496, 499, 501
 - atomic force microscopy, 129
 - carbon atoms, 130
 - field effect transistor, 120
 - MWNT structure, 133
 - sp*² bonding, 131
 - SWNT, 131
 - synthesis, 134–135
 - tip probes, 96
- Catalytic decomposition, 134
- Cell and tissue mechanics, 464
- Cells, 327, 351
 - cadherin, 352
 - ergtoxin-1 (ErgTx1), 354
 - hERG K⁺ channels, 355
 - MyEnd cells, 352
- Centroid tracking, 397–398
- Characteristics/characterization, CNT
 - tips, 141
 - attachment angle, 147
 - FM-AFM, 144
 - high aspect ratio, 146
 - imaging quality, 143
 - MWNT buckling, 143
 - SEM characterization, 142
 - tip resolution, 141
- Charge emission to the oxide layer, 108
 - discharge behavior, 109
- Charge inhomogeneity, 520, 525
- Charge injection, 95, 103
 - continuous discharge, 109
- Charge transfers at nanotube-metal interface, 119
- Charges trapped into the oxide layer, 125
- Checkerboard, 522, 542, 553
- Chemical force microscopy (CFM), 262
 - adhesion force measurements, 149
 - CNT tips, 148
 - covalent modification, CNT, 149
 - electrical testing, 151
 - hydrophillic/hydrophobic surfaces, 150
- Chemical mechanical polishing/planarization (CMP), 874, 897
- Chemical potential, 527, 529, 535–537, 540, 541, 544, 549, 550
- Chemical vapor decomposition (CVD), 134
- Chemisorption, 883, 884, 892
- Chromatin structure, 738, 739
- Chromosome fragments, 744
- Closed loop AFM, 202

- Cluster gas, 508
- Clusters, 556
 - bipartite, 540, 548, 549, 552, 554–557, 561, 563
 - face centered square (FCS), 559
 - nonbipartite, 540, 542, 545, 547, 548, 558
 - octahedron, 542, 547, 552, 559
 - pyramid, 526, 559
 - square, 512, 513, 547, 548, 552, 553, 555, 557–559, 562
 - tetrahedron, 542, 547, 552, 556–563
 - topology, 547, 556
- Coating, 218
- Coefficient of restitution method, 248–250
- Coherence, 516, 518, 524, 546
- Coherent BCS pairing, 555
- Colloidal AFM, 895
- Colossal magnetoresistance (CMR), 527, 541, 556
- Comb drive, 300, 316
- Compaction modeling, 895
- Complex conductivity, 617
- Complex dielectric function, 616, 627
- Compliance method, 289, 290, 313
- Compliant materials, 206
- Computational fluid dynamics (CFD), 429, 445, 448
 - finite volume method (FVM), 449
 - transport phenomena, 448
- Computed tomography (CT), 429, 430
 - micro-computed tomography (micro-CT), 431, 438, 442–444, 447, 453, 468, 472, 476
- Computer-aided-tissue-engineering (CATE), 429
- Condensation, 524–558
 - bose, 555, 556, 564
 - coherent, 524
 - electron, 559, 561
 - incoherent, 524–526, 548, 554, 555
 - pair, 559, 561
- Configuration, 547
 - cluster, 544, 545, 558
 - electron, 521, 550, 553
 - ionic, 544
- Confinement, 508, 511, 542
- Constant-amplitude (CA), 4
- Contact area, 210
- Contact lenses, 374
- Contact mechanics, 238–241, 583
- Contact potential difference, 93, 121
 - ionization probe, 693–694
 - Kelvin probe, 691–692, 694, 699, 700, 713
 - non-vibrating probe, 692–693
- Contact-resonance atomic force microscopy, 587
 - 1D nanostructures, 587
 - precision and accuracy, 585
 - technique, 579
- Contact-sliding, 818
 - friction, 818, 820
 - mechanical stimulation, 819, 820
- Contact stiffness, 11, 14, 210, 281
 - lateral, 288, 290, 297, 313
- Cooling rates, 207, 219
- Core-shell model, 595, 599
- Cornea, 370
- Coulomb repulsion, 554
- Critical temperatures, 543, 550, 551, 553, 561, 564
- Critical volume for stress concentration, 206
- Cross-polarization, 73, 80
- Cross-polarized, 67
- Crossbeam, 287, 312
- Crossover, 542, 544, 546, 551, 561, 562
- Crosstalk, 292
 - conversion factor, 292, 293, 302
 - mechanical, 274, 282, 290, 291, 294, 302, 318
 - optical, 275, 276, 282, 286, 291, 302
- Crystallization, 833–835, 857–858, 862
- CuO, 614, 629
- Cuprates, 507, 513–526, 537, 540, 541, 550, 564
- Curie, 528, 546
- Damping, 242–244
- Debye behavior, 619
- Debye length, 621
- Decoupling, 529
- Deformation, 270
 - in-plane, 270, 271, 274, 275, 290
 - torsional, 271, 274, 276, 288
- Degeneracy, 534

- Kramers, 534, 548
 - permanent, 548
- Delocalization of injected charges, 105
- Depolarization, 73, 75, 80, 83
- Depolarization factor, 73
- Depolarization measurements, 73
- Depth Sensing Instruments (DSI), 200, 202, 219
- Derjaguin-Muller-Toporov (DMT)
 - model, 263
- Diamagnetic lateral force calibrator (D-LFC), 302, 316
- Diamagnetic levitation, 302, 303, 318
- Dielectric relaxation, 617
- Diffusive and ballistic transport in
 - CNT, 121
- Dimensionality, 192
- Dipole moment, 509
- Dip-pen nanolithography (DPN), 796, 797, 902
- Directional effect, 282, 302
- Discharge behavior, 107
- Disulfide bond reductions, 412
- Domain, 545, 549, 550, 559, 564
- Double functionality patterning, 805
- Double layer, 619
- Double-Well model, 662, 672–674, 680
- Dynamic force spectroscopy, 15
- Dynamic quantized fracture mechanics (DQFM), 488–489, 493
- Elastic bending constant, 144
- Elasticity, 191, 192, 495–496
 - contact models, 370
 - measurements, 368–369, 730
 - visco, 365, 368, 369
 - Young's modulus of (E), 370
- Elastic modulus, 176, 179, 183, 184, 188, 201, 203, 213, 219
- Elastic recovery, 204
- Electric field-induced mechanical resonance, 573
- Electric force microscopy (EFM), 152–154
- Electric modulus M^* , 628
- Electrical characterization, 147
- Electrical conduction, 133
- Electrochemical etching, 62
- Electrode–dielectric interfaces, 616
- Electrode polarization, 616, 620
- Electron-hole pairs, 531
- Electron repulsion, 510, 519, 522
- Electron work function
 - friction monitoring, 698–700, 704–713
 - surface mapping, 714, 715
- Electrostatic excitation, 93
- Electrostatic force, 235
- Electrostatic force microscopy (EFM), 90
 - ac-EFM, 121
 - of carbon nanotubes, 97, 104, 117
 - dark and bright features, 106
 - of DNA, 100, 101
 - of native charges, 102
 - signals, 91, 92
 - $1\omega/2\omega$ EFM, 93
- Electrostatic lever-arm of CNT, 124
- Electrostatics interactions in SWCNTs, 115
- Enamel, 204
- Energy gap, 510, 518, 543, 545, 548, 556, 559
- Energy level alignments, 121
- Ensemble, 507, 513, 519, 525, 541, 544, 549
- Enzymatic reaction, 729
- Erythrocytes, 349
 - cystic fibrosis, 349
 - cystic fibrosis transmembrane conductance regulator (CFTR), 349
- Escherichia coli nucleiod, 743
- Etching, 63
- Eukaryotic genome architecture, 731
 - DNA-binding proteins, 733
 - fundamental structures, 735
 - mitotic chromosomes, 738
- Eutectic, 820
- Excentricity, 305
- Excitations, 539
 - charge, 539
 - electron, 519
 - energy, 561
 - quasiparticle, 519
 - spin, 520, 539, 544, 561
- Fabrication of CNT, 137
 - batch fabrication process, 137–139

- electron microscopy, 135–136
- gluing, 135
- MEMS microgripper, 137
- “pick-up method”, 137
- post-growth attachment, 139–141
- Face centered geometry, 559
- Far-field background, 82
- Far-field suppression, 79
- Fast-scanning, 728
- Fermi
 - energy, 523, 534
 - impulse, 534
 - surface, 520, 522, 527
- Ferromagnetic, 509, 511, 518
- Ferromagnetism, 509, 512, 515, 527, 529, 537, 547, 548, 554, 556–558, 563
- Fibronectin, 806–808
- Field enhancement, 58, 67, 68, 70, 72, 80, 82, 775, 777, 780
- Field enhancement factor, 84
- Films
 - LbL, 917, 918, 920, 921, 923, 931, 935, 936
 - nanostructured, 916, 922, 924, 925, 928
 - thin, 915, 917, 923, 924, 931, 940
- Fingerprinting biomolecules, 405
- Finite element analysis (FEA), 277, 290, 300
- Finite element method (FEM), 455, 458, 466
 - lattice method, 457
 - multi-phase lattice Boltzman method (MLBM), 463
- Finite shear elasticity, 172, 178
 - simple liquid approach, 172
- Fission, 508
- Flow porosimetry, 469, 470, 472
- Flow stress, 869, 873, 874, 876, 878, 890–892, 894, 903, 904
- Fluid film parameter, 882–884
- Fluid/liquid cell, 273, 307
- Fluid properties, 174
 - linear rheology, 174–176
 - non-linear rheology, 176–177
 - viscoelastic approach, 173–174
- Fluorescence resonance energy transfer (FRET), 395, 397, 398
- Focused ion beam (FIB), 279, 280, 301
- Force balance, 264, 284, 294
- Force curve, 202, 213, 928, 933, 935
- Force-distance curve (FDC), 262, 281, 285, 286, 311
 - lateral, 272, 274, 314
- Force gradient, 90, 91, 93, 106, 113
- Force measurement, 727
- Force sensing, 23, 24, 29–33, 38–45
- Force spectroscopy, 142
- Force transducer (FT), 202
- Force volume, 326, 351
- Fourier series, 7
- Fourier transform infrared spectroscopy (FTIR), 802, 803
 - cantilever oscillation, 6
 - equation of motion, 6
 - frequency-modulation (FM) technique, 7
 - tip-sample forces, 7, 10
- Frequency-modulation (FM) technique
 - equation of motion, 4
- frequency-modulation (FM) technique, 5
- Fresnel matrix, 66
- Friction, 856, 859–860
- Friction coefficient, 263, 281, 283, 286, 301
- Friction force, 263, 274, 277, 278, 283, 284, 299, 302
- Friction force microscopy (FFM), 152, 260, 883, 885–886
- Friction loop, 273, 276, 277, 281, 282, 288, 289, 302, 309, 310
- Frictionless contact, 240
- Fullerenes, 509
- Gas adsorption, 469, 470, 472
- Gas pycnometry, 469–471
- Geometrical optics method, 272, 273, 307
- Giant- κ ceramics, 629
- Giant permittivity, 614
- Glass fibre, 298, 299, 315
- Gold, 821
- Grain orientation, 873, 892–894
- Graphene sheet, 130
- Grazing impact, 251
- Grinding, 872, 896

- Hardness, 210
 definition, 870
 indentation testing, 869, 871, 899–900
 macro-scale testing, 870
 scratch testing, 874–878, 889–891,
 896, 899–900, 903–904
- Hartree-Fock, 521
- Hertz equation, 207
- Higher-order architectures, 725
- Highest occupied molecular and lowest unoccupied molecular orbital (HOMO-LUMO), 508, 544
- High-resolution imaging, 746
- High-resolution patterning, 801, 806, 808
- High T_c
 cuprates, 523, 564
 doped, 541, 546
 superconductivity, 508–511, 513,
 515–520, 526–529, 541, 546–548,
 550, 551, 554, 559, 563
- Hitler-London, 520
- Hole-rich and hole-poor, 516, 549, 553,
 555, 558, 562
- Holon, 520
- Hopping conduction, 617
- Horizontal lever, 286, 287, 312, 313
- Hubbard insulator, 563
- Hubbard model, 511, 529–539, 541, 548
- Hydrodynamic force, 244
- Hysteretic behavior of CNTFETs, 124
- Image charge effects, 92
- Imaging capability, 144
- Imaging the density of states in CNT,
 118
- Impedance spectroscopy, 622
- Impedance Z^* , 628
- Incoherent pairing (IP), 507, 519, 523,
 553, 554, 558, 562–564
- Incommensurate phase, 531, 538, 539
- Incommensurate state, 537
- Indentation, 368–370
- Indentation size effect (ISE), 872–874,
 879, 889, 899, 904, 905
- Indentation size scale effect (ISE), 212
- Inhomogeneity, 516, 519, 521, 525, 540,
 549
- Inner-shell charging of CNTs, 112, 114
- Instability, 528
 electron, 528, 530, 537, 555
 ferromagnetic, 556
 Nagaoka, 558
 separation, 556
- Instrument compliance, 209
- Instrumented indentation, 577
- Insulating blocking layers, 625
- Inter-configuration, 544
 energy gap, 544
 fluctuation, 525, 544, 554
- Intermittent contact, 248
- Intermittent contact mode, 294, 304
- Internal barrier layer capacitor (IBLC)
 model, 613, 622, 629
- Intraocular lenses, 379
- Ionization, 544
- Ironing, 902–904
- Jahn-Teller, 508
- Johnson-Kendall-Roberts (JKR) model,
 263
- Kelvin force microscopy (KFM), 93
 amplitude modulation, 93, 540
 on CNTFET, 122
 frequency modulation, 93
 voltage drop, 123, 522, 564
- Kelvin probe, 692
 friction, 692, 699–701
 wear, 695, 703
- Kelvin probe force microscopy (KPFM),
 613, 623, 633
- Kondo model, 531, 539
- k -space, 518, 519, 521, 528, 534, 551
- Lateral electrical nanobalance (LEN),
 300, 301, 316
- Lateral force microscopy (LFM),
 260–264, 271–273, 275, 277, 279,
 280, 282, 286, 288, 291, 304, 306,
 318, 319
- Lateral resolution in EFM and KFM,
 94, 96, 97
- L- Dipalmitoyl-phosphatidyl-choline
 (DPPC), 12
- Lennard-Jones (LJ) potential, 244, 246
- Lens, 376
- Level, 507

- crossings, 543, 545, 548, 549, 555, 559
- discrete, 507, 509, 510, 542, 553
- multiple, 548
- nanoscale, 527, 540
- Linear charge densities, 113, 117
- Linear elastic fracture mechanics (LEFM), 488, 489, 492
- Linearisation, 283
 - hardware, 283
 - software, 283
- Linear nanostructures, 23, 24, 33–38, 917
 - semiconductor nanowires, 28, 33–38
- Lithography, 758, 762
- Lorentz force, 306, 317
- Luttinger liquid, 215, 869, 882–887, 890–892

- Magnetic force microscopy (MFM), 155, 513
 - Fe-MWNT, 156
 - image resolution, 157
 - magnetic storage systems, 155
- Magnetic resonance imaging (MRI), 429, 430
- Magnetic susceptibility, 303
- Magnetic trap, 397, 399
- Magnetization, 530
 - longitudinal, 514, 530, 532, 537, 539, 545
 - transverse, 537
- Magnetostrictive material, 305
- Magnification factor, 243
- Manganites, 507, 515, 520, 526, 537, 541, 564
- Mapping mechanical properties, 202, 204, 219
- Mass moment of inertia, 304, 305
- Maximum likelihood methods (MLM), 412
- Maxwell's equations, 617
- Maxwell-Wagner, 616, 629
- Maxwell/Wagner/Sillars polarization, 616, 619
- Mechanical experimental characterization, 473
 - tensile, compression, bending test, 473, 475
- Mechanical properties of 1D nanostructures, 604
 - applications, 604
 - measurement techniques, 572
 - surface effects, 601
- Mechanochemical reactions, 405
- Mechano-transduction, 428
- Melting, 835, 843, 845, 851–852
- Mercury porosimetry, 469–471
- Metallic shell, 114
- Metallic tips, 62
- Metal rolling, 887–889
- Micro-electromechanical systems (MEMS), 260, 300, 301
 - device, 299, 302, 316
- Micro-lateral force sensor (MLFS), 300, 316
- Micro/Nanoelectromechanical systems (MEMS, NEMS), 867, 897, 901
- Misalignment, 275, 276, 291
- Misfit angle, 292
- Mixed valence, 520, 550, 554
- Modeling of indenter, 875–877
- Modulation, 526
 - pair, 564
 - spatial, 521
 - spin, 533
- Molecular recognition force spectroscopy (MRFS), 328
 - Hooks law, 334
- Moment arm, 269, 286, 291
- Mott-Hubbard, 518, 522, 528, 544, 546–549, 551, 555, 556, 559, 563
- Multiferroic, 507, 509, 528
- Multi fractal scaling law (MFSL), 505
- Multiscale-modelling, 429, 438, 444, 448, 454, 459, 461–463, 466
- MWNT tips, 160

- Nagaoka theorem, 548, 550, 555, 558
- Nano-MIS, 643
- Nanoalloying, 813, 819, 820
- Nanocluster, 507, 519, 528, 540, 544, 553, 564
- Nanocrystallization, 819, 820
- Nanoelectrodes, 151
- Nanoelectromechanical systems (NEMS), 260
- Nanofibers, 916, 922, 924

- Nanofluidics, 181, 183
- Nanoindentation, 199
- Nanolithography, 812, 815, 901
 - AFM lithography, 815
 - scanning probe alloying nano-
lithography, 812, 821,
822
- Nanomechanical testing, 260
- Nanoparticles, 260, 261, 922, 933, 939
 - array, 261
- Nanoscale resolution, 728
- Nanoscale Weibull Statistics (NWS),
499, 504
- Nanoscopy, 487
- Nanostructure, 826
- Nanotechnology, 915, 916, 940
- Nanotensile tests, 499–502
- Nanotubes, 509, 520
 - bundle, 493, 494, 496, 504
 - charge, 110
 - fatigue, 494–495
 - fracture strength, 492–493
 - hierarchical simulations, 503–505
 - impact strength, 493
- Nanowires, 596
 - Te, 596
- Near field, 758, 768, 769, 777, 785
- Neutral plane. See Neutral point, 889
- Neutral point, 868, 878, 889
- Newtonian fluid, 176
- NiO, 614
- Non-contact AFM (NC-AFM), 242
- Nonohmic contacts, 623
- Non-vanishing elasticity, 187
- Normal state, 522
- Nuclear magnetic resonance (NMR),
561
- Nuclear pore complex (NPC), 731
- Numerical models, 70

- Oliver and Pharr procedure, 200, 206,
207, 209
- One DOF system, 244, 245, 248, 250
- Optical interference, 282
- Optical near-field microscopy, 158–159
- Optical resolution, 767
- Optical trapping, 24–29, 33–38, 45–53
 - gradient force, 25, 26
 - scattering force, 26
- Optical tweezers, 23, 24, 29–31, 39, 40,
42–45
 - Raman tweezers, 23, 45–53
- Orange peel, 869, 886–888
- Order, 518
 - long-range, 518, 533, 539, 541, 543
 - parameter, 519, 527, 530, 543, 546,
549, 550, 561
 - short-range, 513, 523, 524, 534
- Orientation imaging microscopy (OIM),
893
- Oscillatory shear frequency scanning
method, 183
- Overall conversion factor, 265, 267, 278,
280, 281, 306

- Pairing, 556
 - BCS-like, 555, 556, 558
 - charge, 513, 524, 541, 548, 550, 553,
559, 561
 - coherent, 524, 553–562, 564
 - incoherent, 524, 542, 563, 565
 - spin, 524, 526, 540, 543, 545, 553–561,
563, 564
- Pairs, 563
 - charged, 563
 - cooper, 518, 519, 553
 - electron, 541, 546
 - hole, 526
 - preformed, 562
 - spin, 553
- Parallel scan method, 283, 311
- Paris' law, 490–491, 494–495
- Particle tracking, 24, 30–33, 36–44
- Peierls distortion, 508
- Permittivity ϵ^* , 628
- Perovskite, 515, 547, 559, 615
- Perturbation, 543
 - cluster, 543
 - theory, 512, 529
- Phase, 98, 507
 - diagram, 507, 518, 522, 534, 537–539,
558, 560, 565
 - incommensurate, 533, 539
 - separation, 507, 516, 526–529, 537,
540, 542, 549–552, 556–558, 562
 - shift, 91, 98
 - transition, 508, 512, 539, 542–546,
549–551, 561, 564

- Phase separated materials, 619
 Photoemission, 522, 523, 525
 Photolithography, 759, 761
 Photonic force microscopy, 23, 24, 39–45
 Physisorption, 883, 884, 892
 Pico-Newton mechanical forces, 729, 730
 Piezoelectric actuator, 287, 304, 305
 Piezoresistive material/sensor, 295–298, 315
 Pile up, 203, 214, 215
 Placzek's formalism, 64
 Polarization, 768, 774, 778, 782, 786
 Polarization control, 61
 Polycarbonate, 213, 218
 Polyethylene, 214
 Polyethylenterephthalate, 214
 Polymers, 759, 765, 775, 781, 785
 Polypropylene, 203, 214
 Polystyrene, 213, 837, 856–861
 Position-sensitive detector (PSD), 262–265, 267, 268, 271–278, 280, 282, 287
 Powder metallurgy (PM), 894–896
 Probability, 553
 Probes, 62, 274, 767
 aperture, 280, 773, 779, 780
 apertureless, 777, 778
 cantilever, 767, 770, 773
 colloidal, 274, 278, 288, 290, 294, 295, 306, 313
 fiber, 773
 target, 275, 290
 test, 274, 290
 Probing the band structures of nanotubes, 116
 Prokaryotic genome architecture, 739
 bacterial DNA-binding proteins, 739, 741
 Protein, 407–408
 folding, 401
 unfolding, 401, 407, 408, 412
 Proximity effect, 522, 563
 Pseudogap
 charge, 542, 561, 563
 spin, 542, 555, 559, 561
 ptcr, 623
 Pull-off force, 280, 281, 309, 310
 Pyramid, 526, 542, 552, 559
 Pyrochlore, 547
 Pyrolytic graphite (PG), 303
 Quality factor, 94
 Quantized fracture mechanics (QFM)
 atomistic simulations, 496–498
 Paris' law, 490–491
 Weibull statistics, 491
 Quantum coherent state, 553
 Quantum critical points (QCP), 513, 545, 548, 554, 559, 563
 Quantum Monte Carlo, 528, 539, 540
 Quasi-hole, 537
 Quasiparticles, 513–520, 534
 Raman scattering, 64
 Raman scattering tensors, 65
 Raman spectroscopy, 24, 45–53, 58
 red blood cells, 48
 surface enhanced Raman spectroscopy, 52, 53
 Random phase approximations (RPA), 511, 528
 Rapid prototyping (RP), 434
 computer aided system for tissue scaffolds (CASTS), 436, 437
 Rare earth, 521
 Recognition imaging, 732
 Reference beam, 298
 horizontal, 298
 vertical, 294, 295
 Reference spring, 299, 300, 302
 Refractive index, 273, 274
 Refractive surgery, 374
 Regime prior, 189
 Relaxation modulus, 190
 Residual stress, 236
 Resonance, 7
 flexural, 270, 304
 frequency, 90
 primary, 248
 subharmonic, 248, 249
 superharmonic, 249
 torsional, 270, 275, 304, 305
 Resonance valence bond (RVB), 520
 Retina, 381
 Rf circuits, 614
 Roughness, 917, 918, 921, 922, 927, 931, 935

- Rouse model, 184
r-space, 519
Ruthenocuprate, 522
- Saturated ferromagnetism, 555–558
Scaffold, 428
 permeability, 427
Scanning-by-probe (SBP), 277, 284
Scanning-by-sample (SBS), 277
Scanning capacitance microscopy and SIM, 636
Scanning electron microscopy (SEM), 469–471
Scanning impedance microscopy (SIM), 613, 629
Scanning probe microscope (SPM), 473, 487, 494, 726
 atomic force microscopy (AFM), 474
 nanoindentation (NI), 475
 scanning force microscopy (SFM), 474
Scanning tunneling microscopy (STM), 154–155, 520, 522, 523, 526, 541, 553, 561, 563, 565, 613, 630
Schottky diodes, 616
Schottky-type barriers, 623
Sector
 charge, 539, 548, 555
 spin, 545, 548
Segregation, 549, 555
Selection rule, 67
Selective enhancement, 80
Self-assembled monolayers (SAMs), 796
Sensitivity, 262–265, 267, 268, 271–277, 280, 283, 286, 289, 294, 296, 299, 306
Sensors, 916, 920, 922, 924–928, 932, 940
Shear centre, 292, 293
Shear modulus, 268, 271, 283
Side capacitance effects, 94
Side illumination, 59
Signal-to-noise ratio (SNR), 294, 301
Silicon nitride, 816, 817, 822
Silicon nitride tips, 147
Single dextran monomers, 15
 chain-like molecules, 15
 single-click model, 17
Single functionality patterning, 805
Single-walled carbon nanotube (SWNT), 131
Size scale, 205
S-layer, 346
Sliding contact, 240
Slip-line field models, 869, 873
Slippage, 186
Smudge, 887, 889
Sneddon elastic approach, 205, 206, 208, 210
SNOL, 758, 770, 772
SNOM, 767
Softening effect, 244
Soil, 916, 927, 932, 935, 937
Solid-like correlations, 172, 179, 180, 182, 190, 193
Space, 520
 momentum, 520, 523, 534, 551
 real, 509, 515–517, 543, 551, 554
Spatial inhomogeneities, 507, 540, 551–553, 559
Spatial resolution, 153, 726
Spectrum, 553
 discrete, 553
 excitation, 553
Spin-charge separation, 513–516, 520, 523, 527, 539, 540, 543, 551, 555, 558, 563
Spin liquid, 555, 556
Spinodal, 561
Spinon, 520
Spin pseudogap, 522, 543, 546, 553, 561–563
Spontaneous transition, 520, 521, 543, 549, 550, 556, 562
Spring constant/stiffness, 268
 lateral, 268, 269, 271, 299
 normal, 269, 278, 280, 283
 torsional, 262, 264, 265, 280, 294, 304–306
Stick-slip transition, 286, 302
STM imaging, 154
Strength, 261
 adhesion, 261, 262
 shear, 260, 262, 288
Stribeck curve, 883
Structural maintenance of chromosomes (SMC), 728
Substrate, 271

mirrored, 271–273
 sloped, 284
 Substrate effect, 217
 Super-ferroelectric, 509
 Super-paramagnetic, 509
 Surface barrier layer capacitor (SBLC)
 model, 616, 629
 Surface charge, 624
 Surface effects, 188
 Surface energy, 240
 Surface force apparatus (SFA), 181, 182
 Surface potential, 92, 94
 Surface stress, 235, 236
 Susceptibility, 545
 charge, 545, 550, 560
 peak, 560
 spin, 525, 545, 550–551, 553, 560, 562
 Suspended platform
 magnetically, 302–303
 mechanically, 299
 Suspension stiffness, 301
 Suspension system, 301–302
 Symmetry, 546–548
 breaking, 546, 564
 electron-hole, 530–531, 542, 545, 547,
 549, 558, 562, 563
 Tapping mode, 3, 5, 12
 TERS setup, 60
 Tetrahydropyran (THP), 802, 804
 Thermal vibration, 145
 Thermochemical nanolithography
 (TCNL), 797–798
 covalent functionalization and
 molecular recognition, 805–809
 unmasking amines groups, 802
 unmasking carboxylic acid groups,
 799–801
 wettability modification, 802–805
 Thermococcus kodakaraensis, 744
 Thin coatings, 215
 Three-point bending test, 575
 Throughput, 772, 773, 788
 Time-lapse imaging, 729
 Tip, 275
 apex, 299, 302, 318
 convolution, 204
 damage, 272, 306, 309, 315, 318
 geometry, 370

 orientation, 878–879
 shank, 275, 297, 299, 301
 Tip chemistry, 328
 aminofunctionalization, 328
 3-aminopropyltriethoxysilane
 (APTES), 328, 329
 benzaldehyde, 329
 benzophenon, 331
 ethanolamine, 328, 329
 fluoren-9-yl-methoxycarbonyl-
 protected (FMOC), 331
 His6 tag protein, 331
 hormones, 327, 331
 lysine residue, 329
 poly(ethylene glycol) (PEG), 330
 thiol linker, 328
 tris-NTA, 331
 Tip-amplification tensor *A* expressing,
 67
 Tip-Enhanced Field Modeling, 67
 Tip-Enhanced Raman Scattering, 58
 Tip sample contact, 227, 237
 Tip-sample interaction forces, 12
 DPPC monolayers, 12, 15
 Fourier series, 7
 frequency shift, 8, 10
 large amplitude approximation, 9, 10
 phase shift, 8
 Tissue engineering (TE), 429
 Topography and recognition imaging
 (TREC), 344
 amplitude variation, 344
 cross-talk, 340
 driving frequency, 336, 341
 feedback loop, 335
 full amplitude, 334
 half amplitude, 334
 magnetic AC mode (MAC mode),
 327, 331
 oscillation amplitude, 333
 Q-factor, 332, 334
 resonance frequency, 331, 341
 specificity proof, 332, 344
 Topography feedback, 273, 278, 280,
 283, 284, 304
 Torque, 262, 264–267, 269, 274, 277,
 282, 286–288, 291, 299, 306, 312,
 313
 Torsional arm, 266, 286, 296, 318

- Torsional resonance, 263, 264, 294, 304
- Torsional Sader method, 304, 317
- Total internal reflection fluorescence (TIRF), 405
- Transition matrix (T-Matrix), 25
 - Maxwell stress tensor, 27, 28
- Transition metal, 509, 515, 527, 541
- Transmission mode, 59
- Transverse polarizability of nanotube, 113
- Triple functionality patterning, 805
- Triplet, 547, 555–557
- Tunneling, 507, 509, 510, 513, 515, 520–524, 526, 529, 564
- Tunnelling current, 300, 301
- Two energy, 523
- Type III restriction enzyme, 729
- Ultra high molecular weight polyethylene (UHMWPE), 212
- Ultrasonic force microscopy, 578
- Uncertainty, 269, 272, 280, 283, 287, 290, 298, 301, 306, 318
- Unloading exponent, 214
- User-friendliness, 306
- van der Waals (vdW) force, 244–246
- Vertical lever, 314, 315
- Viscoelasticity scanning method, 183
 - Ge, 185
 - Maxwell model, 184
- Viscoelastic nanoindentation, 214
 - behavior, 201, 214
 - models, 201
- Viscoelastic nose, 215
- Viscous fluids, 187, 188
- Viscous modulus, 175–177, 179, 183, 184, 190–192
- Wear, 263, 284, 822
 - abrasive, 818, 823, 826
 - adhesive, 818, 826
 - modes, 880
 - nanowear, 818, 825
 - scar, 823, 825
 - tip, 817
- Wear tests, 880
 - area, 880–882
 - line, 878–879, 896
- Wedge method, 263, 277–282, 286, 290, 301, 309, 310, 318
- Weibull statistics, 491, 499
- Wettability modification, 802–805
- Wetting, 186, 187
- Wick theorem, 529
- Worm-like chain model, 403
- Xenopus oocyte, 725
- $YBa_2Cu_3O_6 + x$ (YBCO), 518, 524
- Zeeman, 527