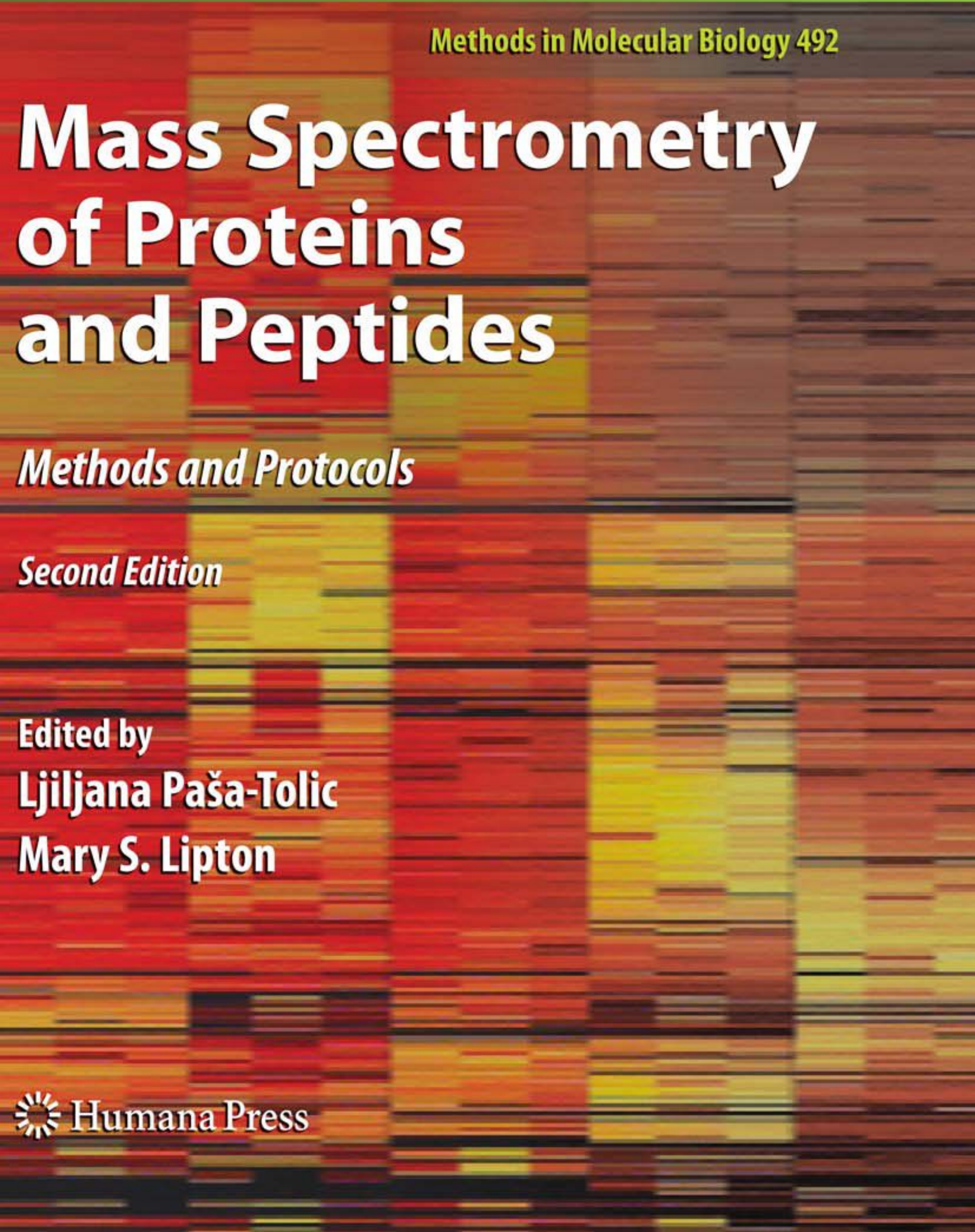
Springer Protocols

Methods in Molecular Biology 492



Mass Spectrometry of Proteins and Peptides

Methods and Protocols

Second Edition

Edited by
Ljiljana Paša-Tolic
Mary S. Lipton

Humana Press

Mass Spectrometry of Proteins and Peptides

Series Editor
John M. Walker
School of Life Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For other titles published in this series, go to
www.springer.com/series/7651

METHODS IN MOLECULAR BIOLOGY™

Mass Spectrometry of Proteins and Peptides

Methods and Protocols

Second Edition

Edited by

Mary S. Lipton and Ljiljana Paša-Tolic

Pacific Northwest National Laboratory, Richland, WA, USA

Humana Press

Editors
Mary S. Lipton
Pacific Northwest National Laboratory
Fundamental & Computational
Sciences Directorate
Richland, WA
USA

Ljiljana Paša-Tolić
Pacific Northwest National Laboratory
Environmental Molecular
Sciences Laboratory
Richland, WA
USA

ISBN:978-1-934115-48-0 e-ISBN: 978-1-59745-493-3
ISSN: 1064-3745 e-ISSN: 1940-6029
DOI: 10.1007/978-1-59745-493-3

Library of Congress Control Number: 2008940986.

© Humana Press, a part of Springer Science+Business Media, LLC 2009
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Humana Press, c/o Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013 USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.
While the advice and information in this book are believed to be true and accurate at the date of going to press, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Cover illustration: Chapter 3, Figure 2: The above experiment represents the protein abundance values obtained for four culture conditions of Y. pestis (the bacteria agent responsible for plague). The growth conditions were meant to mimic growth conditions of the bacteria when in the flea vector (ambient temperature) and after infection into the mammalian host (37 °C with millimolar amounts of divalent calcium as experience in the bloodstream and extracellular fluid, and an absence of free divalent calcium as experience inside the host cells) The abundance values were transfomed into z-scores for the whole cell set using the average protein value for the whole cell set. The z-score values were calculated separately for the pellet and supernatant data together using their average abundance. The abundance values thus are represented by color for 992 Y. pestisproteins identified in triplicate analyses of each sample. Significant protein abundance value changes due to culture condition and cell fractionation can easily be observed. +Ca2+, growth with 4 mM Ca2+; –Ca2+, growth with no Ca2+; ND, not detected.

Printed on acid-free paper

springer.com

Preface

When the last edition of this book was published in 2000, the field of proteomics was in its infancy. At that time, multidimensional liquid chromatographic separations were being introduced as an alternative to traditional gel-based techniques for separating complex protein and peptide mixtures prior to mass spectrometric detection. Today, this approach – referred to as shotgun proteomics – is considered routine for large-scale global analyses of protein mixtures.

Now in its adolescence, proteomics is fundamentally transforming biological and medical research. Much of this transformation can be attributed to technological advancements, particularly in mass spectrometry. Much wider accessibility of high-resolution and mass measurement accuracy instrumentation in recent years has initiated a new revolution in the field by providing more reliable data and shifting the focus from cataloging proteins to precisely quantifying changes in protein abundance over time and in response to stimuli. Advanced mass spectrometers and novel ion dissociation schemes such as electron transfer/capture dissociation make it possible to venture boldly into the maze of protein posttranslational modifications, which are an integral component of understanding functional proteomics in the spatial and temporal domains. Another area that has benefited from these advancements is top-down proteomics, an emerging method essential for characterizing various protein variants that has potentially high impact in biomedical research. Another breakthrough application that holds great potential in pharmaceutical as well as biomedical research is imaging mass spectrometry, which enables high-resolution spatial characterization of complex materials (such as tissues) with simultaneous information on chemical composition. By virtue of enhanced sensitivity, dynamic range, and throughput, these new tools are applied to generate robust quantitative measurements in support of clinical studies. However, before a proteomics platform can be used for routine clinical applications, certain challenges, such as the ability to provide the throughput needed to enable a statistically meaningful number of analyses for a clinical setting, must be overcome. Ion mobility separations have recently emerged as a means for significantly increasing throughput proteomics and as such are expected to have a transformational impact on the field of proteomics and its applications in broad areas of biological and medical research.

Spanning fields from microbial forensics and clinical applications to protein structure, dynamics, and function, the following chapters present new and updated methods for mass spectrometric characterization of proteins and peptides. As with the first two volumes, the intent of this latest edition of *Mass Spectrometry of Proteins and Peptides* is to provide the reader with step-by-step instructions, along with insight into the pitfalls and nuances of apparently straightforward techniques in the hope of facilitating new discoveries.

Mary S. Lipton
Ljiljana Paša-Tolic

Contents

Preface v

Contributors ix

1. Generation and Analysis of Multidimensional Protein Identification Technology Datasets 1
   Selene K. Swanson, Laurence Florens, and Michael P. Washburn
2. Quantitative Peptide and Protein Profiling by Mass Spectrometry 21
   Alexander Schmidt, Birgit Bisle, and Thomas Kislinger
3. Label-Free Relative Quantitation of Prokaryotic Proteomes Using the Accurate Mass and Time Tag Approach 39
   Kim K. Hixson
4. Classical Proteomics: Two-Dimensional Electrophoresis/MALDI Mass Spectrometry 65
   Ursula Zimny-Arndt, Monika Schmid, Renate Ackermann, and Peter R. Jungblut
5. The Use of Difference In-Gel Electrophoresis for Quantitation of Protein Expression 93
   Rajat Sapra
6. Liquid-Chromatography-Mass Spectrometry of Thylakoid Membrane Proteins 113
   Christian G. Huber, Anna-Maria Timperio, Hansjörg Toll, and Lello Zolla
7. High Accuracy Mass Spectrometry in Large-Scale Analysis of Protein Phosphorylation 131
   Jesper V. Olsen and Boris Macek
8. Manual Validation of Peptide Sequence and Sites of Tyrosine Phosphorylation from MS/MS Spectra 143
   Amy M. Nichols and Forest M. White
9. Assigning Glycosylation Sites and Microheterogeneities in Glycoproteins by Liquid Chromatography/Tandem Mass Spectrometry 161
   Yehia Mechref, Milan Madera, and Milos V. Novotny
10. Structure Analysis of N -Glycoproteins 181
    Stefanie Henning, Jasna Peter-Katalinic´, and Gottfried Pohlentz
11. Capillary Zone Electrophoresis-Mass Spectrometry for the Characterization of Isoforms of Intact Glycoproteins 201
    Christian Neusüß, and Matthias Pelzing

12. Top-Down Proteomics on a High-field Fourier Transform Ion Cyclotron Resonance Mass Spectrometer 215
Séverine A. Ouvry-Patat, Matthew P. Torres, Craig Gelfand, Hung-Hiang Quek, Michael Easterling, J. Paul Speir, and Christoph H. Borchers

13. Capillary Isoelectric Focusing/Reversed Phase Liquid Chromatography/Mass Spectrometry............ 233
Cheng S. Lee and Brian M. Balgley

14. Integrating Accelerated Tryptic Digestion into Proteomics Workflows 241
Gordon W. Slysz and David C. Schriemer

15. Hydrogen/Deuterium Exchange Mass Spectrometry 255
Xuguang Yan and Claudia S. Maier

16. Mass Spectrometry Detection and Characterization of Noncovalent Protein Complexes 273
Sheng Yin and Joseph A. Loos

17. Chemical Cross-Linking for Protein–Protein Interaction Studies.................. 283
Xiaoting Tang and James E. Bruce

18. Tissue Analysis with High-Resolution Imaging Mass Spectrometry 295
A.F. Maarten Altelaar and Ron M.A. Heeren

19. Proteomic Global Profiling for Cancer Biomarker Discovery........................ 309
Vitor Faca, Hong Wang , and Samir Hanash

20. Analysis of Protein Glycosylation and Phosphorylation Using Liquid Phase Separation, Protein Microarray Technology, and Mass Spectrometry............ 321
Jia Zhao, Tasneem H. Patwa, Manoj Pal, Weilian Qiu, and David M. Lubman

21. Transthyretin Mass Determination for Detection of Transthyretin Familial Amyloid............ 353
John F. O'Brien and H. Robert Bergen III

22. Characterization of Microorganisms by MALDI Mass Spectrometry 367
Catherine E. Petersen, Nancy B. Valentine, and Karen L. Wahl

23. Mass Spectrometric Characterization of Neuropeptides............ 381
Stephanie S. DeKeyser, James A. Dowell, and Lingjun Li

24. Peptide and Protein Ion/Ion Reactions in Electrodynamic Ion Traps: Tools and Methods............ 395
Scott A. McLuckey

25. Electron Capture Dissociation LC/MS/MS for Bottom–Up Proteomics 413
Roman A. Zubarev

26. Two-Dimensional Ion Mobility Analyses of Proteins and Peptides 417
Alexandre A. Shvartsburg, Keqi Tang, and Richard D. Smith

27. Proteomics for Validation of Automated Gene Model Predictions 447
Kemin Zhou, Ellen A. Panisko, Jon K. Magnuson, Scott E. Baker, and Igor V. Grigoriev

28. Support Vector Machines for Improved Peptide Identification from Tandem Mass Spectrometry Database Search 453
Bobbie-Jo M. Webb-Robertson

Index 461

Contributors

Renate Ackermann • *Department of Plant Sciences, Texas Agricultural Experiment Station, Texas A&M University, College Station, TX, USA*

A.F. Maarten Altelaar • *FOM Institute for Atomic and Molecular Physics FOM-AMOLF, Amsterdam, The Netherlands*

Scott E. Baker • *Fungal Biotechnology, Pacific Northwest National Laboratory, Richland, WA, USA*

Brian M. Balgleys • *Calibrant Biosystems, Gaithersburg, MD, USA*

H. Robert Bergen III • *Department of Biochemistry and Molecular Biology, Mayo Proteomics Research Center, Mayo Clinic of Medicine, Rochester, MN, USA*

Birgit Bisle • *Department of Membrane Biochemistry, MPI of Biochemistry, Martinsried, Germany*

Christoph H. Borchers • *Department of Biochemistry and Microbiology, Genome British Columbia Proteomics Centre, University of Victoria, Victoria, BC, Canada*

James E. Bruce • *Department of Chemistry, Washington State University, Pullman, WA, USA*

Stephanie S. DeKeyser • *Department of Chemistry, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA*

James A. Dowell • *Department of Chemistry, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA*

Michael Easterling • *Bruker Daltonics, Billerica, MA, USA*

Vitor Faca • *Fred Hutchinson Cancer Research Center, Seattle, WA, USA*

Laurence Florens • *Stowers Institute for Medical Research, Kansas City, MO, USA*

Craig Gelfand • *Becton Dickinson, Franklin Lakes, NJ, USA*

Igor V. Grigoriev • *US DOE Joint Genome Institute, Walnut Creek, CA, USA*

Samir Hanash • *Fred Hutchinson Cancer Research Center, Seattle, WA, USA*

Ron M.A. Heeren • *FOM Institute for Atomic and Molecular Physics FOM-AMOLF, Amsterdam, The Netherlands*

Stefanie Henning • *Institute for Medical Physics and Biophysics, University of Münster, Münster, Germany*

Kim K. Hixson • *Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA*

Christian G. Huber • *Instrumental Analysis and Bioanalysis, Saarland University, Saarbrücken, Germany*

Peter R. Jungblut • *Max Planck Institute for Infection Biology, Berlin, Germany*

Thomas Kislinger • *ROntario Cancer Institute, University Health Network, Toronto, ON, Canada Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada*

CHENG S. LEE • *Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, USA*

LINGJUN LI • *Department of Chemistry, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA*

JOSEPH A. LOO • *Department of Chemistry and Biochemistry and Biological Chemistry, Molecular Biology Institute, University of California, Los Angeles, CA, USA*

DAVID M. LUBMAN • *Department of Chemistry, Comprehensive Cancer Center, Department of Surgery, University of Michigan Medical Center, Ann Arbor, MI, USA*

BORIS MAČEK • *Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany*

MILAN MADERA • *Department of Chemistry, Indiana University, Bloomington, IN, USA*

JON K. MAGNUSON • *Fungal Biotechnology, Pacific Northwest National Laboratory, Richland, WA, USA*

CLAUDIA S. MAIER • *Department of Chemistry, Oregon State University, Corvallis, OR, USA*

SCOTT A. MCLUCKEY • *Department of Chemistry, Purdue University, West Lafayette, IN, USA*

YEHIA MECHREF • *Department of Chemistry, Indiana University, Bloomington, IN, USA*

CHRISTIAN NEUSÜSS • *Department of Chemistry, University of Aalen, Aalen, Germany*

AMY M. NICHOLS • *Biological Engineering Department, Massachusetts Institute of Technology, Cambridge, MA, USA*

MILOS V. NOVOTNY • *Department of Chemistry, Indiana University, Bloomington, IN, USA*

JOHN F. O'BRIEN • *Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, MN, USA*

JESPER V. OLSEN • *Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany*

SEVERINE A. OUVRY-PATAT • *Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC, USA*

MANOJ PAL • *Department of Chemistry, University of Michigan Medical Center, Ann Arbor, MI, USA*

ELLEN A. PANISKO • *Fungal Biotechnology, Pacific Northwest National Laboratory, Richland, WA, USA*

TASNEEM H. PATWA • *Department of Chemistry, University of Michigan Medical Center, Ann Arbor, MI, USA*

MATTHIAS PELZING • *Bruker Daltonics, Melbourne, Australia*

JASNA PETER-KATALINIĆ • *Institute for Medical Physics and Biophysics, University of Münster, Münster, Germany*

CATHERINE E. PETERSEN • *Chemical and Biological Sciences, Pacific Northwest National Laboratory, Richland, WA, USA*

GOTTFRIED POHLENTZ • *Institute for Medical Physics and Biophysics, University of Münster, Münster, Germany*

WEILIAN QIU • *Department of Chemistry, University of Michigan Medical Center, Ann Arbor, MI, USA*

HUNG-HIANG QUEK • *Temasek Polytechnic, Singapore*

RAJAT SAPRA • *Biosystems Research Department, Sandia National Laboratories, Livermore, CA, USA*

MONIKA SCHMID • *Max Planck Institute for Infection Biology, Berlin, Germany*

ALEXANDER SCHMIDT • *Institute for Molecular Systems Biology, ETH, Zürich, Switzerland*

DAVID C. SCHRIEMER • *Department of Chemistry and Molecular Biology, University of Calgary, Calgary, AB, Canada*

ALEXANDRE A. SHVARTSBURG • *Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA*

GORDON W. SLYSZ • *Department of Chemistry and Molecular Biology, University of Calgary, Calgary, AB, Canada*

RICHARD D. SMITH • *Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA*

PAUL J. SPEIR • *Bruker Daltonics, Billerica, MA, USA*

SELENE K. SWANSON • *Stowers Institute for Medical Research, Kansas City, MO, USA*

KEQI TANG • *Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA*

XIAOTING TANG • *Department of Chemistry, Washington State University, Pullman, WA, USA*

ANNA-MARIA TIMPERIO • *Instrumental Analysis and Bioanalysis, Saarland University, Saarbrücken, Germany*

HANSJÖRG TOLL • *Instrumental Analysis and Bioanalysis, Saarland University, Saarbrücken, Germany*

MATTHEW P. TORRES • *Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC, USA*

NANCY B. VALENTINE •*Chemical and Biological Sciences, Pacific Northwest National Laboratory, Richland, WA, USA*

KAREN L. WAHL • *Chemical and Biological Sciences, Pacific Northwest National Laboratory, Richland, WA, USA*

HONG WANG • *Fred Hutchinson Cancer Research Center, Seattle, WA, USA*

MICHAEL P. WASHBURN •*Stowers Institute for Medical Research, Kansas City, MO, USA*

BOBBIE-JO M. WEBB-ROBERTSON • *Computational Biology and Bioinformatics, Pacific Northwest National Laboratory, Richland, WA, USA*

FOREST M. WHITE • *Biological Engineering Department, Massachusetts Institute of Technology, Cambridge, MA, USA*

XUGUANG YAN • *Department of Chemistry, Oregon State University, Corvallis, OR, USA*

SHENG YIN • *Department of Chemistry and Biochemistry, University of California, Los Angeles, CA, USA*

JIA ZHAO • *Department of Chemistry, University of Michigan Medical Center, Ann Arbor, MI, USA*

KEMIN ZHOU • *US DOE Joint Genome Institute, Walnut Creek, CA, USA*

URSULA ZIMNY-ARNDT • *Max Planck Institute for Infection Biology, Berlin, Germany*

LELLO ZOLLA • *Instrumental Analysis and Bioanalysis, Saarland University, Saarbrücken, Germany*

ROMAN A. ZUBAREV • *Laboratory for Biological and Medical Mass Spectrometry, Uppsala Biomedical Centrum, Uppsala, Sweden*

Chapter 1

Generation and Analysis of Multidimensional Protein Identification Technology Datasets

Selene K. Swanson, Laurence Florens, and Michael P. Washburn

Summary

Systems that couple two dimensional liquid chromatography (LC/LC) with tandem mass spectrometry are widely used in modern proteomics. One such system, multidimensional protein identification technology (MudPIT), couples strong cation exchange chromatography and reversed phase chromatography to tandem mass spectrometry in a single microcapillary column. Using database searching algorithms like SEQUEST and additional computational tools, researchers are able to analyze in great detail complex peptide mixtures generated from biofluids, tissues, cells, organelles, or protein complexes. This chapter describes the use of MudPIT on modern mass spectrometry instrumentation and describes a data analysis pipeline designed to provide low false positive rates and quantitative datasets.

Key words: Mass spectrometry, Multidimensional protein identification technology, Proteomics, Database searching, Normalized spectral abundance factor.

1. Introduction

With the large dynamic range of proteomes, it is important to be able to separate protein components of a proteome for analysis and identification. One approach for doing so is the bottom up approach where proteins are first digested into peptides prior to separation and analysis by mass spectrometry. A major challenge with the bottom up approach is that the complexity of the mixture increases greatly when a complex protein mixture is digested into peptides prior to the analysis. To separate such complex mixtures multidimensional separations must be used prior to mass spectrometry analysis. Numerous two dimensional (2-D) chromatographic methodologies have been developed in order to resolve peptide mixtures prior to analysis by

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_1

mass spectrometry (reviewed in *(1)*. In one such approach, strong cation exchange chromatography (SCX) is coupled to reversed phase chromatography (RP) for peptide separation prior to tandem mass spectrometry analysis. In this area, multidimensional protein identification technology (MudPIT), uses a fully integrated SCX/RP/MS/MS approach where a bi- or triphasic microcapillary column packed with SCX and RP HPLC grade materials and loaded with a complex peptide mixture generated from a biological sample *(2–5)*. Peptides are directly eluted off of the microcapillary column, ionized, and analyzed in the tandem mass spectrometer, which is capable of fragmenting peptides in a predictable fashion and allows for the computational determination of the peptide sequence by database searching algorithms like SEQUEST *(6)*. MudPIT analyses of whole proteomes *(7–9)*, differential protein expression in response to variable growth conditions *(10, 11)*, membrane proteins *(11–13)* and large multi-protein complexes *(14–16)* have been described demonstrating the diverse capabilities of the approach. The current review of methods will describe the approaches needed to carry out an analysis of a complex protein mixture using MudPIT.

2. Materials

2.1. Sample Preparation

1. Benzonase (Sigma, St. Louis, MO).
2. Trichloroacetic Acid (TCA).
3. Digestion buffer, 100 mM Tris–HCl, pH 8.5 (stored at 4°C), 8 M urea (added fresh).
4. Tris(2-carboxylethyl)-phosphine hydrochloride (TCEP) (1 M stock, stored at –20°C, diluted 1/10) (Pierce, Rockford, IL).
5. Iodoacetamide (Sigma, St. Louis, MO).
6. Endoproteinase LysC (Roche Applied Science, Indianapolis, IN).
7. Calcium Chloride, 500 mM stock.
8. Trypsin, modified sequencing grade, 0.1 µg/µL stock in water at –20°C (Roche Applied Science, Indianapolis, IN).
9. 90% Formic acid.
10. Elastase stock solution, 10 µg/mL (in 10 mM Tris–HCl, pH 8.5) stored at –20°C (Calbiochem, San Diego, CA).
11. Subtilisin A stock solution, 10 µg/µL (in 10 mM Tris–HCl, pH 8.5), stored at –20°C (Calbiochem, San Diego, CA).
12. Proteinase K stock solution, 15 µg/µL in water, stored at –20°C (Roche Applied Science, Indianapolis, IN).

2.2. Microcapillary Column Construction and Sample Loading

1. 100 μm i.d. × 365 μm o.d. and 250 μm i.d. × 365 μm o.d. polyimide coated fused silica (Polymicro Technologies, Phoenix, AZ).
2. Model P-2000 Laser Puller (Sutter Instrument Co. Novato, CA).
3. Stainless Steel Pressurization Device (Brechbuehler, Inc., Houston, TX, or MTA for blueprints available by request from John Yates, Scripps Research Institute, La Jolla, CA).
4. 5 μm C_{18} Aqua Reversed Phase Packing Material (Phenomenex, Torrance, CA).
5. 5 μm Partisphere Strong Cation Exchange packing material (Whatman, Florham Park, NJ).
6. M-520 Inline Micro Filter Assembly and F-185 Microtight 0.0155 × 0.025 Sleeves (UpChurch Scientific, Oak Harbor, WA).

2.3. Multidimensional Chromatography and Tandem Mass Spectrometry

1. Buffer A, 5% acetonitrile, 0.1% formic acid made with HPLC grade water.
2. Buffer B, 80% acetonitrile, 0.1% formic acid made with HPLC grade water.
3. Buffer C, 500 mM ammonium acetate, 5% acetonitrile, 0.1% formic acid made with HPLC grade water.
4. 50 μm i.d. × 365 μm o.d. polyimide coated fused silica (Polymicro Technologies, Phoenix, AZ).
5. P-775 MicroTee Assemblies and F-185 Microtight 0.0155 × 0.025 Sleeves (UpChurch Scientific, Oak Harbor, WA).
6. Gold wire (Scientific Instrument Services, Inc, Ringoes, NJ).
7. Agilent1100 series G1379A degasser, G1311A quaternary pump, G1329A autosampler, G1330B autosampler thermostat, and G1323B controller (Agilent Technologies, Palo Alto, CA).
8. LTQ tandem mass spectrometer (Thermo Electron, San Jose, CA).
9. Nano electrospray stage (MTA for blueprints available by request from John Yates, Scripps Research Institute, La Jolla, CA). Other options include the Thermo Electron Nanospray II ion source or PicoView Sources from New Objective (Woburn, MA).

3. Methods

3.1. Sample Preparation

The total amount of protein in a starting sample, from a multiprotein complex or whole proteome, is an important piece of information in order to add the proper amount of proteases for digestion. Samples from a variety of biological sources will be analyzed in the same way except for the early step of TCA precipitation. If a large amount of protein is in the starting sample, it should be split into 200–300 μg aliquots and each fraction of the TCA should be precipitated independently. In addition, if the sample is in a large volume, the sample should be split into 100 μL aliquots and each fraction of the TCA should be precipitated independently. After the resuspension step, samples may be recombined for protease digestion that is adjusted to the total protein content. The standard approach for the detection and identification of proteins is the Endoproteinase LysC/Trypsin digestion protocol. For the identification of post-translational modifications, high sequence coverage of individual proteins is needed. In this case, samples will be split into at least three fractions and independently digested with Endoproteinase LysC/Tyrpsin, Elastase, and Subtilisin in a similar fashion to the protocol published by MacCoss et al. *(17)*. In addition, the high pH proteinase K digestion protocol can also be used to generate high sequence coverage data from complex protein mixtures for post-translational modification analysis *(18)*. In each case, when using specific or non specific proteases, the false discovery rate (FDR) of the peptide needs to be carefully evaluated and selection criteria needs to be adjusted to have a FDR of 2% or less *(19–21)*.

3.1.1. TCA Precipitation of Proteins from Solutions (see Notes 1–4)

1. Bring solution to 400 μL with 100 mM Tris–HCl, pH 8.5.
2. Add 100 μL TCA (100%) to cold sample, mix well to give a final TCA concentration of 20%. Reaction should be carried out on ice and the sample left overnight at 4°C.
3. Spin at 14,000 rpm (17,500 × g) for 30 min at 4°C; aspirate the supernatant with gel loading tip, leaving 5–10 μL in the tube so as not to disturb the pellet.
4. Wash with 2 × L 500 μL of cold acetone. After each wash spin for 10 min at 14,000 rpm (17,500 × g).
5. Dry using a speed vac for 5 min.

3.1.2. Protein Denaturation, Reduction, and Alkylation

1. Add 100 mM Tris–HCl, pH 8.5, 8 M urea to TCA precipitated proteins: vortex.
2. Bring solution to 5 mM TCEP with 0.1 M (1 M stock diluted 1/10) incubate at room temperature for 30 min.
3. Bring solution to 10 mM IAM with 0.5 M; incubate at room temperature for 30 min in dark.

3.1.3. Endoproteinase LysC/Trypsin Digestion

1. Add Endoproteinase LysC at 1 μg/μL (1:100) to denatured, reduced, and carboxymethylated proteins; incubate at 37°C for at least 6 h.

2. Dilute to 2 M urea with 100 mM Tris–HCl, pH 8.5.
3. Add $CaCl_2$ to 2 mM (stock at 500 mM).
4. Add Trypsin at 0.1 µg/µL (1:100): incubate at 37°C overnight while shaking.
5. On the next day add 90% formic acid to 5% (final concentration).
6. Store sample at −80°C.

3.1.4 Elastase Digestion

1. Dilute denatured, reduced, and carboxymethylated to 2 M urea with 100 mM Tris–HCl, pH 8.5.
2. Add Elastase to an enzyme-to-substrate ratio of 1:50 (wt/wt); incubate at 37°C, for 6 h while shaking.
3. Add 90% formic acid to 5%.
4. Store sample at −80°C.

3.1.5 Subtilisin A Digestion

1. Dilute denatured, reduced, and carboxymethylated to 4 M urea with 100 mM Tris–HCl, pH 8.5.
2. Add Subtilisin A to an enzyme to substrate ratio of 1:50 (wt/wt); incubate at 37°C for 2–3 h while shaking.
3. Add 90% formic acid to 5% (final concentration).
4. Store sample at −80°C.

3.1.6 High pH Proteinase K Digestion

1. Resuspend TCA-precipitated proteins in 100 mM sodium carbonate, pH 11.5.
2. Add solid urea to 8 M urea, vortex.
3. Bring solution to 5 mM TCEP with 0.1 M (1 M stock diluted 1/10); incubate at room temperature for 30 min.
4. Bring solution to 10 mM IAM with 0.5 M stock; incubate at room temperature for 30 min in dark
5. Add Proteinase K at 0.25 µg/µL, at an enzyme to substrate ratio of 1:100 (wt/wt); incubate a 37°C for 4 h while shaking.
6. Add 90% formic acid to 5% (final concentration).
7. Store sample at −80°C.

3.2. Microcapillary Column Construction and Sample Loading

Currently all samples are desalted on-line using columns similar to the three phase microcapillary columns described in McDonald et al. *(5)*. These columns contain reversed phase material, followed by strong cation exchange material, followed by reversed phase material. In an abbreviated fashion, they are RP/SCX/RP columns. By using these columns one does not need to carry out additional sample clean up and buffer exchange prior to loading. For sample quantities of 400 µg or less, the Triple-phase Fused-silica microcapillary column is used and for samples containing more than 400 µg the Split Triple-phase Fused-silica microcapillary column is used. In addition, the Split Triple-phase column is used for samples that originally contained detergents, allowing for more extensive washing after sample loading.

3.2.1. Pulling Columns

1. Make a window in the center of ~50 cm of 100 µm × 365 µm fused silica capillary by holding it over an alcohol flame until the polyimide coating has been charred. The charred material is removed by gently wiping the capillary with a tissue soaked in methanol.
2. To pull a needle, place the capillary into the P-2000 laser puller. Position the exposed window of the capillary in the mirrored chamber of the puller. Arms on each side of the mirror have grooves and small vises, which properly align the fused silica and hold it in place. Our four-step parameter setup for pulling ~3 µm tips from a 100 µm i.d. × 365 µm o.d. capillary is as follows with all other values set to zero:

Heat = 290, Velocity = 40, and Delay = 200
Heat = 280, Velocity = 30, and Delay = 200
Heat = 270, Velocity = 25, and Delay = 200
Heat = 260, Velocity = 20, and Delay = 200

3.2.2. Triple-Phase Fused-Silica Microcapillary Column

1. Pull tip with laser puller as described in **Subheading 3.2.1**.
2. Place ~20 mg of Aqua Reversed Phase (RP) packing material into a 1.7-mL microfuge tube, add 1 mL of MeOH, and place the tube into a stainless steel pressurization vessel. Secure the pressurization vessel lid by tightening the bolts.
3. The lid has a Swagelok® fitting containing a 0.4 mm Teflon ferrule. Feed the fused silica capillary (pulled end up) down through the ferrule until the end of the capillary reaches the bottom of the microfuge tube. Tighten the ferrule to secure the capillary.
4. Apply pressure to the pressurization vessel by first setting the regulator on the gas cylinder to ~400–800 psi, then open a valve on the pressurization vessel to pressurize it. The packing material will begin filling the pulled needle capillary. If it does not flow right away, gently open the tip using a capillary scriber angled at 45°. Pack the capillary with 9 cm of reversed phase packing material (*see* **Note 5**).
5. Slowly release the pressure from the pressurization vessel so as to not cause the packed RP material to unpack. Open the stainless steel pressurization vessel and remove the microfuge tube containing the RP material in MeOH.
6. Place ~20 mg of Whatman Paritshpere SCX packing material into a 1.7-mL microfuge tube, add ~1 mL of MeOH, and place the tube into the pressurization vessel. Secure the pressurization vessel as described.

7. Apply pressure as described in Subheading 3.2.2, **step 4**. Pack the capillary with 3 cm of the SCX material and then slowly release the pressure as described.
8. Place the tube with Aqua RP particles in methanol back into the pressurization vessel and add 2–3 cm of RP material after the SCX material.
9. Wash with Methanol for at least 10 min.
10. Equilibrate with Buffer A (5%ACN, 0.1% formic acid) for at least 30 min.
11. To get rid of any particulate (which could clog the microcapillary column), spin down the sample to be loaded for 30 min at 14,000 rpm (17,500 × g) and transfer the supernatant to a new 1.7-mL microfuge tube.
12. Sample can then be loaded by placing the 1.7-mL microfuge tube into the pressurization vessel.
13. After the sample is loaded, 500 μL of Buffer A is added to the microfuge tube that contained the sample and the loaded column is washed until installed onto the mass spectrometer (at least 1 h).

3.2.3. Split Triple-Phase Fused-Silica Microcapillary Column

1. The first step in this process is to prepare a double-phase column out of 250 μm fused silica microcapillary. Place approximately 2× 15 cm of 250 μm i.d. × 365 μm o.d. fused silica capillaries on both sides of a M-520 Inline Micro Filter Assembly.
2. As described in Subheading 3.2.2, first pack the column with 3–4 cm of Whatman Partisphere SCX packing material by leading the 250-μm capillary connected to the "fritted" side of the inline micro-filter assembly into the pressurization vessel. The capillary connected to the other end of the inline assembly serves as a waste line flowing into an empty microfuge tube.
3. Then pack with 2–3 cm of Phenomenex Aqua RP packing material.
4. Wash with methanol for at least 10 min.
5. Equilibrate in Buffer A for at least 30 min.
6. Sample can then be loaded by placing the 1.7-mL microfuge tube into the pressurization vessel.
7. After sample is loaded, 1.5 mL of Buffer A is added to the microfuge tube that contained the sample and the loaded column is washed until installed onto the mass spectrometer. Because the split three phase columns are used for larger sample amounts and for samples that contained detergent, longer washing is necessary. Ideally, all 1.5 mL of Buffer A is used for washing.

8. Next, a single-phase RP column, to be added to the other end of the filtered union assembly, is prepared by first pulling a tip using 100 µm i.d. × 365 µm o.d fused silica with the laser puller as described in **Subheading 3.2.1** and packing 9–10 cm of RP material.
9. This single-phase RP column is equilibrated with Buffer A (5%ACN, 0.1% formic acid) for at least 30 min.
10. When both parts of the split-column have been washed, the single-phase 100-µm RP column can be connected to the loaded 250 µm double-phase capillary using the filtered union.

3.3. Multidimensional Chromatography and Tandem Mass Spectrometry

The MudPIT system we use is a combination of the Agilent1100 quaternary pump stack with thermostatted autosampler, LTQ tandem mass spectrometer, and a nanoelectrospray stage that interfaces the two systems. Furthermore, the hand-made microcapillary columns described above are single-use, and the fused silica portion of each column is discarded after every analysis (*see* **Note 6**).

3.3.1. Setup of the Nanoelectrospray Stage

1. A detailed schematic of the nanoelectrospray stage is shown in **Fig. 1** of Florens and Washburn *(22)*. The triphasic microcapillary column or split three-phase column should be attached as shown to the P-775 MicroTee Assemblies.
2. One connection point of the cross contains the transfer line from the HPLC pump. This consists of a piece of 100 µm i.d. × 365 µm o.d. polyimide coated fused silica.
3. A second connection point contains a length of 50 µm i.d. × 365 µm o.d. fused silica capillary that is used as a split/waste line. This split line allows a majority of the flow to exit through the split; therefore, very low flow-rates can be achieved through the packed capillary micro-column. The size and length of this section of capillary depend on the flow-rate from the pump and the length of the micro-column. A good starting point is to use a 12-in. section of 50 µm i.d. × 365 µm o.d. polyimide coated fused silica for the split line.
4. Another connection contains a section of gold wire. This is to allow the solvent entering the needle to be energized to 2,400 V, thus allowing electrospray ionization to occur.
5. Place the packed, loaded, and washed column into a MicroTee on a stage, which in this case is designed for the LTQ mass spectrometer. This stage performs a threefold purpose: to support the MicroTees and hold it in place along with the connections, to electrically insulate the MicroTees from contact with its surroundings when it is held at high voltage

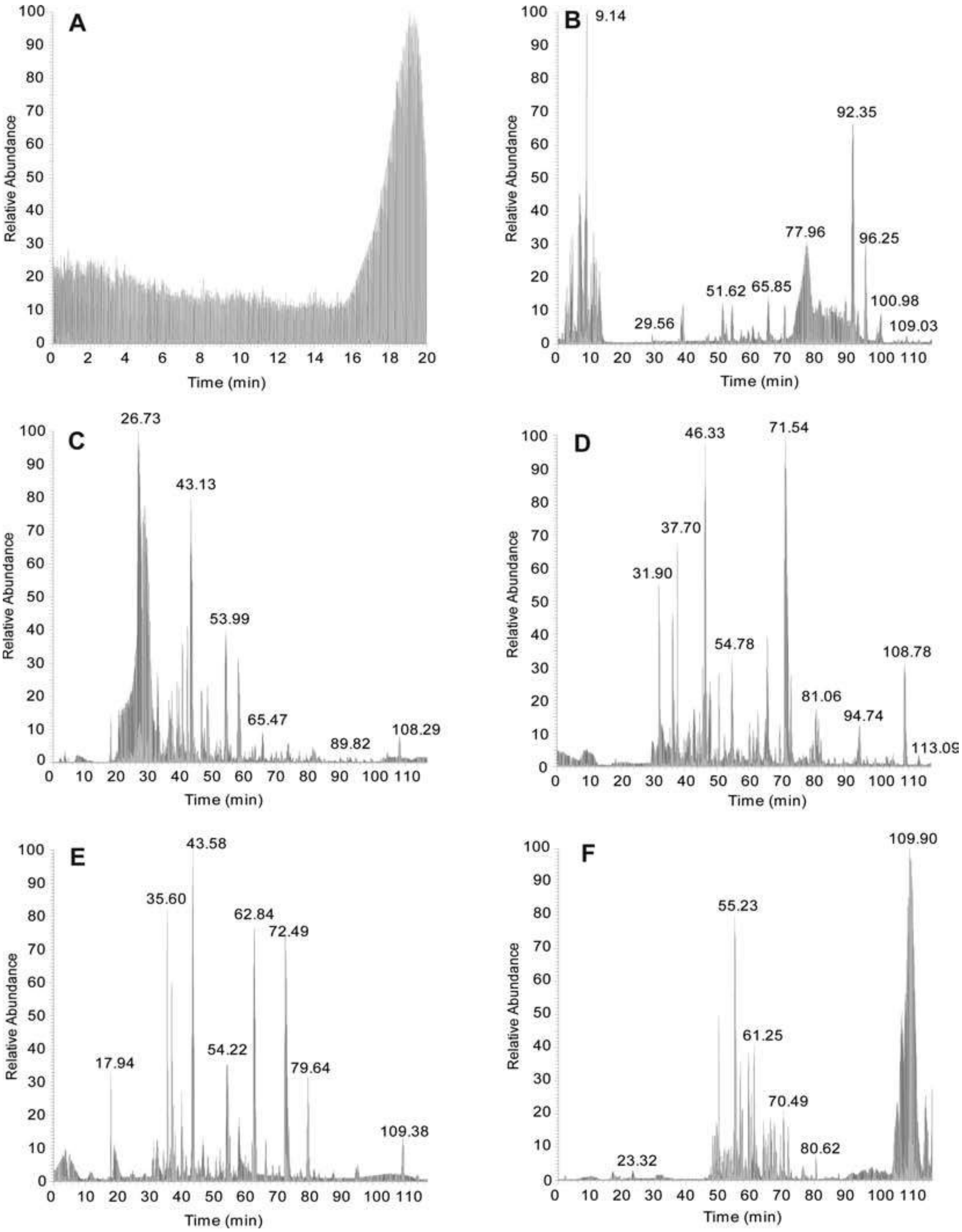


Fig. 1. Chromatographic Profile of a six-step MudPIT Analysis of a Mediator Sample. Final elutions of HeLa cell Mediator preparations were pooled and split into multiple samples for routine monitoring of MudPIT analysis performance. The chromatographic profile of a six-step MudPIT analysis on an LTQ is shown. (**a**) The base peak chromatogram (*m/z* 400–1,400) of the 20 min reversed phase gradient of the first step. (**b**) The base peak chromatogram (*m/z* 400–1,400) of the second step using 15% Buffer C. (**c**) The base peak chromatogram (*m/z* 400–1,400) of the third step using 30% Buffer C. (**d**) The base peak chromatogram (*m/z* 400–1,400) of the fourth step using 50% Buffer C. (**e**) The base peak chromatogram (*m/z* 400–1,400) of the fifth step using 70% Buffer C. (**f**) The base peak chromatogram (*m/z* 400–1,400) of the last step using 100% Buffer C.

potential, and to allow for fine position adjustments of the micro-column with respect to the entrance of the mass spectrometer (heated capillary) by using an XYZ manipulator.

6. Measure the flow from the tip of the capillary micro-column, using graduated glass capillaries. To do this, set the flow rate of the Agilent1100 to 0.1 mL/min from the controller. The target flow rate at the tip should be ~200–300 nL/min and a back pressure on the Agilent1100 of between 30 and 50 bar. If the flow-rate is too fast, cut off a portion of the split line capillary. This will cause more of the flow to exit out of the split and cause less flow through the micro-column. If the flow is too slow, a longer piece of 50-µm capillary or a section with a smaller inner diameter can be used to force more flow through the micro-column. Measuring the flow-rate and adjusting the split line may have to be repeated a number of times until the target flow-rate is reached (*see* **Notes** 7 and 8).
7. Prior to initiating a run, position the micro-column using the XYZ manipulator so that the needle tip is within 5 mm from the orifice of the mass spectrometer's heated capillary.

3.3.2. Instrument Method Design Description (see *Note 9*)

Data dependent acquisition of tandem mass spectra during the HPLC gradient is also programmed through the LTQ Xcalibur™ software. Here we provide guidance for setting the parameters for data dependent acquisition using the LTQ series mass spectrometer. The following settings are for a typical data dependent MS/MS acquisition analysis. The method consists of a continual cycle beginning with one scan of MS (scan one), which records all of the *m/z* values of the ions present at that moment in the gradient, followed by five rounds of MS/MS. Full MS spectra are recorded on the peptides over a 400–1,600 *m/z* range. Dynamic exclusion is activated to improve the protein identification capacity during the analysis.

1. In the main Xcalibur™ software page, select "Instrumental Setup".
2. In the next window select the button labeled, "Data Dependent MS/MS".
3. The general settings for any given method is as follows: segments is set to 1, start delay is set to zero, duration is 117 min, number of scan events is 6, scan event details are as follows: first is MS, second is MS/MS of the most intense ion, third is MS/MS of second most intense ion, fourth is MS/MS of third most intense ion, fifth is MS/MS of fourth most intense ion, and sixth is MS/MS of fifth most intense ion.
4. For scan event one highlight the bar showing "Scan Event 1". Scan event one should contain the normal mass range, the

normal scan rate, the full scan type, positive polarity, centroid data type, a scan range of 400–1,600. In the MsN settings the isolation width is 1.0, the normalized collision energy 20.0, the activation $Q = 0.250$, and the activation time = 30.0 ms.

5. The "Tune Method" box specifies the path for a file containing the parameters for the electrostatic lenses and ion trap. These parameters are established through the "LTQ Tune" as described in the ThermoFinnigan LTQ "Getting Started" manual. The capillary temperature is 200°C and the electrospray voltage is 2.42 kV.
6. When the "Scan Event 2" box is highlighted the following parameters must be set. The mass range is normal, the scan rate is normal, the data type is centroid, and the dependent scan box is checked. Next check the "settings" box and enable the dynamic exclusion, set a repeat count of 2, a repeat duration of 30 s, the exclusion list size to 500, the exclusion duration of 300 s, and the exclusion mass width by mass in the low box to 0.8 and in the high box to 2.20. Scan event two is one the first most intense ion. Scan events 3–6 are the same as scan event two except that they are on the second most intense ion to the fifth most intense ion.
7. Under the "Contact Closure" window, check the contact closure box, set to two contact positions where the position at the start of the run is open and the contact position duration is 0.05 min.

3.3.3. Gradient Profiles for Complex Mixture Analysis

We have two general gradient profiles that we use. One is for analyzing protein complexes and one is for analyzing organelles or other cellular fractions. The method for analyzing protein complexes takes approximately 10 h and the second method takes approximately 20 h. Since we use three phase columns, the first step in any run is a reversed phase gradient to move any bound peptide from the first RP to the SCX material inside the column. Then, successive salt bumps are run to move small amounts of peptides from the SCX onto the last RP, followed by a slow reversed phase gradient to resolve peptides within the last RP before they are eluted off into the mass spectrometer. In the "Gradient Program" window of the instrument setup the following sequences are steps that can be used or adjusted by a researcher. A base peak chromatogram shown in **Fig. 1** is as an example of the profile that should be seen from the six steps of a successful MudPIT analysis. The sample in **Fig. 1** is a control sample that is a mixture of the third FLAG peptide elutions from various mammalian Mediator subunit pull downs *(15)*. This same sample will be used as the example in the data analysis section and **Figs. 2** and **3**.

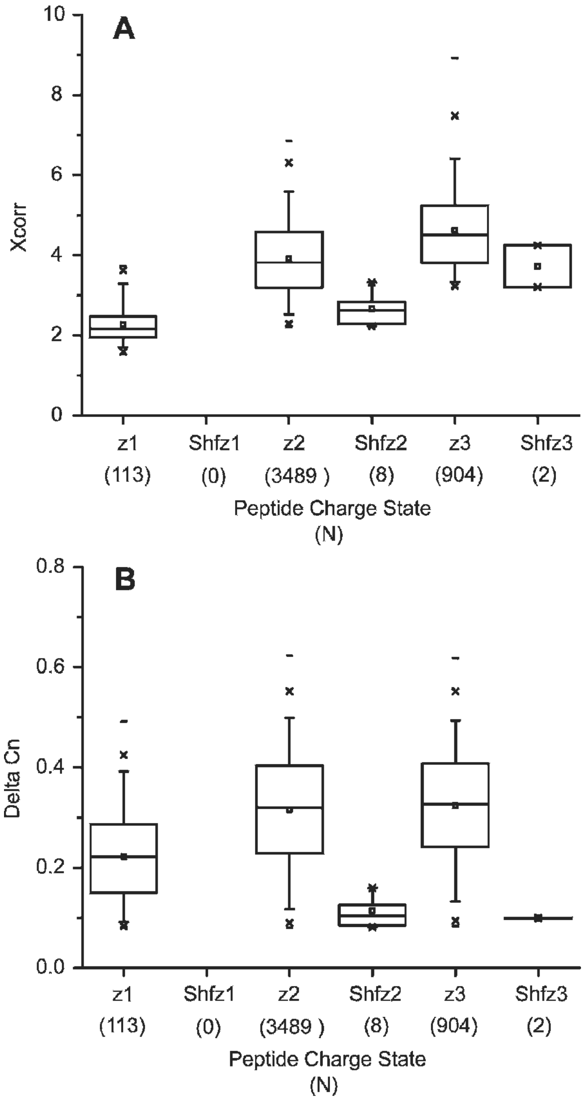


Fig. 2. Summary Percentile Statistics of Peptide Matches from SEQUEST. The Mediator control sample shown in **Fig.1** was searched using SEQUEST. A box plot summary statistic representation is shown for the charge state (z for real protein matches and Shf z for shuffled protein matches) dependent Xcorr (**a**) and DeltaCn (**b**) distributions of peptides. The SEQUEST parameters used were each peptide was fully tryptic, each peptide had to be at least seven amino acids in length, each peptide had to have a DeltaCn of 0.08, and +1 peptides had a minimum Xcorr of 1.5, +2 peptides had a minimum Xcorr of 2.2, and +3 peptides had a minimum Xcorr of 3.2. In a box plot, the 25th and 75th percentile are represented by the upper and lower boundaries of the box, with the median being the line dissecting the box, and the mean being the small square in each box. The 5th and 95th percentiles are shown with lines attached to the box, the "X" represents the 1st and 99th percentiles, and the stand alone "–" represents the complete range. The number (*N*) of total peptide identifications in each category is shown within brackets below each charge state.

1. Gradient Profile of First Step (**Table 1**). This is the first step of an analysis which is simply a reverse phase gradient to move peptides onto the SCX portion of the column.
2. Gradient Profile of ×(2–80)% Buffer C Step. This is the general gradient profile that we use for all additional steps where the % of Buffer C increases with each step (**Table 2**).

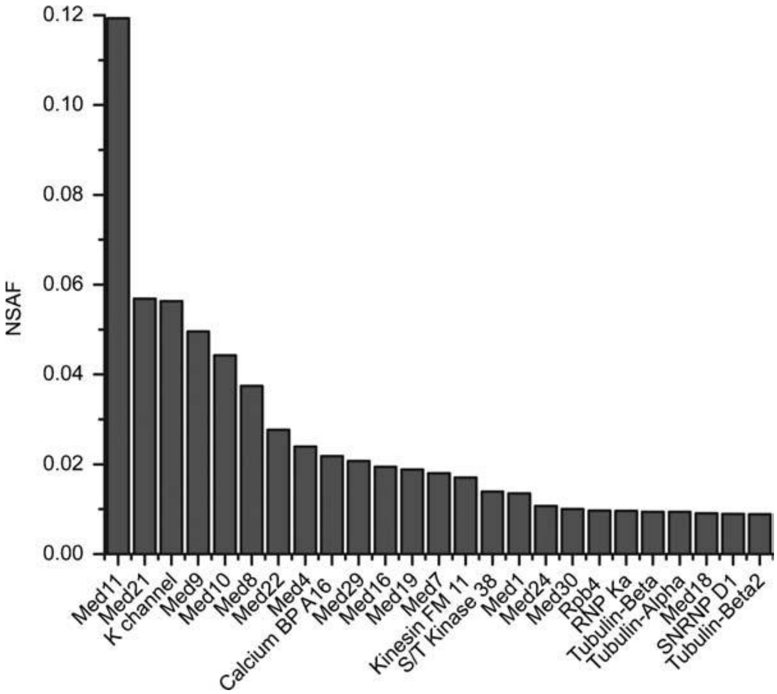


Fig. 3. Normalized Spectral Abundance Factors for the Top 25 Proteins. After SEQUEST analysis and peptide score analysis, a total of 400 nonredundant proteins, 4,516 unique peptides, and 21,825 spectral counts were obtained. Next, the Normalized Spectral Abundance Factor (NSAF) was calculated for each protein in the dataset. The NSAF value for the top 25 proteins is plotted as a function of protein name.

Table 1
Gradient profile of first step

Time (min)	Flow rate (mL/min)	% Buffer A	% Buffer B
0	0.1	100	0
16	0.1	60	40
17	0.1	0	100
20	0.1	0	100

3. Gradient Profile of 100% Buffer C Step. This is the last step for any given MudPIT analysis and is designed to try to remove everything off the column (Table 3).
4. MudPIT Analyses Methods. In the Xcalibur software main page check the sample list button. In the "Sequence Setup" window fill in the "Filename", "Path", and "Inst Method" lines with the appropriate methods made previously. Save the Sequence to the same directory of the paths for the *.RAW files. In the table below are the sequences we use for the analysis of a protein complex (six-step MudPIT) and for more complex mixtures like an organelle (12-step MudPIT). The setups for these methods are found in Table 4.

Table 2
Gradient profile of X(2–80)% Buffer C step

Time (min)	Flow rate (mL/min)	% Buffer A	% Buffer B	% Buffer C
0	0.1	100	0	0
3	0.1	100	0	0
3.1	0.1	98	0	X
5	0.1	98	0	X
5.1	0.1	100	0	0
10	0.1	100	0	0
10.1	0.1	100	0	0
25	0.1	85	15	0
117	0.1	55	45	0

Table 3
Gradient profile of 100% Buffer C step

Time (min)	Flow rate (mL/min)	% Buffer A	% Buffer B	% Buffer C
0	0.1	100	0	0
2	0.1	100	0	0
2.1	0.1	0	0	100
22	0.1	0	0	100
22.1	0.1	100	0	0
27	0.1	100	0	0
37	0.1	80	20	0
85	0.1	30	70	0
90	0.1	0	100	0
90.1	0.1	0	100	0
95	0.1	0	100	0
95.1	0.1	100	0	0
97	0.1	100	0	0

Table 4
MudPIT analyses methods

Method	6-Step MudPIT	12-Step MudPIT
20 min reverse phase	Yes	Yes
2% Buffer C		Yes
6% Buffer C		Yes
10% Buffer C		Yes
15% Buffer C	Yes	Yes
20% Buffer C		Yes
30% Buffer C	Yes	Yes
40% Buffer C		Yes
50% Buffer C	Yes	Yes
60% Buffer C		Yes
70% Buffer C	Yes	
80% Buffer C		Yes
100% Buffer C	Yes	Yes

5. Once everything is setup and the needle is positioned correctly in front of the heated capillary on the mass spectrometer hit “Actions”–“Run Sequence”. A box will come up which should have the “Agilent1100 Quat,” “Agilent1100 Thermostatted AS,” and “LTQ” listed and “Yes” should be typed under the “Start Instrument” box of the “Agilent1100 Thermostatted AS.” The “start when ready” boxed should be checked as also the “After Sequence Set System to –Standby”. Both the “pre-acquisition” and “post-acquisition” boxes under “run synchronously” should contain check marks. Hit “O.K.” and the run will begin. A message is likely to come up saying that there are devices that need to be turned on, hit “O.K.”.
6. Upon the completion of a run, the *.RAW files which have been accumulated by the mass spectrometer need to be converted to *.DAT files for SEQUEST analysis. To do this go to the “Home Page” and hit the “Tools-File Converter” buttons. Follow the instructions to select the jobs which need to be converted to *.DAT files and hit o.k.

3.4 Data Analysis

For complex peptide mixture analysis via MudPIT, database searching is an important component and requires a computer cluster to analyze tandem mass spectra. Nowadays, SEQUEST *(6)* clusters are available from IBM and ThermoElectron (San Jose, CA) and Mascot *(23)* clusters are available from Matrix Science (Boston, MA). In addition, given the necessary expertise, researchers can design their own cluster and purchase licenses for these software platforms to install on their cluster. Lastly, an open source tandem mass spectral searching algorithm named X!Tandem *(24, 25)* is emerging as an alternative to SEQUEST *(6)* and Mascot *(23)*. An important aspect of data analysis includes the determination of the FDR *(19–21)*. To minimize the FDR and maximize the number of peptide and protein identifications, the scoring criteria and parameters used to filter a dataset are important and will probably need to be modified in each individual case to achieve this balance. In the section below the specific example used is the Mediator control sample whose chromatographic profile was shown in **Fig. 1** in **Subheading 3.3.3**.

1. Convert each dat file into an ms2 file *(26)* using *extract-ms* in order to obtain the coordinates of the MS/MS spectra to be analyzed.
2. Subject each ms2 file to the *2to3* software *(27)* to remove spectra of poor quality and assign a tentative charge state to precursor peptides.
3. Set up the sequest.params file so the peptide mass tolerance is 3; no enzyme specificity is required; parent ions are calculated with average masses, while fragment ions are modeled with monoisotopic masses; and cysteine residues are considered fully carboxyamidomethylated (+57 Da) and searched as a static modification.
4. For the Mediator control sample, search datasets with SEQUEST *(6)* against a database of 56,838 protein sequences, combining 28,242 *Homo sapiens* proteins (from the National Center for Biotechnology Information 2005-02-17 release), 177 common contaminants like keratin and immunoglobulins, and their corresponding shuffled *(28)* sequences (i.e., each *H. sapiens* and contaminant sequence was randomized maintaining the same amino acid composition and length). Clearly, the concatenated database, real and shuffled proteins, will vary depending on the species of the sample.
5. Run SEQUEST with no enzyme specificity because this approach has been demonstrated to be superior for minimizing false positive identifications *(21)*. The use of the shuffled sequences is to estimate FDR rates. For every spectrum matching a “shuffled” peptide, there should be one false positive in the “normal” dataset. The Peptide False

Discovery Rate (FDR) is calculated based on spectral count (SpC) as in *(19)*:

$$\mathrm{FDR} = \frac{(2 \times \text{"shuffled"\_SpC})}{\text{Total\_SpC}}$$

7. Use DTASelect *(29)* to parse the peptide information contained in the SEQUEST output files and assemble it into protein level information.
8. Use CONTRAST *(29)* and an in-house developed script, *contrast-report* to compare multiple MudPIT runs.
9. To be retained, MS/MS spectra had to match fully tryptic peptides of at least seven amino acid long, with a normalized difference in cross-correlation scores (DeltCn) of at least 0.08, and minimum cross-correlation scores (Xcorr) of 1.5 for singly charged spectra, 2.2 for doubly charged spectra, and 3.2 for triply charged spectra. The summary percentile statistics of peptide matches from SEQUEST using these criteria is shown in **Fig. 2** for the Mediator control sample analyzed in **Fig. 1**. These strict selection criteria led to a Spectra_FDR of 0.44% in this case indicating that the spectrum/peptide matched could be trusted at least 99.56% of the time.

3.5. Spectral count normalization

In recent years, spectral counts obtained from shotgun proteomic approaches have been shown to be a good estimation of protein abundance *(13, 30–32)*. To account for the fact that larger proteins tend to contribute more peptide/spectra, spectral counts were divided by protein length, defining a Spectral Abundance Factor (SAF) *(31)*. SAF values were then normalized against the sum of all SAFs for a particular run (removing redundant proteins) allowing us to compare protein levels across different runs. Using an in-house developed script (*contrast-report-add-nsaf*), for each protein *k* detected in a particular MudPIT analysis, Normalized Spectral Abundance Factors (NSAFs) were calculated as follow *(11)*:

$$(\mathrm{NSAF})_k = \frac{(\mathrm{SpC/length})_k}{\sum_{i=1}^{N} (\mathrm{SpC/length})_i}$$

NSAF values should range from zero to one, with values closer to one indicating higher protein levels. NSAFs were used to rank proteins within the analyses of the Mediator control sample used an example in this methods review. In total, 400 nonredundant proteins, 4,516 fully tryptic unique peptides, and 21,825 fully tryptic spectral counts were detected and identified using the Mediator control sample. Of those 400 nonredundant proteins, the NSAF value for the top 25 proteins is shown in **Fig. 3**.

4. Notes

1. For samples that probably contain large amounts of DNA, the addition of 0.1 units of Benzonase to the sample prior to TCA precipitation and incubation at 37°C for 30 min is critical. This will decrease the viscosity of the mixture and prevent column clogging during sample loading, washing, or analysis.
2. Using the TCA precipitation protocol as the starting point, the only chemical that cannot be dealt with is glycerol. The TCA precipitation protocol has been useful in detergent removal. When the presence of glycerol cannot be avoided, loading such samples on split-columns with extensive washes may prevent column clogging.
3. With the TCA precipitation protocol and protein complexes, it is common to not see a pellet. For this reason, a small amount of acetone should be left at the bottom of the tube during washing to be sure not to aspirate and discard the pellet.
4. When handling samples, gloves must be worn at all times in order to prevent the addition of large amounts of human keratin and immunoglobulins to a sample.
5. Placing a black 3-ring binder about 6 in. behind the loading column will provide the contrast necessary to see the status of column packing.
6. The microassemblies used for the split three phase columns can be reused by first sonicating in 50% methanol:50% water and allowing them to air dry. However, we discard the filters (M-120X filter end fitting from UpChurch Scientific, Oak Harbor, WA) and use a new one for each sample.
7. If the Agilent1100 reaches the high pressure limit (set at 100 bar), the most likely place for clogging is the small piece of 100 mm i.d. × 365 mm o.d. fused silica tubing which connects the two MicroTees (**Fig. 1** of Florens and Washburn *(22)*. If changing the capillary does not alleviate the pressure, removing the MicroTees and sonicating them in 50% methanol:50% water and allowing them to air dry may solve the problem.
8. For columns that clog during loading of the sample, washing of the sample, or during the flow rate test prior to analysis, heating the length of the column with a Varitemp Heat Gun

(model VT-750C from Master Appliance Corp, Racine, WI) usually makes it flow again. Once a column starts flowing again extensive additional washing with Buffer A is recommended to prevent additional clogging.

9. Especially with whole cell extracts and their analysis by MudPIT, it is critical to clean and tune the mass spectrometer according to the manufacturer's instructions at least every 10 days.

References

1. Evans, C.R., and Jorgenson, J.W. (2004) Multidimensional LC-LC and LC-CE for high-resolution separations of biological molecules. *Anal Bioanal Chem* 378, 1952–1961.
2. Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates, J.R., III (1999) Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 17, 676–682.
3. Washburn, M.P., Wolters, D., and Yates, J.R., III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19, 242–247.
4. Wolters, D.A., Washburn, M.P., and Yates, J.R., III (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 73, 5683–5690.
5. McDonald, W.H., Ohi, R., Miyamoto, D.T., Mitchison, T.J., and Yates, J.R. (2002) Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LCMS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Int J Mass Spectrom* 219, 245–251.
6. Eng, J., McCormack, A.L., and Yates, J.R., III (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Mass Spectrom* 5, 976–989.
7. Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., Witney, A.A., Wolters, D., Wu, Y., Gardner, M.J., Holder, A.A., Sinden, R.E., Yates, J.R., and Carucci, D.J. (2002) A proteomic view of the Plasmodium falciparum life cycle. *Nature* 419, 520–526.
8. Koller, A., Washburn, M.P., Lange, B.M., Andon, N.L., Deciu, C., Haynes, P.A., Hays, L., Schieltz, D., Ulaszek, R., Wei, J., Wolters, D., and Yates, J.R., III (2002) Proteomic survey of metabolic pathways in rice. *Proc Natl Acad Sci U S A* 99, 11969–11974.
9. Pan, Y., Kislinger, T., Gramolini, A.O., Zvaritch, E., Kranias, E.G., MacLennan, D.H., and Emili, A. (2004) Identification of biochemical adaptations in hyper- or hypocontractile hearts from phospholamban mutant mice by expression proteomics. *Proc Natl Acad Sci U S A* 101, 2241–2246.
10. Washburn, M.P., Koller, A., Oshiro, G., Ulaszek, R.R., Plouffe, D., Deciu, C., Winzeler, E., and Yates, J.R., III (2003) Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae. Proc Natl Acad Sci U S A* 100, 3107–3112.
11. Zybailov, B.L., Mosley, A.L., Sardiu, M.E., Coleman, M.K., Florens, L., and Washburn, M.P. (2006) Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J Proteome Res* 5, 2339–2347.
12. Wu, C.C., MacCoss, M.J., Howell, K.E., and YatesIII, J.R., (2003) A method for the comprehensive proteomic analysis of membrane proteins. *Nat Biotechnol* 21, 532–538.
13. Zybailov, B., Coleman, M.K., Florens, L., and Washburn, M.P. (2005) Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal Chem* 77, 6218–6224.
14. Graumann, J., Dunipace, L.A., Seol, J.H., McDonald, W.H., Yates, J.R., III, Wold, B.J., and Deshaies, R.J. (2004) Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast. *Mol Cell Proteomics* 3, 226–237.
15. Sato, S., Tomomori-Sato, C., Parmely, T.J., Florens, L., Zybailov, B., Swanson, S.K., Banks, C.A., Jin, J., Cai, Y., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2004) A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol Cell* 14, 685–691.

16. Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., and Workman, J.L. (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581–592.

17. MacCoss, M.J., McDonald, W.H., Saraf, A., Sadygov, R., Clark, J.M., Tasto, J.J., Gould, K.L., Wolters, D., Washburn, M., Weiss, A., Clark, J.I., and YatesIII, J.R., (2002) Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc Natl Acad Sci U S A* 99, 7900–7905.

18. Wu, C.C., MacCoss, M.J., Mardones, G., Finnigan, C., Mogelsvang, S., YatesIII, J.R., and Howell, K.E. (2004) Organellar proteomics reveals golgi arginine dimethylation. *Mol Biol Cell* 15, 2907–2919.

19. Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* 2, 43–50.

20. Elias, J.E., Haas, W., Faherty, B.K., and Gygi, S.P. (2005) Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat Methods* 2, 667–675.

21. Xie, H., and Griffin, T.J. (2006) Trade-off between high sensitivity and increased potential for false positive peptide sequence matches using a two-dimensional linear ion trap for tandem mass spectrometry-based proteomics. *J Proteome Res* 5, 1003–1009.

22. Florens, L., and Washburn, M.P. (2006) Proteomic analysis by multidimensional protein identification technology. *Methods Mol Biol* 328, 159–175.

23. Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567.

24. Craig, R., Cortens, J.P., and Beavis, R.C. (2004) Open source system for analyzing, validating, and storing protein identification data. *J Proteome Res* 3, 1234–1242.

25. Craig, R., and Beavis, R.C. (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20, 1466–1467.

26. McDonald, W.H., Tabb, D.L., Sadygov, R.G., MacCoss, M.J., Venable, J., Graumann, J., Johnson, J.R., Cociorva, D., and YatesIII, J.R., (2004) MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. *Rapid Commun Mass Spectrom* 18, 2162–2168.

27. Sadygov, R.G., Eng, J., Durr, E., Saraf, A., McDonald, H., MacCoss, M.J., and YatesIII, J.R., (2002) Code developments to improve the efficiency of automated MS/MS spectra interpretation. *J Proteome Res* 1, 211–215.

28. Higdon, R., Hogan, J.M., Van Belle, G., and Kolker, E. (2005) Randomized sequence databases for tandem mass spectrometry peptide and protein identification. *OMICS* 9, 364–379.

29. Tabb, D.L., McDonald, W.H., and YatesIII, J.R., (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J Proteome Res* 1, 21–26.

30. Liu, H., Sadygov, R.G., and YatesIII, J.R., (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76, 4193–4201.

31. Powell, D.W., Weaver, C.M., Jennings, J.L., McAfee, K.J., He, Y., Weil, P.A., and Link, A.J. (2004) Cluster analysis of mass spectrometry data reveals a novel component of SAGA. *Mol Cell Biol* 24, 7249–7259.

32. Old, W.M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K.G., Mendoza, A., Sevinsky, J.R., Resing, K.A., and Ahn, N.G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* 4, 1487–1502.

Chapter 2

Quantitative Peptide and Protein Profiling by Mass Spectrometry

Alexander Schmidt, Birgit Bisle, and Thomas Kislinger

Summary

Proteomics may be defined as the systematic analysis of proteins expressed in a given organism (Electrophoresis 16:1090–1094, 1995). Important technical innovations in mass spectrometry (MS), protein identification methods, and database annotation, over the past decade, now make it possible to routinely identify thousands of proteins in complex biological samples (Nature 422:198–207, 2003). However, to gain new insights regarding fundamental biological questions, accurate protein quantification is also required. In this chapter, we present methods for the biochemical separation of different cellular compartments, two-dimensional chromatographic separation of the constituent peptide populations, and the recently published *Spectral Counting Strategy*, a label-free MS-based protein quantification technology (Cell 125:173–186, 2006; Anal Chem 76:4193–4201, 2004; Mol Cell Proteomics 4:1487–1502, 2005; Cell 125:1003–1013, 2006; Methods 40:135–142, 2006; Anal Chem 77:6218–6224, 2005; J Proteome Res 5:2339–2347, 2006). Additionally, highly accurate protein quantification based on isotope dilution, describing the isotope coded protein label (ICPL) – method shall be explained in detail (Mol Cell Proteomics 5:1543–1558, 2006; Proteomics 5:4–15, 2005).

Key words: Proteomics, Profiling, Mass spectrometry, Quantification, Spectral counting, Isotope labeling, MudPIT, ICPL.

1. Introduction

Proteomics is the detection of proteins expressed in a given biological system (e.g., an organism, tissue, cell, organelle, or protein complex) *(1)*. With the availability of the genomic sequences of human and many eukaryotic and prokaryotic organisms, the goal of proteome research is the qualitative, quantitative and functional analysis of protein expression. The tool of choice for the detection of proteins in systems biology is the mass spectrometer (MS) *(2)*.

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_2

The many technical innovations of mass spectrometers in recent years have allowed scientists to rapidly and systematically detect hundreds to thousands of proteins in complex biological samples. Especially, the *Mu*lti*d*imensional *P*rotein *I*dentification *T*echnology (MudPIT), an elegant technology pioneered by the laboratory of Dr. John Yates III, has significantly increased to number of proteins detected by shot-gun proteomics *(3, 4)*. However, to fully understand the biological processes accurate quantification of proteins is required *(5)*. In 1999, Gygi and colleagues have developed a new approach for accurate protein quantification within complex mixtures using stable isotope labeling of proteins *(6)*. The method has shown to overcome several drawbacks of 2-DE based studies usually carried out for differential proteome analysis. Since then, the strategy has become increasingly popular and several groups have adopted the principle of this powerful methodology to generate additional strategies with their own strength and weaknesses. To date, three different ways of stable isotope labeling of proteins/peptides are utilized, that is chemically *(6–10)*, metabolically *(11, 12)* or enzymatically *(13)*. Since all methods are based on the same principle, we will describe the basic workflow and general challenges of protein quantification by stable isotope labeling on the basis of the recently developed isotope coded protein label (ICPL) approach *(7)*. ICPL is based on stable isotope tagging at the frequent free amino groups of isolated intact proteins and is therefore applicable to any protein sample, including membrane proteins, tissues extracts or body fluids. The following chapter we describe some of the recent developments in quantitative MS-based proteomics. The main focus will be on the detailed description of two alternative quantitative proteomics technologies (1) *Spectral counting* (SpC) in combination with MudPIT-based proteomics and (2) *ICPL* for quantitative proteomics with stable isotopes.

2. Materials

All materials used are of the highest quality. HPLC-grade solvents (water, methanol, acetonitrile) were from Fisher scientific. Proteomics grade enzymes (Endoproteinase Lys-C and trypsin) were obtained from Roche Diagnostics (Laval, QC, Canada). All standard proteins and *N*-(2-hydroxyethyl) piperazine-*N′*-(2-ethanesulfonic acid) (HEPES) were purchased from Sigma. All solid chemicals were from Fluka and of the highest purity available. $^{12}\text{C}_6$- (light ICPL reagent), $^{13}\text{C}_6$-nicotinoyl succinimide (heavy ICPL reagent) were purchased from Serva (Serva, Heidelberg, Germany)/Bruker (Bruker Daltonics Inc., MA, USA).

a-Cyano-4-hydroxycinnamic acid (HCCA) was obtained from Bruker (Bremen, Germany). Tris-(hydroxymethyl) aminomethane (Tris) was purchased from Bio-Rad (Bio-Rad, Munich, Germany) and trifluoroacetic acid (TFA) was obtained from Applied Biosystems (Framingham, MA, USA).

2.1. MudPIT Analysis

1. Fused silica, 100 mm inner diameter (Polymicron Technologies, Phoenix, AZ).
2. P-2000 Laser puller (Sutter Instruments, Novato, CA).
3. Reversed phase beads – ZORBAX Eclipse XDB-C18 5 mm (Agilent Technologies, Mississauga, ON, Canada).
4. Partisphere strong cation exchange resin (Whatman, Clifton, NJ).
5. Pressure vessel – made in-house or commercially available from Proxeon Biosystems (Odense, Denmark) or Brechbuehler (Houston, TX).
6. OMIX solid phase extraction cartridges (Varian, Mississauga, ON, Canada).

2.2. HPLC Buffers for MudPIT Analysis

1. Buffer A: 95% water/5% acetonitrile/0.1% formic acid.
2. Buffer B: 80% acetonitrile/20%water/0.1% formic acid.
3. Buffer C: 500 mM ammonium acetate in Buffer A.

2.3. Tissue Homogenization and Protein Extraction

1. *Tissue homogenization buffer*: 250 mM sucrose, 50 mM Tris–HCl, pH 7.4, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), and 1 mM phenylmethylsulfonyl fluoride (PMSF). DTT and PMSF are freshly added before every use.
2. *Sucrose cushion solution 1*: 0.9 M sucrose, 50 mM Tris–HCl, pH 7.4, 5 mM $MgCl_2$, 1 mM DTT, and 1 mM PMSF.
3. *Sucrose cushion solution 2*: 2.2 M sucrose, 50 mM Tris–HCl, pH 7.4, 5 mM $MgCl_2$, 1 mM DTT, and 1 mM PMSF.
4. *Nuclear extraction buffer 1*: 20 mM HEPES, pH 7.8, 1.5 mM $MgCl_2$, 450 mM NaCl, 0.2 mM EDTA, and 25% glycerol.
5. *Nuclear extraction buffer 2*: 20 mM HEPES, pH 7.8, 1.5 mM $MgCl_2$, 450 mM NaCl, 0.2 mM EDTA, 25% glycerol, and 1% Triton-X-100.
6. *Mitochondrial extraction buffer 1*: 10 mM HEPES, pH 7.8.
7. *Mitochondrial extraction buffer 2*: 20 mM HEPES, pH 7.8, 1.5 mM $MgCl_2$, 450 mM NaCl, 0.2 mM EDTA, 25% glycerol, and 1% Triton-X-100.

2.4. Protein Digestion and Preparation for MudPIT Analysis

1. Approximately 150 μg of total protein extract in extraction buffer.
2. Ice-cold biotechnology grade acetone.
3. 8M urea, 50 mM Tris–HCl, pH 8.5.

4. 50 mM ammonium carbonate, pH 8.5.
5. Stock solution of 100 mM $CaCl_2$.
6. Endoproteinase Lys-C (Roche Diagnostics).
7. Trypsin, recombinant, proteomics grade (Roche Diagnostics).
8. OMIX solid phase extraction cartridges (Varian, Mississauga, ON, Canada).

2.5. HPLC Buffers and MALDI-Matrix for ICPL Labeling

1. Loading buffer: 0.1% trifluoroacetic acid.
2. Buffer A: 0.05% trifluoroacetic acid.
3. Buffer B: 80% acetonitrile/20%water/0.04% trifluoroacetic acid.
4. Matrix buffer 1: 10 mg/mL HCCA in 50% acetonitril/0.1% trifluoroacetic acid.
5. Matrix buffer 2: 10 mM ammoniumdihydrogen phosphate in 50% acetonitril/0.1% trifluoroacetic acid.

2.6. Cell Culture and Membrane Protein Preparation

1. Complete medium (1 L): 250 g NaCl, 20 g $MgSO_4$, 3 g sodium citrate, 2 g KCI and 10 g Oxoid Bacteriological Peptone L 37 H (Colab Laboratories, Glenwood, IL).
2. Basal salt buffer: 4.3 M NaCl, 81 mM $MgSO_4$ and 27 mM KCl.
3. 10 and 60% sucrose solution.
4. 1 M and 0.5 M NaCl solution.
5. Chloroform.
6. Methanol.

2.7. Isotope Labeling

1. *Sample buffer*: 6 M guanidine HCl and 0.1 M HEPES, pH 8.5 ± 0.1*.
2. *Reduction buffer*: 0.2 M tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and 0.1 M HEPES, pH 8.5 ± 0.1*.
3. *Alkylation buffer*: 0.4 M iodacetamide and 0.1 M HEPES, pH 8.5 ± 0.1*.
4. *Alkylation stop buffer*: 0.5 M *N*-acetyl-cysteine and 0.1 M HEPES, pH 8.5 ± 0.1*.
5. *Labeling solution* (ICPL light): 0.15 M *N*-^{12}C-6-nicotinoyl-NHS in DMSO.
6. *Labeling solution* (ICPL heavy): 0.15 M *N*-^{13}C-6-nicotinoyl-NHS in DMSO.
7. *Stop solution*: 1.5 M hydroxylamine hydrochloride.

* Use 0.1 M NaOH or 0.1 M HCl for pH adjustment.

2.8. Protein Digestion and Sample Preparation for Mass Spectrometry

1. Approximately 100 μg of total labeled protein extract in labeling buffer.
2. 50 mM Tris–HCl, pH 8.5.
3. Trypsin, recombinant, proteomics grade (Roche Diagnostics).
4. Reprosil-Pur 120 ODS, C18 particles (Dr. Maisch GmBH, Germany, Cat#:r13.93).

3. Methods

Traditionally, two-dimensional gel electrophoresis (2-DE) followed by silver staining and MS identification of a separated gel spot, was the method of choice for proteome profiling. However, 2-DE is biased against the detection of membrane proteins, and proteins with extremes in molecular weight and isoelectric point. In recent years, several groups have presented gel-free approaches to overcome some of these limitations. Multidimensional protein identification technology (a.k.a. MudPIT) pioneered by the laboratory of John Yates III – allows for the systematic identification of hundreds to thousands of proteins in complex mixtures *(3, 4)*. However, the extreme complexity of the proteome of higher mammals (several 100,000 proteins including post-translational modifications), pushes modern proteomics to its limits. Even high resolution technologies such as MudPIT are thus not capable of identifying every protein present. Sample fractionation and minimization of sample complexity is therefore extremely important *(14)*.

We describe, below, a basic protocol for the fractionation of mammalian tissue (House mouse; *Mus musculus*) into defined organelle fractions, and their preparation for MudPIT analysis. The quantitative estimation of protein abundance based on *Spectral Counting* (SpC) is discussed *(14, 15)*. Additionally, a basic protocol for the isolation, isotopic labeling and quantitative analysis of membrane proteins from *Halobacterium salinarium* and the use of MALDI-TOF-MS/MS analysis for data dependent MS-analysis is discussed *(16)*.

3.1. Tissue Homogenization and Organelle Fractionation

1. Mice are CO_2-asphyxiated and sacrificed. The tissue of interest is removed, carefully minced with a razorblade and washed three times with ice-cold PBS. Minced tissue samples are homogenized in ice-cold *tissue homogenization buffer* in a dounce homogenizer, with at least 15 strokes. All subsequent centrifugation steps are performed at 4°C. Tissue lysate is centrifuged for 15 min at 800 × *g*. The supernatant (crude cytoplasm) is subsequently used for the isolation of mitochondria,

membranes and cytosol. The pellet contains crude nuclei, which are further purified. The crude nuclear pellet is resuspended in tissue homogenization buffer, layered onto *sucrose cushion solution 1*, and centrifuged for 15 min at 1,000 × *g*. To further purify nuclei, the pellet is resuspended in 8 mL of *sucrose cushion solution 1* and carefully layered onto 4 mL of *sucrose cushion solution 2* in a 13-mL ultracentrifugation tube. The mixture is pelleted at 100,000 × *g* for 60 min using a Beckman SW40.1 rotor. The pellet containing purified nuclei is resuspended in *nuclear extraction buffer 1* and incubated on ice for 15 min. Nuclei are lysed by ten passages though an 18-gauge needle, followed by centrifugation at 13,000 rpm, 21,000×g for 30 min. The supernatant is *nuclear extract 1*. The pellet is resuspended in *nuclear extraction buffer 2*, incubated on ice for 30 min, and then centrifuged at 13,000 rpm for 30 min. The resulting supernatant is *nuclear extract 2*.

2. Mitochondria are isolated by centrifugation of the crude cytoplasmic fraction at 8,000 × *g* for 20 min. The supernatant is collected for the isolation of mixed membranes and cytosol. The mitochondrial pellet is washed twice in *tissue homogenization buffer*, resuspended in *mitochondrial extraction buffer 1*, and incubated on ice for 30 min, followed by brief sonication. The solution is centrifuged at 13,000 rpm, 21,000×g for 30, min and the supernatant collected as *mitochondrial extract 1*. The pellet is resuspended in *mitochondrial extraction buffer 2* and incubated on ice for 30 min, followed by centrifugation at 13,000 rpm for 30 min. The supernatant is collected as *mitochondrial extract 2*.
3. Finally, the mixed membrane fraction is isolated from the crude cytoplasmic supernatant by centrifugation at 100,000 × *g* for 60 min (Beckman SW40.1 rotor), and the resulting pellet extracted in *mitochondrial extraction buffer 2* for 30 min on ice. This preparation is spun at 13,000 rpm 21,000×g, for 30 min and the supernatant collected as *mixed membrane extract*. The supernatant from the final ultracentrifugation at 100,000 × *g* is considered the *cytosol* (*see* **Notes 1** and **2**).

3.2. Protein Digestion and Preparation for MudPIT Profiling

1. One hundred fifty micrograms of total protein are precipitated over night at −20°C with 5 volumes of ice-cold acetone.
2. The solution is centrifuged at 13,000 rpm, 21,000×*g* for 15 min at 4°C.
3. The protein pellet is washed once with 150 mL of ice-cold acetone.
4. The protein pellet is carefully solubilized in 8 M urea, 50 mM Tris–HCl, pH 8.5, 2 mM DTT (dithiothreitol) at 37°C for 30 min.
5. IAA (iodoaceteamide) is added to a final concentration of 8 mM and the solution is incubated in the dark at 37°C for 30 min.

6. The sample is diluted to 4 M urea with 100 mM ammonium bicarbonate, pH 8.5 and digested with endoproteinase Lys-C (ratio 1:150) at 37°C for 6 h.
7. The digestion mixture is further diluted to 2 M urea with 50 mM ammonium bicarbonate, pH 8.5, $CaCl_2$ is added to a final concentration of 1 mM and the solution digested with trypsin (ratio 1:150) at 37°C over night.
8. The resulting peptide mixture is solid phase extracted with Varian OMIX cartridges according to the manufacturer's instructions and stored at –80°C until further use.

3.3. MudPIT Analysis

1. A fully automated MudPIT is used to analyze each sample, as previously described *(14, 15)*. Each MudPIT consists of several independent chromatography separation steps that form a sequence. Fused silica microcapillary columns (i.d. 100 mm) are pulled to a fine tip (~5–10 mm) using a Sutter P-2000 laser puller. The columns are custom packed with ~10 cm ZORBAX Eclipse XDB-C18 5 mm, followed by ~6 cm Partisphere SCX resin using a pressure vessel. Samples are loaded onto the column using the same pressure vessel and placed in-line with a capillary HPLC system (Agilent 1100 series or Thermo Finnigan Surveyor). The HPLC pumps are operated at a constant flow rate of 100 mL/min with a precolumn flow splitter. The effective flow rate at the column is ~200–400 nL/min. Peptides are directly eluted into the mass spectrometer. We use the Thermo Finnigan LTQ linear iontrap equipped with a Proxeon Biosystems Nano Electrospray Ion Source. Chromatographic elution profiles and individual salt concentration steps for a typical MudPIT analysis are shown below **(Table 1)** (*see* **Note 3**).
2. The mass spectrometer is operated with the following settings.
   (a) Distal spray voltage: 2.3 kV.
   (b) Full scan mass spectrum: 400–1,400 *m/z*.
   (c) Six data-dependent MS/MS scans at a 35% normalized collision energy.

3.4. Sequest Searches and Validation by the STATQUEST Algorithm

1. MS data is searched using the Sequest algorithm, against publicly available protein sequence databases. We use the Swiss-Protein/TrEMBL (http://ca.expasy.org/) and IPI (http://www.ebi.ac.uk/IPI/IPIhelp.html) databases. To objectively estimate our false positive rate, we use the "target/decoy" database strategy *(14, 17)*, in which every protein sequence in the native "target" protein sequence database is reversed to generate a "decoy" database. Briefly, the number "decoy" proteins appearing in a list of protein identifications will provide a rough estimate of the false positive rate.

Table 1
Elution profiles

Time (min)	Buffer A (%)	Buffer B (%)	Buffer C (%)	Flow rate (mL/min)
0	100	0	0	100
2	100	0	0	100
2.01	90	0	10	100
7	90	0	10	100
7.01	100	0	0	100
12	100	0	0	100
12.01	95	5	0	100
85	30	70	0	100
88	100	0	0	100
90	100	0	0	100

Salt steps

Step	Buffer A (%)	Buffer B (%)	Buffer C (%)
1	100	0	0
2	90	0	10
3	80	0	20
4	70	0	30
5	65	0	35
6	60	0	40
7	50	0	50
8	45	0	55
9	40	0	60
10	20	0	80
11	0	0	100

2. Sequest search results are validated using an in-house probability-based algorithm termed STATQUEST *(14)*. This computer tool automatically assigns a defined percentage likelihood of correct peptide identification to every Sequest search result. For LTQ data we only accept peptide identifications with a confidence interval of ≥99%.
3. As MudPIT profiling generates very large data files, we recommend parsing all search results and their associated Sequest and STATQUEST scores into a relational database. This step significantly speeds up subsequent data analysis.

3.5. Quantitative Proteomics Based on SpC

In principal, semi-quantitative proteomics based on SpC is a simple and straightforward methodology. It is important, however, that only spectra matching to peptides identified with high confidence (e.g., based on STATQUEST *(14)* or PeptideProphet *(18)*) are considered for comparison.

1. The number of peptides (or spectra) confidently identified and matched to a protein in the sequence database are used for quantitative comparison.
2. The number of spectra matching the same protein in different samples is compared.
3. In **Fig. 1** we present the data from a recently published manuscript *(15)*. Proteins identified in several subcellular fractions isolated from several healthy mouse tissues are clustered based on their SpC values (**Fig. 1a**). Western blotting results against specific cellular markers and their associated SpC values are presented (**Fig. 1b**). The two types of data are highly correlated (*see* **Note 4**).

3.6. ICPL Workflow

The complete workflow is shown in **Fig. 2** as recently published *(16)*. Membrane proteins obtained from aerobically and phototrophically grown cells, respectively, were first individually reduced, alkylated and labeled with either the "*light*" or "*heavy*" version of the ICPL reagent. After combining both mixtures, proteins were cleaved into peptides using trypsin and separated by nano-reversed phase HPLC. The eluting peptides were mixed with MALDI-matrix and directly spotted onto MALDI plates for data dependent MS/MS analysis. Since peptides of identical sequence derived from the two differentially labeled protein samples differ in mass. They appear as doublets in the acquired MS-spectra. From the ratios of the ion intensities of these sister peptide pairs, the relative abundance of their parent proteins in the original samples can be determined. Subsequently, isotopic peptide pairs that differ in abundance were automatically selected for collision-induced dissociation (CID) and identified by correlation with sequence databases using the MASCOT search algorithms *(19)*.

3.7. Bacterial Strain and Growth Conditions

Halobacterium salinarum (strain R1, DSM 671) was grown in complete medium *(20)* as previously described *(21)*. Briefly, for preparation of a starter culture, Halobacterium was grown aerobically in the dark at 37°C in 1 L of complete medium to the stationary phase. For protein preparation, Halobacterium was grown through three successive transfers to ensure uniform cell state. For the first two transfers, 35 mL of fresh medium was inoculated with 1 mL of the previous culture, for the third transfer 1 L medium in a 2 L flask was inoculated with 35 mL of the previous culture. The cells were grown to late log-phase (30–40 Klett units), either aerobically in the dark or phototrophically

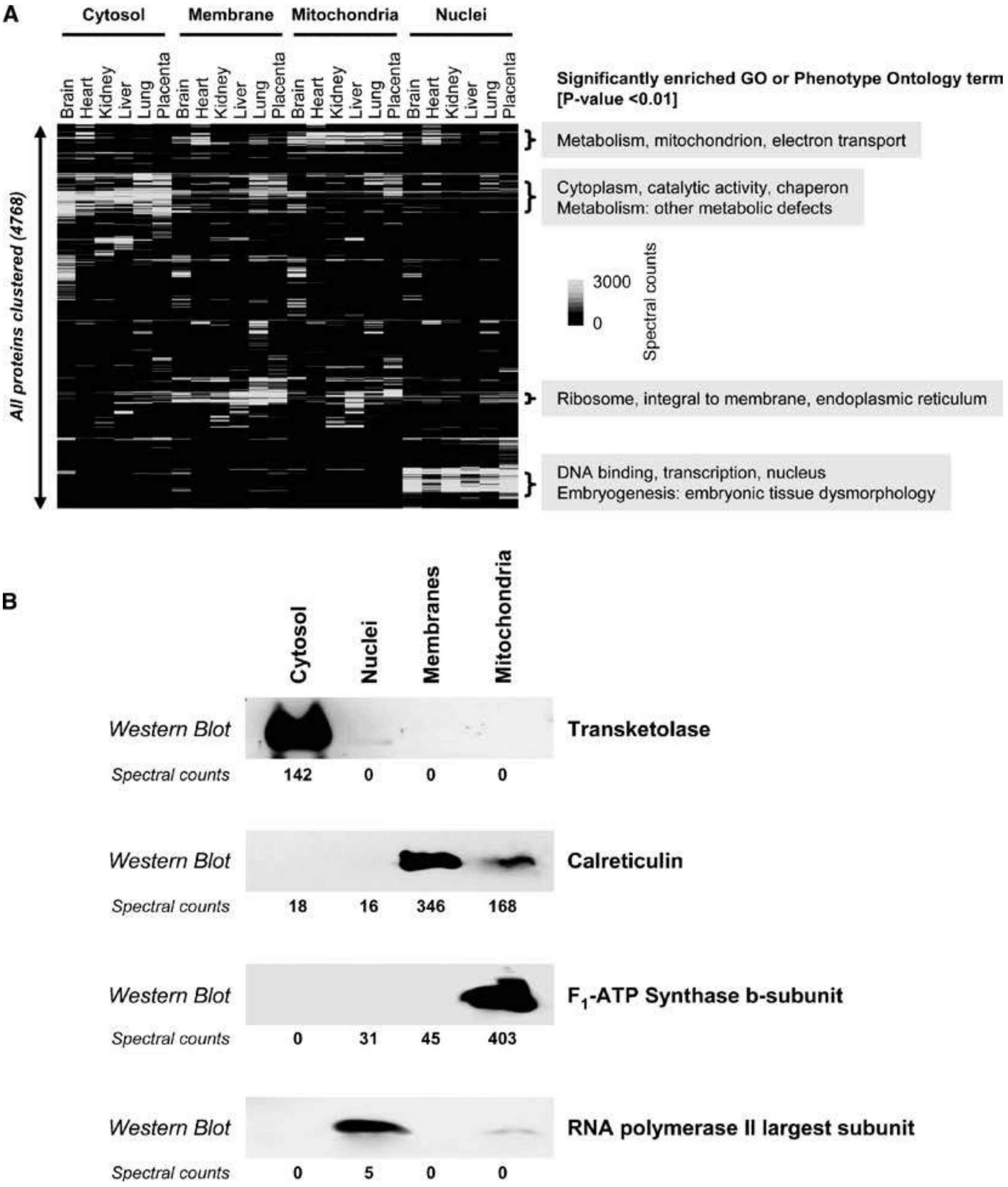


Fig. 1. Quantitative proteomics by SpC. (a) Confidently identified proteins are clustered based on SpC. (b) Selected marker proteins are shown by Western blotting together with their corresponding SpC.

with light as energy source. For the latter, flasks were closed after inoculation so that residual oxygen was consumed and growth continued under anaerobic conditions.

3.8. Membrane Protein Preparation

Two liters of cell culture were centrifuged for 50 min at 4,000 × *g* and cells were resuspended in 40 mL Basal salt buffer before cell

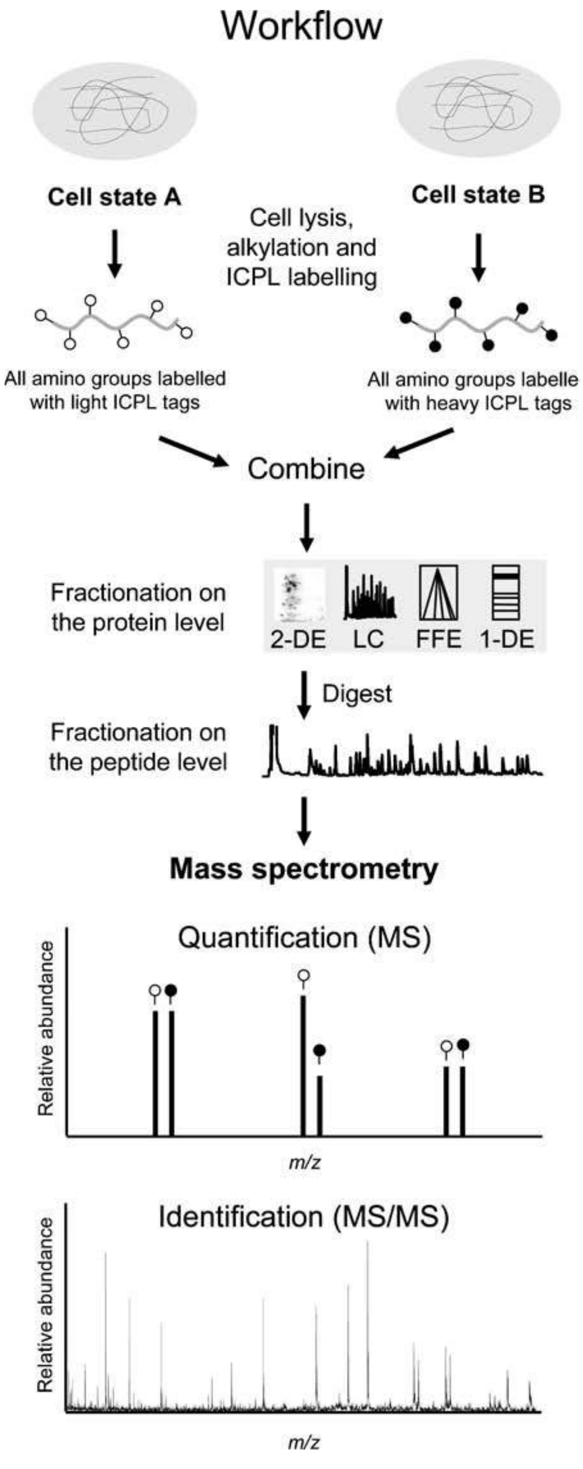

Workflow
Cell state A
Cell state B
Cell lysis, alkylation and ICPL labelling
All amino groups labelled with light ICPL tags
All amino groups labelle with heavy ICPL tags
Combine
Fractionation on the protein level
2-DE
LC
FFE
1-DE
Digest
Fractionation on the peptide level
Mass spectrometry
Quantification (MS)
Relative abundance
m/z
Identification (MS/MS)
Relative abundance
m/z

Fig. 2. Quantitative proteomics by ICPL. Protein extracts are labeled with either the “light” or the “heavy” version of the ICPL label. Protein extracts are combined, fractionated and analyzed by MS. Quantitation is achieved by comparing the relative intensity of each peptide pair in MS-mode.

rupture by sonication (3 × 1 min on ice, 50% Duty Cycle, Branson Sonicator). Solid debris was removed by a short centrifugation step (5,800 rpm, 10 min, 10°C, 4,000 × *g*). The vesicles were centrifuged (30,000 rpm, 1 h, 4°C, 65,000 × *g*), the pellet was resuspended in 2 mL Basal salt buffer, layered over a linear sucrose density gradient (10–60% sucrose in Basal salt buffer (w/w)) and centrifuged for 14 h (25,000 rpm, 4°C, 80,000 × *g*). The colored vesicle band was collected and sucrose was removed by dilution with 1 M NaCl and pelleting of vesicles by centrifugation. This step was repeated with 500 mM NaCl and the final pellet was resuspended in 2 mL H_2O. The membranes were delipidated with chloroform/methanol as described *(22)* and precipitated proteins were subsequently lyophilized (*see* **Notes 5** and **6**).

3.9. Isotope Labeling

One hundred micrograms of total protein obtained from aerobically and phototrophically growing cells are dissolved in 20 mL Sample buffer, respectively, and if necessary, the pH is adjusted to 8.5 ± 0.1 using 1 M NaOH or HCl.

3.9.1. Alkylation

Both samples are equally processed.

1. 0.5 mL of *Reduction buffer* is added and the solution kept at 60°C for 30 min.
2. After cooling and sinning down, 0.5 mL of *Alkylation buffer* is added and the solution is incubated in the dark at 25°C for 30 min.
3. The alkylation is stopped by adding 0.5 µL of *Alkylation stop buffer* and incubated for another 15 min at room temperature.

3.9.2. Isotope Labeling

1. Three microliters of *Labeling solution* light are added to aerobically grown sample and 3 mL of *Labeling solution* heavy are added to the phototrophically grown sample and the mixtures are incubated for 2 h at 25°C.
2. Then, 2 µL of Stop solution are added to each sample.
3. After 15 min at room temperature, the two samples are combined.
4. The pH is adjusted to 11.9–12.0 by adding 2 M NaOH and incubated for 20 min at 25°C.
5. Then, an equimolar amount of HCl is added to the sample to lower the pH to its original value (*see* **Notes 7** and **8**).

3.10. Sample Digestion and Preparation for LC-MS

1. The sample was diluted with 25 mM Tris (pH 8.5) to a final guanidine HCl concentration of 0.5 M and digested overnight at 37°C with trypsin (substrate-to-enzyme ratio = 50:1) (*see* **Note 9**).
2. The resulting peptide mixture was acidified with 10 mL 1% TFA and the volume reduced by evaporation to approximately 30 µL.

3. The resulting peptide mixture was solid phase extracted using self-made C18-columns packed with Reprosil-Pur 120 ODS, C18 particles (Dr. Maisch GmBH, Germany, Cat#:r13.93) as described *(23)*, dried under vacuum and stored at −80°C until further use.

3.11. LC-MALDI-TOF/TOF Analysis

1. All peptide separations were performed utilizing a capillary liquid chromatography system (Ultimate, LC Packings) containing a reversed-phase column (LC Packings Pepmap reversed-phase C18 column, 75 mm i.d., 15 cm) coupled directly online with a MALDI target spotter (Probot, LC Packings). A sample volume of 50 mL was injected and the peptides were trapped on a short reversed-phase column (300 mm i.d., 5 mm) using Loading buffer at a flow rate of 20 mL/min. For the separation of the peptides, a 65 min linear gradient from 10 to 45% B at a flow rate of 200 nL/min was used, followed by a 20 min wash step of the column with 100% B. The analytical column was directly connected to a MicroTee (Upchurch, Oak Harbor, WA) where the eluent was mixed with MALDI matrix solution (freshly prepared 1:1 mixture of Matrix buffer 1 and 2) at a flow rate of 1.3 μL/min and deposited onto a blank MALDI plate. The LC-eluent was automatically spotted in 10 s fractions over a time period of 66.66 min resulting in 400 spots per MALDI target plate. The sample spots were allowed to dry at room temperature.
2. The mass spectrometer is operated with the following settings.
   - (a) 2,500 laser shots for each MS spectrum and 1,500 shots for each MS/MS spectrum.
   - (b) Deflection cut off range was 700 *m/z*.
   - (c) Focus mass was 2,100 *m/z*.
   - (d) CID spectra were acquired using collision energy of 1 keV and nitrogen as collision gas.
3. MS/MS-data acquisition was done in a data dependent manner. First, MS-spectra of each LC-fraction were acquired. Then, isotopic labeled peptide pairs were automatically detected and quantified using the Peakpicker software (Applied Biosystems). To keep the analysis time as low as possible, only differentially regulated peptide pairs and from these only the more intense MS-peak were selected for MS-sequencing (*see* **Note 10**).

3.12. Data Analysis

1. All MS/MS-spectra obtained were searched against the Halobacterium protein sequence database that was exported from the HaloLex database (http://www.halolex.mpg.de) *(24)*, using an in house version of Mascot *(19)* in combination with the GPS-Explorer™ 2.0 software (Applied Biosystems). For the

database search, carbamidomethylation was set as a required cysteine modification, whereas oxidation of methionine was considered as a variable modification. Further potential modifications include $^{12}C_6$- and $^{13}C_6$-nicotinoylation of lysine and the protein N terminus. It is important to note that trypsin does not cleave the labeled lysines. Therefore, the enzyme Arg-C should be selected for database searching (*see* **Note 9**).

2. To objectively estimate our false positive rate, we use the "target/decoy" database strategy *(25)*, in which every protein sequence in the native "target" protein sequence database is reversed to generate a "decoy" database. Thereby, the false positive rate was estimated to be below 2%.

3.13. Calculation of Peptide Ratios

The ratio for each peptide pair was calculated using the program Peakpicker (Applied Biosystems). Ratios for each protein were determined by averaging all quantified peptides of one protein utilizing the DecisionSite 8.0 software (Spotfire AB, Goeteborg, Sweden) on the basis of the raw data. The median of the complete set of quantified peptides was determined and used for a computational normalization of the original ratios. Finally, regulation factors were computed for each protein such that the same extent of positive or negative regulation results in an identical absolute value of the regulation factor. To provide symmetric regulation factors these ratios were inversed and multiplied by −1. This scale excludes any values between 1 and −1.

The application of the ICPL methodology to the analysis of highly purified membranes of the halophilic archaeon *Halobacterium salinarum* resulted in the accurate quantification of over 150 membrane proteins *(16)*. Importantly, the comparison of the ICPL results to DIGE labeling in combination with an improved two-dimensional 16-BAC/SDS–PAGE procedure showed excellent correlation between both complementary technologies. In a proof-of-principle experiment two different growth conditions (aerobic vs. anaerobic/phototrophic) were compared by quantitative proteomics. Several differentially regulated proteins involved in photosynthesis and energy metabolism could be detected.

3.14. Conclusions

Advanced MS-based proteomics and allied bioinformatics tools enable biologists to confidently identify hundreds to thousands of proteins in complex biological samples. Direct quantitative comparison of different conditions (e.g., disease or development) on a global scale is very important to fully understand biological processes and mechanisms in a non-hypothesis driven manner. Several methodologies have been developed in recent years to accomplish this goal, possibly the simplest approach being SpC. Isotope labeling can provide significantly more reliable and accurate quantification over label-free quantitative approaches, but the ongoing progress in both, instrumentation and software, will further increase the quality of such approaches. Therefore, it is to be expected that label-free quantification of proteins using methods based on spectral counting or MS-signal intensities in concert with their simple workflows and low costs will become widely used in the near future. The applications of these state-of-the-art technologies to diverse biological settings provide a unique opportunity for a more complete biological understanding.

4. Notes

1. The isolation of pure organelles is a challenging task, if not impossible. The protocol presented above was applied in our lab and works well with MS-based analysis. It provides reasonably clean or "enriched" organelle fractions, although some cross-contamination between the mixed membrane and mitochondrial fractions was observed. Nevertheless, one should always consider the possibility that some proteins are actually present in more than one cellular location, even if annotation is only available for one subcellular fraction.
2. Organelles should only be isolated from fresh tissues and cells.
3. The extreme complexity of mammalian proteome samples is daunting. Even high resolution procedures such as MudPIT are unable to detect every peptide present. The process is further complicated by the skewed range of overall protein abundance, with some proteins being present in very high abundance, and most proteins present in lower abundance. The detection of these "low abundance" proteins is like "finding a needle in a haystack." Several strategies for the enrichment of low abundance peptides have been presented in the scientific literature *(26–28)*.
    (a) Sample fractionation prior to MS analysis
        (i) Organelle fractionation
        (ii) Biochemical fractionations (e.g., ion exchange chromatography)
    (b) Repeat analysis of the same sample also increases the overall detection depth. We highly recommend analyzing several technical replicates: "random sampling effect" *(29)*.
4. Quantitative comparisons based on SpC are new and relatively unproven. Although several papers were recently published suggesting good correlation between relative protein abundance and SpC *(15, 29–34)*. We suggest caution in interpretation of quantitative data based on SpC. This is especially true if

low spectral counts and small differences between samples are observed. We suggest validating results using alternative methodologies (e.g., Western blotting, if antibodies are available). Normalization strategies of SpC values and statistical approaches for the comparison have also been described recently. (*(30, 31, 34, 38)*. Spectral counting is further complicated by the "protein inference" problem, where confidently identified peptides are shared among different database entries. Programs such as *ProteinProphet (35)* or *Isoform Resolver (30)* can help in grouping these proteins.

5. To effectively remove guanidine HCl, a charged molecule that impairs sample preparation steps, and to concentrate your sample for 1D-SDS–PAGE, acetone precipitation can be used. Therefore, the ICPL-labeled sample solution is diluted 1:1 (v/v) with water followed by 5 volumes (based on the volume of sample and water) of ice-cold acetone. After incubation at –20°C over night, the sample is centrifuged at 14,000 rpm, 21,000 × *g* for 30 min at 4°C. The supernatant is removed, the pellet is gently washed once with a ice-cold solution of 80% acetone and the sample is once more centrifuged at 14,000 rpm at 4°C for 15 min. Finally, the supernatant is discarded and the protein pellet dried under vacuum. The sample can either be stored at –80°C or further processed.
6. However, the comprehensive precipitation of hydrophobic membrane proteins is a difficult task. In our experiment, a significantly higher number of membrane proteins was identified by skipping the precipitation step and by direct digestion of the sample in solution like described above.
7. The pH of the solution is very important to obtain a specific and complete modification of all amino groups. It should therefore be tested and, if required, adjusted before adding the ICPL-reagent.
8. The labeling procedure has no impact on protein phosphorylation sites *(7)*. Therefore, the ICPL-method can be combined with phosphopeptide enrichment strategies, like IMAC *(36)* or TiO_2 *(11)*, to quantify changes of protein phosphorylation.
9. Since the all lysines are blocked, trypsin does not cleave after this amino acid. As a result, the enzyme Arg-C must be selected for database searches and endoproteinase Lys-C cannot be used at all to digest ICPL-labeled proteins. However, database searches with trypsin selected, offer an effective opportunity to determine the ICPL-labeling efficiency because no lysine terminating peptides should be identified, if all amino groups have been modified.

10. The number of modified lysines, can easily be calculated from the mass gap of each isotopic peptide pair, serves as a strong constraint for database searching and highly increases protein identification confidence of labeled peptides *(16)*.

Acknowledgements

The author would like to thank Brian Raught for critical reading of the manuscript.

References

1. Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., and Humphery-Smith, I. (1995) Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* 16, 1090-4.
2. Steen, H., and Mann, M. (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* **5**, 699-711.
3. Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., III (1999) Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 17, 676-82.
4. Washburn, M. P., Wolters, D., and Yates, J. R., III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19, 242-7.
5. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 1, 252-62.
6. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17, 994-9.
7. Schmidt, A., Kellermann, J., and Lottspeich, F. (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 5, 4-15.
8. Munchbach, M., Quadroni, M., Miotto, G., and James, P. (2000) Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety. *Anal Chem* 72, 4047-57.
9. Regnier, F. E., and Julka, S. (2006) Primary amine coding as a path to comparative proteomics. *Proteomics* 6, 3968-79.
10. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3, 1154-69.
11. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635-48.
12. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376-86.
13. Staes, A., Demol, H., Van Damme, J., Martens, L., Vandekerckhove, J., and Gevaert, K. (2004) Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18. *J Proteome Res* 3, 786-91.
14. Kislinger, T., Rahman, K., Radulovic, D., Cox, B., Rossant, J., and Emili, A. (2003) PRISM, a generic large scale proteomic investigation strategy for mammals. *Mol Cell Proteomics* 2, 96-106.
15. Kislinger, T., Cox, B., Kannan, A., Chung, C., Hu, P., Ignatchenko, A., Scott, M. S., Gramolini, A. O., Morris, Q., Hallett, M.

T., Rossant, J., Hughes, T. R., Frey, B., and Emili, A. (2006) Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell* 125, 173-86.

16. Bisle, B., Schmidt, A., Scheibe, B., Klein, C., Tebbe, A., Kellermann, J., Siedler, F., Pfeiffer, F., Lottspeich, F., and Oesterhelt, D. (2006) Quantitative profiling of the membrane proteome in a halophilic archaeon. *Mol Cell Proteomics* 5, 1543-58.
17. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* 2, 43-50.
18. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74, 5383-92.
19. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551-67.
20. Oesterhelt, D., and Stoeckenius, W. (1974) Isolation of the cell membrane of Halobacterium halobium and its fractionation into red and purple membrane. *Methods Enzymol* 31, 667-78.
21. Oesterhelt, D., and Krippahl, G. (1983) Phototrophic growth of halobacteria and its use for isolation of photosynthetically-deficient mutants. *Ann Microbiol (Paris)* 134B, 137-50.
22. Wessel, D., and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* 138, 141-3.
23. Rappsilber et al. (2003) *Anal Chem* **75**(3), 663-70.
24. Tebbe, A., Klein, C., Bisle, B., Siedler, F., Scheffer, B., Garcia-Rizo, C., Wolfertz, J., Hickmann, V., Pfeiffer, F., and Oesterhelt, D. (2005) Analysis of the cytosolic proteome of *Halobacterium salinarum* and its implication for genome annotation. *Proteomics* 5, 168-79.
25. Moore, R. E., Young, M. K., and Lee, T. D. (2002) Qscore: an algorithm for evaluating SEQUEST database search results. *J Am Soc Mass Spectrom* 13, 378-86.
26. Andersen, J. S., and Mann, M. (2006) Organellar proteomics: turning inventories into insights. *EMBO Rep* 7, 874-9.
27. Chen, E. I., Hewel, J., Felding-Habermann, B., and Yates, J. R., III (2006) Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT). *Mol Cell Proteomics* 5, 53-6.
28. Simpson, D. C., Ahn, S., Pasa-Tolic, L., Bogdanov, B., Mottaz, H. M., Vilkov, A. N., Anderson, G. A., Lipton, M. S., and Smith, R. D. (2006) Using size exclusion chromatography-RPLC and RPLC-CIEF as two-dimensional separation strategies for protein profiling. *Electrophoresis* 27, 2722-33.
29. Liu, H., Sadygov, R. G., and Yates, J. R., III (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76, 4193-201.
30. Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., Resing, K. A., and Ahn, N. G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* 4, 1487-502.
31. Roth, A. F., Wan, J., Bailey, A. O., Sun, B., Kuchar, J. A., Green, W. N., Phinney, B. S., Yates, J. R., III, and Davis, N. G. (2006) Global analysis of protein palmitoylation in yeast. *Cell* 125, 1003-13.
32. Roth, A. F., Wan, J., Green, W. N., Yates, J. R., and Davis, N. G. (2006) Proteomic identification of palmitoylated proteins. *Methods* 40, 135-42.
33. Zybailov, B., Coleman, M. K., Florens, L., and Washburn, M. P. (2005) Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal Chem* 77, 6218-24.
34. Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L., and Washburn, M. P. (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J Proteome Res* 5, 2339-47.
35. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75, 4646-58.
36. Kokubu, M., Ishihama, Y., Sato, T., Nagasu, T., and Oda, Y. (2005) Specificity of immobilized metal affinity-based IMAC/C18 tip enrichment of phosphopeptides for protein phosphorylation analysis. *Anal Chem* 77, 5144-54.
37. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207.
38. Gramolini A.O. et al. 2008 Comparative proteomics profiling of a phospholamban mutant mouse model of dilated cardiomyopathy reveals progressive intracellular stress responses. *Mol Cell Proteomics.* 7(3):519-533.

Chapter 3

Label-Free Relative Quantitation of Prokaryotic Proteomes Using the Accurate Mass and Time Tag Approach

Kim K. Hixson

Summary

Prokaryotic protein expression changes in detectable amounts due to the environmental stimuli encountered by the organism. To understand the underlying mechanisms involved it is necessary to comprehensively detect both the proteins present and their relative abundance under the growth conditions of interest. LC-MS based accurate mass and time (AMT) tag method along with the use of clustering software can provide a visual and more comprehensive understanding of significant protein abundance increases and decreases. These data then can be effectively used to pin-point proteins of interest for further genetic and physiological studies. This method allows for the identification and quantitation of thousands of proteins in a single mass spectrometric analysis and is more comprehensive than two dimensional electrophoresis and shotgun approaches.

Key words: Proteomics, Accurate mass and time tag, Label-free, Relative quantitation, LC-MS, Bacteria.

1. Introduction

Prokaryotic organisms are prevalent in our environment. They carry out many important roles including nitrogen fixation, carbon cycling, energy production, and pathogenesis *(1–4)*. Prokaryotes are the simplest of self replicating organisms. Their cell structure consists primarily of either one or two membranes, usually supported by a cell wall. While the DNA is considered to be the unchanging blueprint or parts list of the cell, the proteins (translated from select portions of the DNA) are actually the little machines and structures within the cell that allow it to function. The unchanging DNA inside the cell dictates what proteins can

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_3

possibly be made but the specific type and quantity of protein produced in a cell changes in response to its environment and growth state. Understanding these protein changes allows insight into the biology of the cell. Here we present a label-free proteome characterization of two different bacteria using the accurate mass and time (AMT) tag method *(5, 6)*. In addition to the organisms presented in this chapter, the AMT tag method has been shown to be useful for the proteome characterizations of many other bacteria, eukaryotic tissues, and viruses *(1, 7–12)*.

The AMT tag global proteomics strategy allows for the rapid characterization of a large number of proteins by identifying the unique molecular masses of many peptides resulting from an enzymatic digestion of the proteins *(13)*. Central to the AMT tag approach is the use of a high resolution capillary LC peptide separation combined with tandem mass spectrometry (MS/MS) and liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry (LC-FTICR MS). An initial reference database, constructed from MS/MS (often from an ion trap mass spectrometer) analyses provides a cursory list of peptide identifications and corresponding LC elution times for all detected peptides in a sample set. The amino acid sequences obtained from these MS/MS spectra are typically identified by the software SEQUEST *(14)* or MASCOT *(15)* which use the organism's annotated genome to match the theoretical spectra with the experimental spectra. Once this initial reference database is constructed, routine MS/MS analyses are no longer needed and all subsequent data used for comparisons of protein abundance can be obtained from high-throughput LC-FTICR MS analyses only. LC-FTICR MS is capable of measuring and discriminating between thousands of high and low abundant peptides due to its inherent high mass measurement accuracy (MMA) and enhanced sensitivity. The peptides from the LC-FTICR MS analyses that match the MS/MS peptide reference database with respect to both an accurate mass (e.g., ±1 ppm) and normalized elution time (e.g., ±1–2%) are then defined as AMT tags. The intensity values obtained from the LC-FTICR MS data for each AMT tag can be used for relative quantitation. Proteins with significantly increased or decreased abundance values due to a specific growth condition can be grouped together and placed in a visually comparative color map. This technique proves useful for narrowing down potential proteins of interest from high-throughput organism-wide analyses.

Here we demonstrate this method on both *Yersinia pestis (16)*, a bacterial pathogen that causes plague, and *Geobacter sulfurreducens (17)*, a soil bacterium that reduces iron in anaerobic environments to produce energy. To obtain the proteomic data for both organisms a trypsin digestion was performed on proteins obtained from a whole cell lysate (for *Geobacter* and *Yersinia*), and two sub-cellular fractions ("integral membrane proteins," and "soluble proteins") (for *Yersinia* only) (*see* **Fig. 1**). Proteins

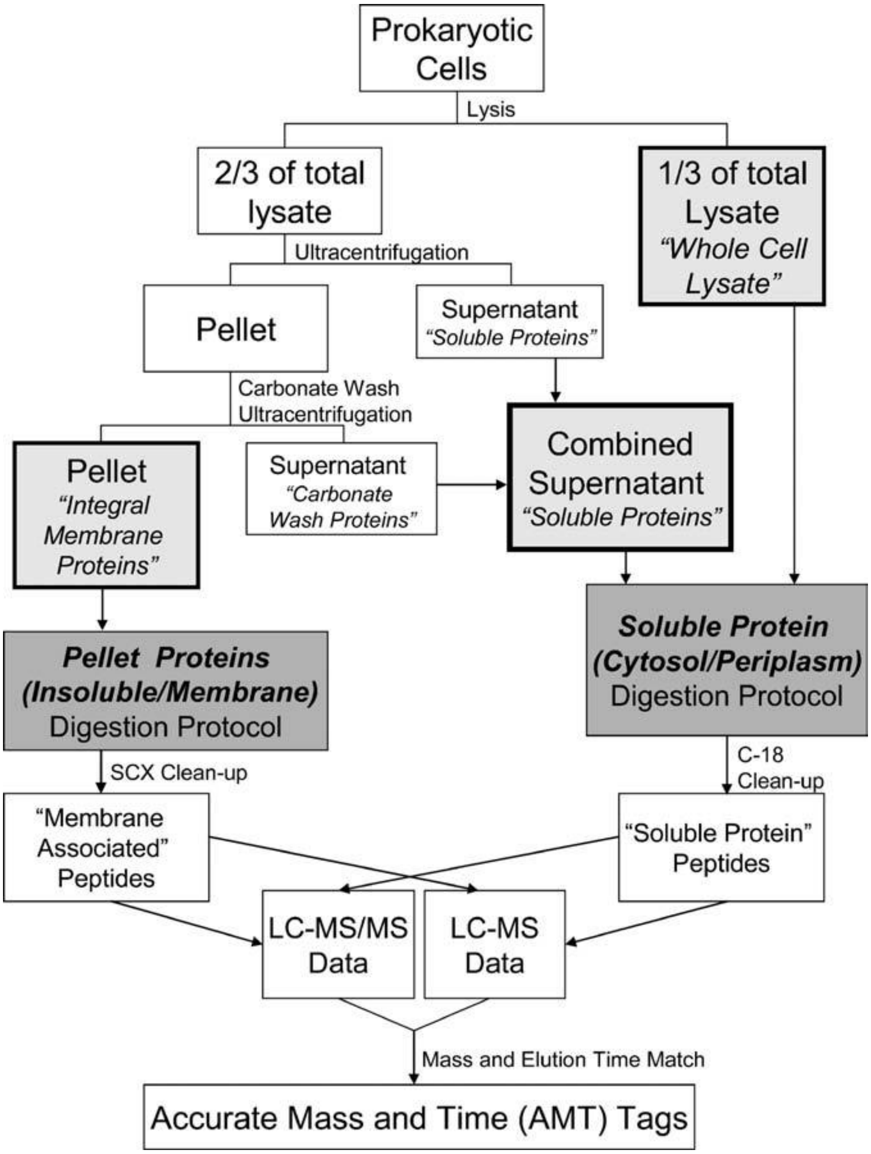

Prokaryotic Cells
Lysis
2/3 of total lysate
1/3 of total Lysate "Whole Cell Lysate"
Ultracentrifugation
Pellet
Supernatant "Soluble Proteins"
Carbonate Wash Ultracentrifugation
Pellet "Integral Membrane Proteins"
Supernatant "Carbonate Wash Proteins"
Combined Supernatant "Soluble Proteins"
Pellet Proteins (Insoluble/Membrane) Digestion Protocol
Soluble Protein (Cytosol/Periplasm) Digestion Protocol
SCX Clean-up
C-18 Clean-up
"Membrane Associated" Peptides
"Soluble Protein" Peptides
LC-MS/MS Data
LC-MS Data
Mass and Elution Time Match
Accurate Mass and Time (AMT) Tags

Fig. 1. Work flow chart of the sample preparation of prokaryotic cells for label-free proteomic analysis of prokaryotic cells

were digested with trypsin to produce specific polypeptides. Polypeptides are more compatible with detection in a mass spectrometer since they are smaller and more soluble than whole proteins, allowing for identification of most proteins in the relatively small mass/charge range of 400–2,000. Even integral membrane proteins (IMP), which are notoriously insoluble, can be identified due to the fact that most IMPs contain soluble portions that are likely to be liberated and detected. The digestion of the "whole cell lysate" produces soluble peptides from all soluble proteins and the soluble (externally exposed to the membrane) portions of integral membrane proteins. Thus proteomic output from the "whole cell lysate" shows the overall protein output of the cell. As in the case with the *Yersinia* proteome analysis, the digestion of the sub-cellular fractions was also performed. Data obtained from sub-cellular fractions provide a greater number of peptide identifications allowing for a more complete observation of the proteome and for more detail with regard to specific protein changes generally occurring in either the membrane or cytosolic/periplasmic portions of the cell.

2. Materials

2.1. Cell Lysis

1. Water (*see* **Note 1**).
2. Mini-beadbeater (Biospec, Bartlesville OK, cat# 3110BX).
3. 0.1 mm zirconia/silica beads for beadbeater (Biospec, Bartlesville OK, cat# 11079101z).
4. Siliconized, sterile individually wrapped microcentrifuge tubes, 1.5 and 2.0 mL (Fisher Scientific, cat# 02-681-332, cat# 02-681-331) (*see* **Note 2**).
5. Falcon 15-mL polypropylene conical tubes (Fisher Scientific, cat# 14-959-49B) (*see* **Note 3**).
6. Becton Dickinson 26 G needles (Fisher Scientific, cat# 14-826-10).

2.2. Subcellular Fractionation and "Integral Membrane Protein Enrichment"

1. Beckman Optima TL ultracentrifuge and rotor TLA 120.1 (Beckman, Fullerton CA)
2. 8 mm × 34 mm polycarbonate centrifuge tubes (Beckman, Fullerton CA, cat# 343776)
3. 100 mM sodium carbonate (*see* **Note 1**).
4. Branson 1510 sonicating bath (Fisher Scientific, cat# 15-337-22C)
5. Siliconized, sterile individually wrapped microcentrifuge tubes, 1.5 and 2.0 mL (Fisher Scientific, cat# 02-681-332, cat# 02-681-331) (*see* **Note 2**).
6. 20% Glacial acetic acid (*see* **Note 1**).
7. ColorpHast pH 0–14 pH indicator strips (Fisher Scientific, cat# S60169) (*see* **Note 4**).

2.3. Coomassie Blue Assay to Determine Protein Concentration

1. Coomassie Plus Bradford Assay Kit (Pierce, Rockford IL, cat# 23236) (*see* **Note 1**).
2. Siliconized, sterile individually wrapped microcentrifuge tubes, 1.5 and 2.0 mL (Fisher Scientific, cat# 02-681-332, cat# 02-681-331) (*see* **Note 2**).

3. Breathe-EASIER Tube Membrane (Diversified Biotech, Boston MA, cat# BMTM-2000) (*see* **Note 14**).

2.4. Internal Reference Proteome

1. Apomyoglobin, from equine (Sigma, St. Louis MO, cat# A8673-1VL)
2. Cytochrome *c*, from bovine (Sigma, St. Louis MO, cat# C2037-50MG).
3. Glyceraldehyde-3-phosphate dehydrogenase, from rabbit (Sigma, St. Louis MO, cat# G2267-1KU).

2.5. Reduction and Denaturation of "Integral Membrane Protein" Samples

1. Integral Membrane Protein Denaturation Solution: 8 M high purity urea (480 mg urea) (*see* **Note 6**) and 2% CHAPS (20 mg CHAPS) in 1 mL of 50 mM ammonium bicarbonate, pH 7.5–8.5. Made fresh each time before use (Sigma, St. Louis MO, cat# U1250, C3023, and 09830) (*see* **Notes 1**, **17–20**).
2. Alternative membrane protein denaturation solution: 8 M guanidine-HCl solution with 2% CHAPS (20 mg CHAPS per 1 mL) (Sigma, St. Louis MO, cat# C3023, and Pierce, Rockford IL, cat# 24115) (*see* **Note 21**).
3. Bond-Breaker tris(2-carboxyethyl)phosphine (TCEP) neutralized pH solution (Pierce, Rockford IL, cat#77720) (*see* **Note 22**).

2.6. Trypsin Digestion and Alkylation of Proteins

1. 50 mM ammonium bicarbonate (*see* **Notes 1** and **19**).
2. 1 M calcium chloride (*see* **Notes 1** and **24**).
3. Sequencing Grade Modified Trypsin (Promega, Madison WI, cat# V5111).
4. 195 mM iodoacetamide (36 mg/1 mL of water). Made fresh right before use and placed in tin foil to limit exposure to light (*see* **Note 25**).
5. 20% Formic acid in water.
6. ColorpHast pH 0–14 pH indicator strips (Fisher Scientific, cat# S60169) (*see* **Note 10**).

2.7. SCX Clean-Up of "Membrane Associated Peptides"

1. HPLC grade Acetonitrile
2. 1 mL/100 mg Supelco Supelclean SPE SCX tubes (Supelco, St. Louis MO, cat# 504920).
3. Visiprep SPE Vacuum Manifold, 12 port (Supelco, St. Louis MO, cat# 57044)
4. KNF Loboport solid PTFE vacuum pump (Supelco, St. Louis MO, cat# Z26,226-9)
5. ColorpHast pH 0–14 pH indicator strips (Fisher Scientific, cat# S60169) (*see* **Note 10**).

6. 20% Formic acid (*see* **Note 1**).
7. SCX Wash Buffer: 10 mM ammonium formate in a 25:75 volume ratio of HPLC grade acetonitrile:water, adjusted to pH 3.0 with formic acid before addition of acetonitrile (*see* **Note 1**).
8. SCX Elution Buffer: 80:15:5 volume ratio of HPLC grade methanol:water:ammonium hydroxide (Fisher Scientific, cat# A669-500) (*see* **Note 1**).
9. Breathe-EASIER Tube Membrane (Diversified Biotech, Boston MA, cat# BMTM-2000) (*see* **Note 14**).

2.8. Reduction and Denaturation of "Whole Cell Lysate," and "Soluble Proteins"

1. Soluble protein denaturation solution: 8 M high purity urea in 50 mM ammonium bicarbonate (480 mg urea per 1 mL of 50 mM ammonium bicarbonate, pH 7.5–8.5). Make fresh right before use (Sigma, St. Louis MO, cat# U1250, C3023, and 09830) (*see* **Notes 1**, **8**, **9**, and **34**).
2. Alternative soluble protein denaturation solution: 8 M guanidine-HCl Solution (Pierce, Rockford IL, cat# 24115) (*see* **Note 21**).
3. Bond-Breaker tris(2-carboxyethyl)phosphine (TCEP) neutralized pH solution (Pierce, Rockford IL, cat#77720) (*see* **Note 22**).
4. Breathe-EASIER Tube Membrane (Diversified Biotech, Boston MA, cat# BMTM-2000) (*see* **Note 14**).

2.9. Trypsin Digestion and Alkylation of Soluble Proteins "Whole Cell Lysate," "Soluble Proteins," and "Carbonate Wash Proteins"

1. 50 mM ammonium bicarbonate (*see* **Note 1**).
2. 1 M calcium chloride (*see* **Notes 1** and **24**).
3. Sequencing Grade Modified Trypsin (Promega, Madison WI, cat# V5111).
4. 195 mM iodoacetamide (36 mg/1 mL of water). Made fresh right before use and placed in tin foil to limit exposure to light (*see* **Note 25**).
5. 20% Formic acid in water.
6. ColorpHast pH 0–14 pH indicator strips (Fisher Scientific, cat# S60169) (*see* **Note 10**).

2.10. Concentration and Clean-Up of Soluble Proteins (Whole Cell Lysate and Supernatant Fraction)

1. HPLC grade acetonitrile
2. 1 mL/100 mg LC-18 SPE tubes (Supelco, St. Louis MO, cat# 504270).
3. Visiprep SPE Vacuum Manifold, 12 port (Supelco, St. Louis MO, cat# 57044)
4. KNF Loboport solid PTFE vacuum pump (Supelco, St. Louis MO, cat# Z26,226-9)
5. C-18 Resin Preparation Buffer: 0.1% trifluoroacetic acid in water (*see* **Note 1**).

6. C-18 Wash Buffer: 5% acetonitrile and 0.1% trifluoroacetic acid in water (*see* **Note 1**).
7. C-18 Elution Buffer: 80% acetonitrile in water (*see* **Note 1**).
8. Breathe-EASIER Tube Membrane (Diversified Biotech, Boston MA, cat# BMTM-2000) (*see* **Note 14**).

2.11. Bicinchoninic Acid Assay to Determine Peptide Concentration

1. BCA Protein Assay Kit (Pierce, Rockford IL, cat# 23227) (*see* **Note 1**).
2. Siliconized, sterile individually wrapped microcentrifuge tubes, 1.5 and 2.0 mL (Fisher Scientific, cat# 02-681-332, cat# 02-681-331) (*see* **Note 2**).
3. Breathe-EASIER Tube Membrane (Diversified Biotech, Boston MA, cat# BMTM-2000) (*see* **Note 14**).

2.12. Strong Cation Exchange Offline Fractionation for Potential Mass Tag Development

1. Polysulfoethyl A 200 × 9.4 mm column, (5 μm, 300 Å) preceded by a 10 × 10 mm guard column (PolyLC, Columbia MD)
2. Mobile Phase A: 10 mM ammonium formate in 25% acetonitrile, pH 3.0
3. Mobile Phase B: 500 mM ammonium formate in 25% acetonitrile, pH 6.8

2.13. High Performance Liquid Chromotography and Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

1. Reversed-phase (C-18) pack capillaries (150 μm i.d. × 360 μm o.d., with 5 μm Jupiter brand C18 particles with 300-Å pore size) (Phenomenex, Torrance CA, and Polymicro Technologies, Phoenix, AZ, emitter made in house from tubing).
2. Mobile Phase A: 0.2% acetic acid and 0.05% trifluoroacetic acid in water (Sigma Aldrich, St. Louis MO, cat# A6283-500ML and T6508-10× 10AMP) (*see* **Note 1**).
3. Mobile Phase B: 0.1% trifluoroacetic acid in 90% acetonitrile and 10% water (Sigma Aldrich, St. Louis MO, cat# T6508-10× 10AMP and 34998-1L) (*see* **Note 1**).
4. Chemically etched open tubular electrospray emitter: 150 μm o.d. 20 μm i.d., fused silica tubing (Polymicro Technologies, Phoenix, AZ, emitter made in-house from tubing).
5. 49% (v/v) Hydrofluoric acid in water (Sigma Aldrich, St. Louis MO, 244279-25G) (*see* **Note 37**).
6. Zero dead-volume union (Valco Instruments, Houston TX, cat# ZU1XC-S6)
7. PEEK tubing (380 μm i.d.) (Upchurch Scientific, Oak Harbor, WA)
8. 2 μm pore sized steel screen for making micro bore LC columns (Valco Instruments, Houston TX)

3. Methods

The methods below, which are summarized in the flowchart seen in **Fig. 1**, focus on the preparation of samples for the relative quantification of proteins using mass spectrometry. The specific instrumentation used to obtain an abundance value is equally important and must be appropriate for quantitative analysis. In this chapter, we highlight the use of an LC-FTICR MS but other instruments which offer high mass measurement accuracy can be used (e.g., ThermoFinnigan Orbitrap MS, ThermoFinnigan LTQ-FT, or a quadrupole time-of-flight (Q-TOF) mass spectrometer).

The most fundamental requirement for label-free comparative proteomics is consistency in everything from sample preparation to analysis on the mass spectrometers. In preparing samples it is ideal to acquire all cells which will be compared and perform all sample processing at the same time. After the quantity of protein in each sample after lysis has been measured, the addition of a reference proteome is also advised. This can be any standard protein or peptide mixture which will not interfere with protein identification in the samples of interest. This internal reference verifies the success and efficiency of the trypsin digestion and can provide a measurable indicator of replicate reproducibility and change in the performance of the mass spectrometer in time. If a change in performance is detected the reference proteome can aid in the normalization of replicate data. With large multiple sample comparisons, replicate analyses of the same sample is crucial. For statistical purposes it is ideal to have at least 5 replicate LC-MS runs performed on the same column if possible. Depending upon how many samples are being compared or how much total sample is available, this amount of replication may not be practical. At the very least three replicate analyses of the same sample should be performed as long as the normalized replicate analyses prove to have a high correlation (i.e., $R^2 \geq 0.9$).

In addition to an internal reference, variability due to a sample change or LC-MS performance change in time can be minimized by using a Latin square experimental design to dictate the order in which the analyses are to be run. The Latin square design order minimizes systemic bias due to both possible carry over from run to run and a change in instrument performance over time. The Latin square design essentially places each sample to be compared into a block for each replicate (i.e., for four replicate analyses there would be four blocks). Within each block the order of the samples is randomized. This prevents any possible systemic bias due to sample carryover. The added benefit of using the Latin square design is that if an event such as an instrument failure occurs during a block, only that one block would need to be repeated instead of the entire study.

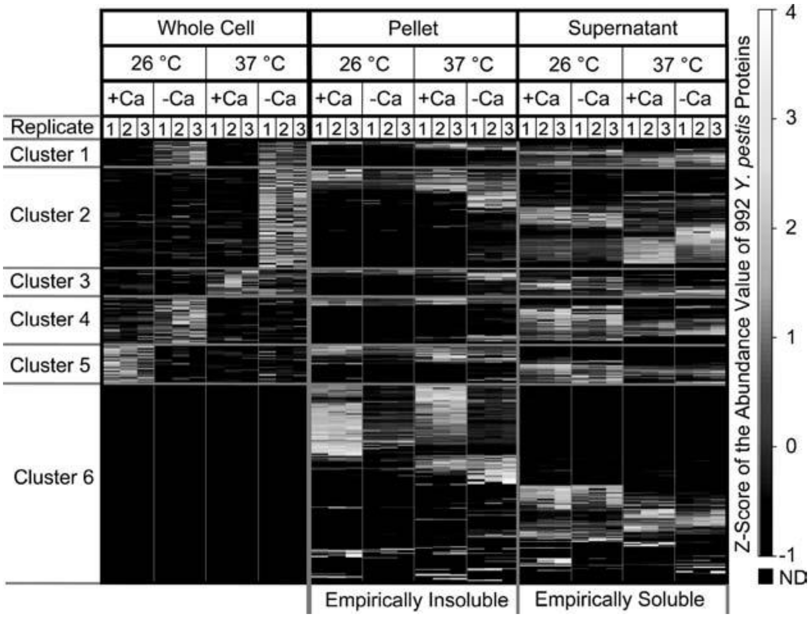


Fig. 2. The above experiment represents the protein abundance values obtained for four culture conditions of *Y. pestis* (the bacteria agent responsible for plague). The growth conditions were meant to mimic growth conditions of the bacteria when in the flea vector (ambient temperature) and after infection into the mammalian host (37°C with millimolar amounts of divalent calcium as experience in the bloodstream and extracellular fluid, and an absence of free divalent calcium as experience inside the host cells) The abundance values were transformed into *z*-scores for the whole cell set using the average protein value for the whole cell set. The *z*-score values were calculated separately for the pellet and supernatant data together using their average abundance. The abundance values thus are represented by color for 992 *Y. pestis* proteins identified in triplicate analyses of each sample. Significant protein abundance value changes due to culture condition and cell fractionation can easily be observed. *+Ca*$^{2+}$ growth with 4 mM Ca^{2+}; *−Ca*$^{2+}$ growth with no Ca^{2+}; *ND* not detected.

After several MS analyses have been performed for a given study it is usually necessary to normalize all replicate runs. There are many different normalization approaches for removing systemic biases specific to LC-FTICR MS data. The specific approach required may depend on the data at hand. The paper by Callister et al. highlights four normalization approaches that are applicable to LC-FTICR MS analyses *(18)*. Once the data is normalized, it is useful to display the significant protein changes in a color map after clustering together the data with similar protein abundance changes. This way the eye can quickly identify proteins of interest due to each growth condition investigated. Figures 2 and 3 show the color – protein associations obtained from the analyses of *Y. pestis* and *G. sulfurreducens.* These figures were produced using the clustering software OminVis™. This software is expensive and not readily available but similar clustering and colored protein association figures can be obtained from the free software MeV (http://www.tm4org/mev.html). MeV contains tools for doing statistical analyses (e.g., Student's *t* test, ANOVA, etc.) and for

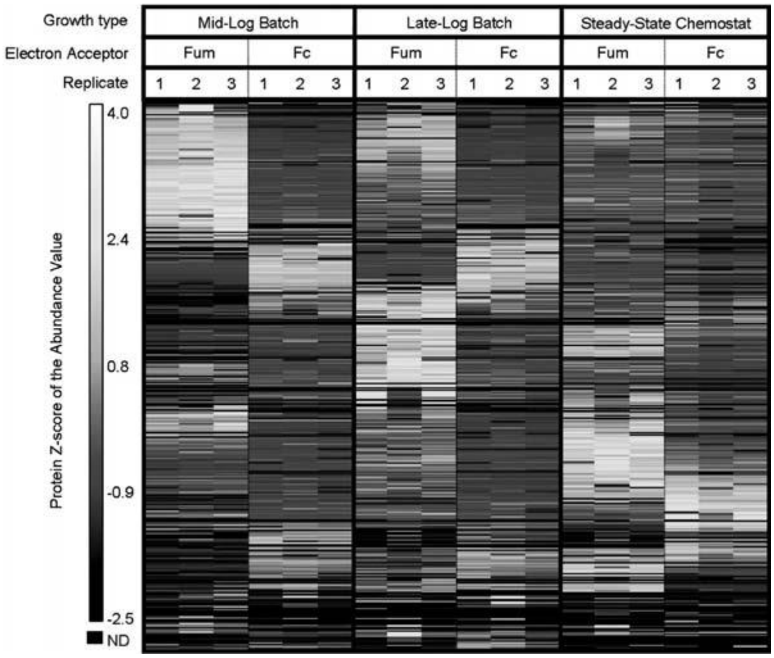


Fig. 3. *G. sulfurreducens* is an anaerobic soil bacteria that uses iron as an end electron acceptor for producing energy much like humans use oxygen as an end electron acceptor for energy production. The researchers here wanted to know what proteins were involved in reducing either fumarate (non-iron electron acceptor, Fum) or iron (studied here as iron citrate, Fc). Also they wanted to know if there was a difference in protein output due to different growth phases/conditions of the bacteria (mid-log phase, late-log phase, and growth in a chemostat). The relative *z*-score of each triplicate run from each culture condition is also shown side by side using the average value of each protein in all runs as the reference. The *darker colored* data represent proteins that have a significantly lower abundance than average. The *lighter color* represents higher *z*-scores and thus higher relative abundances for each observed protein. A *z*-score of 0 represents a protein abundance value that is equal to the average abundance value of all abundances for that given protein in the six culture conditions. Identifications represented in *black* are proteins that were not detected. A good resource is an online tutorial of analyzing multiple experiments with MeV (http://jbpc.mbl.edu/GenomesCourse/media/200510281330-braisted.pdf).

clustering or grouping together proteins with similar abundance change profiles (e.g., K-means, hierarchical, etc.). This software was made for microarray data but it can just as easily be applied to proteomics data that is formatted into a tab delimited file.

In the case of **Figs. 2** and **3**, the abundance values were transformed into a *z*-score. The *z*-score is useful for identifying protein changes that are significant. The *z*-score represents the units of standard deviation of the protein's abundance which is compared to the average value found for that protein in all analyses. The *z*-score transformation is useful for the comparison of samples when there is no specific reference sample or when there is a desire to see the overall protein change of multiple samples as compared with an average value observed for that protein. The *z*-score can be directly converted to a *p*-value as well. A *p*-value is a statistical term that is seen often in microarray data analyses. It is a measure of probability that a difference between groups (e.g., a control and a treated sample) during an experiment happens by chance. For example, a *p*-value of 0.01 ($p = 0.01$) means that there is 1 in 100 chance that the result occurred by chance. The lower the *p*-value the more likely it is that the difference between groups was caused by the treatment. The *p*-value can be obtained by MeV after doing a *t*-test or ANOVA analyses after data has been log (base 2) transformed and normalized.

3.1. Cell Lysis

1. Obtain cells at defined growth phase or state of interest (*see* Note 4).
2. Pellet the cells with centrifugation (10 min at 3220 rcf). Obtain an approximate pellet volume of ~150 μL for each cell set of interest that will later be compared.
3. Resuspend the 150 μL volume pellet in 450 μL of water (*see* Note 1).
4. Rupture the 600 μL cell/buffer suspension in a BioSpec mini-beadbeater for 180 s at 4,200 rpm with 0.1 mm zirconia/silica beads in a 1.5-mL sterilized siliconized microfuge tube (*see* Notes 5 and 6).
5. Immediately (while tube is still warm) after bead beating, carefully poke a hole in the bottom of the microfuge tube with a 26 guage needle by applying gentle pressure while twisting the needle back and forth while pointing the bottom of the microcentrifuge tube down toward an open 15 mL sterile capless Falcon tube (*see* Notes 7 and 8).
6. Place the microfuge tube full of lysate and beads into the 15 mL sterile capless Falcon tube.
7. Centrifuge the 1.5-mL tube in the Falcon tube at 805 rcf (805 × *g*) for 2 min.
8. Place extracted lysate contained in the Falcon tube on ice to inhibit proteolysis.
9. Divide lysate into a 2/3 and a 1/3 volume fraction to be used for the sub-cellular fractions ("integral membrane protein," "soluble protein," and the "carbonate wash protein") and the "whole cell lysate" protein digestion procedures, respectively (*see* Note 9).
10. Quick freeze "whole cell lysate" 1/3 volume fraction from Subheading 3.1, **step 9** and store at −80°C until ready for Subheading 3.8, **step 2** (*see* Note 9)

3.2. Subcellular Fractionation and "Integral Membrane Protein Enrichment"

1. Dilute the "2/3 volume lysate" fraction from Subheading 3.1, **step 9** (~400 μL) to about 1 mL with water (*see* Note 1).
2. Ultracentrifuge diluted lysate divided between two ultracentrifuge tubes at 100,000 rpm (355,040 avg. rcf) for 20 min.
3. Quickly after ultracentrifugation and carefully (without sucking up pellet) suck off supernatant from like tubes and combine the supernatants back together into a 1.5-mL siliconized,

sterile microcentrifuge tube. (This is your "soluble protein" fraction.) (*see* **Note 11**).

4. Keep supernatant on ice until you get to **Subheading 3.2**, **step 13**.
5. Add 50 mL of water to each pellet and sonicate (with a vigorous bath sonicator or probe sonicator) until you have an emulsion (*see* **Notes 1** and **12**).
6. To the emulsion add 600 µL of 100 mM sodium carbonate and gently pipette up and down to mix (*see* **Note 13**).
7. Incubate in an ice bath for 10 min.
8. Ultracentrifuge protein solution at 100,000 rpm (355,040 avg. rcf) for 10 min.
9. Carefully suck off supernatant. This is your "carbonate wash protein" fraction.
10. To the pellets add 50 µL of ice cold water into each tube and resuspend with vigorous sonication into an emulsion (*see* **Note 12**).
11. Add 600 µL of ice cold water to the emulsion and pipette up and down to mix.
12. Ultracentrifuge again at 100,000 rpm (355,040 avg. rcf) for 10 min.
13. Carefully suck off supernatant and add to the "carbonate wash protein" fraction from **Subheading 3.2**, **step 9** and add this combined supernatant to the supernatant from **Subheading 3.2**, **step 3**. This is your final "Soluble Protein" fraction.
14. Neutralize the combined carbonate wash fraction with enough 20% acetic acid until sample is about pH 8. Determine pH using a small amount of the sample on pH paper.
15. Quick freeze "Soluble Protein fraction" and store at −80°C until you get to **Subheading 3.8**, **step 1**.
16. To the pellets, add 50 mL of ice cold water and resuspend with vigorous sonication into an emulsion (*see* **Note 12**).
17. Combine the emulsion of like pellet samples together into a 1.5-mL siliconized, sterile microcentrifuge tube. This is your "Integral Membrane Protein" fraction.
18. If needing a stopping point, the "integral membrane protein" fraction can be quick frozen with liquid nitrogen and stored at −80°C until ready to do the next step.

3.3. Coomassie Blue Assay to Determine Protein Concentration

1. Make bovine serum albumin standards (BSA) of 0, 0.025, 0.125, 0.25, 0.5, 0.75, and 1.0 mg/mL using the 2 mg/mL bovine serum albumin standard supplied with the Coomassie Blue assay kit using sterile siliconized microcentrifuge tubes and diluting with water (*see* **Note 1**).

2. Gently shake your coomassie plus reagent solution (it can settle in time).
3. For Standards made in Subheading 3.3, **step 1**, mix 1 mL coomassie plus reagent + 50 μL of vortexed standard for each concentration in a clean, sterile, siliconized 1.5-μL microcentrifuge tube.
4. For each "integral membrane protein" sample, mix 1 mL coomassie plus reagent + 45 μL water + 5 μL of vortexed sample (*see* Note 1).
5. Vortex all assay tubes and incubate at room temperature for 10 min.
6. Transfer 1 mL to a 1 mL cuvette (if using a UV–vis spectrophotometer that uses cuvettes) or transfer 200 μL of each tube to a microplate well (if using a microplate reader).
7. Obtain your absorbance at 595 nm using the 0 mg/mL bovine serum albumin standard as your blank.
8. Calculate your sample concentrations using the BSA standard curve. Values should fall within the curve.
9. Multiply the sample concentration values obtained by 10 since they were a tenfold dilution compared with the standards. This is the sample concentration in mg/mL
10. Using a pipetter, determine the total volume of the sample (it should be close to 100 μL). Or for more accuracy, pipette an exact amount (e.g., 90–95 μL) of sample to a new 1.5-μL microfuge tube.
11. Multiply the volume by the sample concentration. This gives you an estimate of the total protein mass in the sample.

3.4. Internal Reference Proteome

1. Combine 1 mg apomyoglobin (equine), 0.363 mg cytochrome *c* (bovine) and 0.210 mg G-2-PDH (rabbit) per 1 μL total volume of water. This is your stock protein reference solution. Verify the concentrations of each standard with a BCA assay (*see* Notes 1 and 15).
2. Add 10 μL (15.73 μg) of the stock protein reference solution from Subheading 3.4, **step 1** to every 984.27 μg of sample (*see* Note 16).
3. Speed Vac samples to dryness (*see* Note 14).

3.5. Reduction and Denaturation of "Integral Membrane Protein" Samples

1. Add 100 μL of Integral Membrane Protein Denaturation Solution or the Alternative Membrane Protein Denaturation Solution if concerned about carbamylation (Make fresh right before use) (*see* Note 23).
2. Vortex or sonicate pellet into solution.

3. Add 1 µL Bond Breaker TCEP solution (TCEP solution in sample is now 5 mM).
4. Vortex.
5. Incubate protein solution for 30–45 min at 60°C.

3.6. Pellet Proteins (Insoluble/Membrane) Digestion Protocol

1. Dilute sample tenfold to 1.0 mL with 50 mM ammonium bicarbonate.
2. Add 2 µL of 1 M calcium chloride for every 100 µL of Integral Membrane Protein Denaturation Solution used in **Subheading 3.5**, **step 1**.
3. Add 40 µL of resuspension buffer (supplied with trypsin) to 20 µg of Sequencing Grade Modified Trypsin and vortex gently.
4. Add a 1:100 mass ratio of trypsin to sample (i.e., the mass calculated from **Subheading 3.3**, **step 11**) (*see* **Note 26**).
5. Incubate at 37°C for 4 h.
6. Intermittently during incubation vortex and sonicate any forming pellet back into solution.
7. Remove sample from incubator and quick freeze in liquid nitrogen and store at −80°C if needing a stopping point or proceed to the next step.
8. Speed vac down to ~100 µL.
9. Add a small amount of 20% formic acid to the sample until it is ≤pH 7.0. You can use a small amount of pH paper to check the pH (*see* **Note 27**).
10. Add 1 µL of Bond Breaker TCEP solution (TCEP solution in sample is now 10 mM).
11. Incubate for 5 min at 37°C.
12. Add 10 µL of 195 mM iodoacetamide. (Iodoacetamide concentration in sample is ~20 mM.)
13. Wrap sample in foil to avoid too much exposure to light (*see* **Note 28**).
14. Incubate at room temperature with rotation/mixing for 30 min.

3.7. SCX Clean-Up of "Membrane Associated Peptides"

1. Add 1 mL of HPLC grade acetonitrile to a 1 mL/100 mg LC-SCX SPE tube using the vacuum manifold to suck liquid through at 20 mmHg of vacuum stopping liquid just at the top of the resin.
2. Add 1 mL of SCX Wash Buffer stopping liquid just at the top of the resin.
3. Speed Vac all samples to dryness and add 1 mL of the SCX Wash Buffer and enough 20% formic acid (if needed to further reduce pH)to the sample so that pH is 3 (test with ~10–20 µL of sample on pH paper) (*see* **Notes 29** and **30**).

4. Centrifuge sample at 16,000 rcf for 1–2 min to spin down any precipitated material (ppt) and discard ppt (*see* **Note 31**).
5. Apply supernatant through the tube.
6. Add 6–8 mL of SCX Wash Buffer, stopping liquid just at the top of the resin (*see* **Note 32**).
7. Break the vacuum and put collection siliconized clean/sterile 1.5-mL microcentrifuge tubes below SPE tube to collect peptides into.
8. Add 2 aliquots of 300 μL SCX Elution Buffer to SPE tube stopping liquid just at the top of the resin.
9. Add 400 μL of acetonitrile and elute all liquid from the SPE tube.
10. Speed vac sample to dryness (*see* **Notes 14** and **33**).
11. Store at −80°C until ready for **Subheading 3.11**, **step 1**.

3.8. Reduction and Denaturation of Soluble Proteins: "Whole Cell Lysate," and "Soluble Proteins"

1. Remove the "Whole Cell Lysate," and "Soluble Protein" fractions from the freezer.
2. Speed vac the samples down to ~200 μL (*see* **Note 14**).
3. Perform all steps in **Subheading 3**, **steps 3** and **4** for each sample.
4. Speed vac to dryness.
5. Add 150 μL of Soluble Protein Denaturation or the Alternative Soluble Protein Denaturation Solution.
6. Add 1.5 μL of Bond Breaker neutralized TCEP solution.
7. Vortex and sonicate to resuspend lysate into solution.
8. Incubate at 60°C for 30–45 min.

3.9. Trypsin Digestion and Alkylation of Soluble Proteins: "Whole Cell Lysate," "Soluble Proteins," and "Carbonate Wash Proteins"

1. Dilute sample tenfold to 1.5 mL 50 mM ammonium bicarbonate (*see* **Note 23**).
2. Add 2 μL of 1 M calcium chloride for every 100 μL of Soluble Protein Denaturation Solution used in **Subheading 3.8**, **step 5** (3 μL total).
3. Add 40 μL of resuspension buffer (supplied with trypsin) to 20 μg of Sequencing Grade Modified Trypsin and vortex gently.
4. Add a 1:50 mass ratio of trypsin to sample (i.e., the mass calculated from **Subheading 3.3**, **step 11**) (*see* **Note 26**).
5. Incubate at 37°C for 4 h.
6. Remove sample from incubator and quick freeze in liquid nitrogen and store at −80°C if needing a stopping point or proceed to the next step.
7. Speed vac down to ~150 μL.

8. Add a small amount of 20% formic acid to the sample until it is ≤ pH 7.0. You can use a small amount of pH paper to check the pH (*see* **Note 27**).
9. Add 1.5 μL of Bond Breaker TCEP solution (TCEP solution in sample is now 10 mM).
10. Incubate for 5 min at 37°C.
11. Add 15 μL of 195 mM iodoacetamide. (Iodoacetamide concentration in sample is ~20 Mm.)
12. Wrap sample in foil to avoid too much exposure to light (*see* **Note 28**).
13. Incubate at room temperature with rotation/mixing for 30 min.

3.10. C-18 Clean-Up of "Soluble Protein" Peptides

1. Add 2 mL of HPLC grade acetonitrile to a 1 mL/100 mg Discovery DSC-18 SPE tube using the vacuum manifold to suck liquid through at 20 mmHg of vacuum stopping liquid just at the top of the resin.
2. Add 2 mL of C-18 Resin Preparation Buffer stopping liquid just at the top of the resin.
3. Centrifuge sample at 16000 rcf for 5 min. Discard pellet if there is one.
4. Apply supernatant through the tube.
5. Add 4–5 mL of C-18 Wash Buffer, stopping liquid just at the top of the resin.
6. Break the vacuum and put siliconized clean/sterile 1.5-mL microcentrifuge collection tubes below SPE tube to collect peptides into.
7. Add 2 aliquots of 300 μL C-18 Elution Buffer to SPE tube stopping liquid just at the top of the resin.
8. Add 300 μL of acetonitrile and elute all liquid from the SPE tube.
9. Speed vac sample to dryness (*see* **Notes 14** and **33**).

3.11. Bicinchoninic Acid Assay to Determine Peptide Concentration

1. Collect all peptide sample sets ("whole cell lysate peptides," "integral membrane peptides" from **Subheading 3.7**, **step 11**, and "soluble peptides"). If samples are dried down, resuspend with sonication and vortexing with 100 μL of water.
2. Perform BCA assay according to instructions in **Subheading 3.11**, **steps 4–13** below to determine peptide concentration and correct concentration with an additional speed vac session or addition of water to make 1.3 mg/mL (*see* **Notes 1** and **35**).
3. Perform BCA assay again according to instructions in **Subheading 3.11**, **steps 4–13** below to verify protein concen-

tration and adjust concentration again to be 1.0 mg/mL (*see* **Note 36**).

4. Make bovine serum albumin standards (if you don't already have them from **Subheading 3.3**, **step 1**) of 0, 0.025, 0.125, 0.25, 0.5, 0.75, and 1.0 mg/mL diluting with water (*see* **Note 1**).
5. Mix the 1:50 ratio of Reagent B to Reagent A from the BCA assay kit for the total amount of reaction reagent needed. Assume 1 mL for each sample and each standard.
6. For each sample and each standard place 1 mL of the combined reagents into a sterile/clean siliconized 1.5-mL microcentrifuge tube.
7. For each standard, add 50 μL of standard to each of the 1 mL of reagent.
8. For each sample, add 45 μL of water + 5 μL of sample to each of the 1 mL of reagent.
9. Vortex and incubate at 37°C for 30 min or 2 h at room temperature.
10. Transfer 1 mL to a 1 mL cuvette (if using a UV–vis spectrophotometer that uses cuvettes) or transfer 200 μL of solution from each tube to a microplate well (if using a microplate reader).
11. Obtain your absorbance at 562 nm using the 0 mg/mL standard as your blank.
12. Calculate your sample concentrations using the BSA standard curve. Values should fall within the curve.
13. Multiply the sample concentrations obtained by 10 since they were at a tenfold dilution compared with the standards.
14. Quick freeze with liquid nitrogen and store at −80°C until ready for LC-MS analysis.

3.12. Strong Cation Exchange Offline Fractionation for Reference Peptide Database

1. A database of all detectable peptides must be made using MS/MS data. To identify low abundant peptide it is necessary to do a fractionation of the sample offline that will be orthogonal to the downstream C-18 peptide separation that goes straight into the mass spectrometer.
2. Obtain 500 μg of peptide material from each sample or a pooled sample if sample is limited.
3. Concentrate the sample to dryness using a speed vac.
4. Add 450 μL of buffer A, sonicate into solution, and centrifuge at 16,000 × *g* for 3 min.
5. Inject the sample onto the strong cation exchange (SCX) column

6. Start the gradient isocratically at 100% buffer A for 10 min.
7. Then start a gradient from 0 to 50% buffer B over the next 40 min, increasing to 100% buffer B in the following 10 min.
8. At 20 min, hold buffer B at 100% to ensure elution of any remaining peptides.
9. Follow with a 30 min water wash of the column and re-equilibrate with 100% buffer A.
10. Determine when the peptides are coming out using a UV detector measuring 280 nm and collect 25 equal time spaced fractions.
11. Dry down the fractions in a speed vac and add 15 μL of water to resuspend the peptide to each fraction.
12. Vortex and sonicate into solution.
13. Run each fraction on a ThermoFinnigan LTQ mass spectrometer in MS/MS mode (*see* Subheading 3.12, **step 32**).

3.13. High Performance Liquid Chromatography and Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

1. It is important to have high quality electrospray emitters for high performance mass spectrometry using electrospray. Our lab makes these emitters in-house *(19)* by first taking a piece of 150 μm o.d., 20 μm i.d. fused silica tubing and burning about 1 cm of the coating off from one end.
2. Then hook up the fused silica tubing to a syringe pump that is pumping water at a flow rate of 0.1 μL/min. The water should be slowly emerging from the bare silica end.
3. Place about 1 mm length of the bare capillary vertically into a 49% aqueous hydrofluoric acid solution. The water flow prevents etchant from entering the capillary interior (*see* Notes 37 and 38).
4. Once the silica touching the hydrofluoric acid reservoir is completely removed, the etch process is automatically halted.
5. Once the etching is finished, rinse in water and cut to the desired length (we use 4–6 cm) from the unetched side.
6. LC columns were made "in-house" using a "in-house" made slurry packer (an analytical slurry packer is available from Alltech, Deerfield IL, cat# 1666) by the following method *(20)*.
7. Connect one end of the fused silica capillary (150 μm i.d. × 360 mm i.d × 60 cm long) to a zero dead-volume union using PEEK tubing (380 mm i.d., Upchurch) to position a steel screen having 2 μm pores.
8. At the other end of the capillary, connect it to the slurry packer which contains the C-18 packing material in a solvent mixture of acetonitrile/H_2O (90:10, v/v).

9. Pack the packing material into the column by gradually increasing the pressure from 500 to 10,000 psi.
10. Place column in an ultrasonic bath for about 5 min while still under pressure to help settle the packing material.
11. Allow column to depressurize overnight.
12. Connect the newly packed column to a high pressure LC system (one that can maintain 5,000–10,000 psi) and connect the end with the filter to the electrospray emitter. The LC system used in these studies was an automated LC system that consisted of a pair of syringe pumps (100-mL ISCO model 100DM) and controller (series D ISCO) and an in-house manufactured mixer, capillary column selector and sample loop. The mobile-phase selection valve was switched from position A to B 10 min after injection, creating an exponential gradient as solvent A was displaced by B in the mixer.
13. This work used a constant pressure LC system with a steel mixer containing a magnetic stirrer to mix solvents A and B. Use and optimize your system to provide a good separation starting with solvent A and increasing solvent B where the peptides will elute based upon increasing hydrophobicity. Flow through the capillary LC column in this study was ~1.8 μL/min when equilibrated to 100% solvent A.
14. In this study, a Bruker 9.4 T FTICR MS was used for the *Yersinia pestis* work and a ThermoElectron LTQ-FT was used for the *Geobacter sulfurreducens* work.
15. A total of 5 μg of total peptide material was analyzed for each FTICR analysis.
16. For MS/MS identification for building the reference database, samples were analyzed on a ThermoElectron LTQ.
17. Identify the MS/MS spectra using Sequest, Mascot, or X!tandam peptide searching software (X!tamdem is free and located at http://www.thegpm.org/TANDEM/)
18. Populate the reference peptide database with these identifications.
19. Match the FTICR data to the Potential Mass Tag (PMT) database using elution time and accurate mass *(5,14,21)*.
20. Use the FTICR MS summed ion current of peptide peaks as an abundance value for comparison. In this study the top 1/3 most abundant peptides for each protein were averaged for a final abundance value used in the comparison color figures.

4. Notes

1. Unless stated otherwise, all solutions should be prepared in water that has a resistivity of 18.2 MΩ cm and total organic content of less than five parts per billion. This standard is referred to as "water" in this text.
2. It is important to use this type of tube for two reasons. The siliconized tubes provide more flexibility under the stress of bead beating and proteins and peptides are less likely to stick irreversibly to the tubes. It is important that the tubes are sterile and individually wrapped since "bulk" tubes usually contain a lot of protein contaminates picked up in the manufacturing process (e.g., keratin).
3. Do not store any sample or liquid used in the sample processing in a falcon tube for longer than a day. In time polypropylene in the tubes will leach out into the sample causing polymer interference in the downstream mass spectrometric analysis.
4. Cells may be previously frozen or fresh. Ideally, there should be no protease inhibitors in the sample. If desired, cells may be washed with a phosphate buffered saline or other compatible buffer. The most important dogma from cell collection through the sample processing is that all samples are treated the same. If one cell set was previously frozen then all cells to be compared should also be frozen previously.
5. Fill beads to about 90% of tube after the addition of the cell suspension. Make sure not to have beads stuck under the lip of the tube or lysate will escape under increased pressure from heat caused by friction. Wipe such beads away with a moist Kimwipe and carefully close the tube.
6. Alternatively use mechanical lysis method of choice (i.e., sonication, French press, or Barocycler). The most important thing about the lysis choice is just that all cells that are to be compared are lysed with the same method.
7. Recover lysate from beads immediately while the tube is still warm because as the lysate cools down the proteins will start to stick to the beads. Be prepared for the lysate to initially squirt out after piercing with the needle due to the increased heat and pressure in the tube. Point the end that will be pierced toward the open end of the collection tube while piercing the microfuge tube.
8. If working with a pathogen, an alternative lysis method may be preferred due to chance of infection while working with needles. Recoveries using the bead beating method follow-

ing the needle pierce and centrifugation are quite good. If your alternative method of lysis does not have as good a recovery than you will have to lyse a larger amount of cells.

9. The sub-cellular fractionation isn't obligatory. The whole cell lysate proteomes by themselves can be analyzed. If this is desired then 100% of the lysate can be used and the protocol can be continued at **step 2** – **Subheading 3.7**.
10. Just the bottom square of the pH paper is needed to determine the approximate pH of the sample. This limits the amount of sample required for a pH determination.
11. The pellet is the insoluble fraction. This contains an enrichment of your membranes and some soluble protein contaminants or peripheral membrane proteins that have loose association with the membrane or due to the lysis and centrifugation end up stuck in the membrane fraction.
12. Most sonicating water baths have a suggested fill line. At this fill line the sonicating water is not vigorous enough to get the pellet into an emulsion. Instead add only about an inch of water to the sonicator bath or enough water to produce a rough and vigorous visual sonication and hold the sample contained in the ultracentrifuge tube down into what looks like the "sweet spot" of the sonication. Since water is splashing around it also helps to place a small piece of parafilm over the tube to keep the sample unadulterated.
13. The sodium carbonate solution is at a pH of roughly 11.5. This high pH serves to make most proteins express a net negative charge. This causes proteins that are not anchored into the membrane to lose attraction to all other proteins and pieces of DNA. Subsequently, proteins that are loosely associated to the membrane due to polar charge attraction are repelled from the integral membrane proteins *(22)*.
14. During the process of putting microcentrifuge tubes in and out of the speed vac there is a risk to accidentally knock over tubes and spill them. To avoid this and maintain the integrity of the samples it is optional but convenient to use the Breath-Easier Tube membranes.
15. The apomyoglobin after being added to the sample will be at 1% (w/w) total sample protein. Cytochrome *c* will be at twofold less in molarity than apomyoglobin and G-3-PDH will be 10 times less than apomyoglobin in terms of molarity. Apomyoglobin = 16,950 g/mol, Cytochrome *c* = 12,300 g/mol, and G-3-PDH = 35,600 g/mol. The total stock solution is 1.573 mg/mL.
16. This reference protein should not be used if working with mammalian samples of bovine, equine, or rabbit origin.

17. Urea when heated with proteins may cause carbamylation since urea in solution is in equilibrium with ammonium cyanate. The form isocyanic acid then can react with protein amino groups. Take this modification into consideration when determining your peptide sequence.
18. This protocol is for the alkylation of the cysteine residues using iodoacetamide. Alkylation of cysteines is necessary if identification of cysteine containing peptides is desired. If not alkylated, cysteines will randomly reform their disulfide bonds in time and these newly created species will not be identified. If cysteine containing peptide identification is not needed, alternatively you may use a denaturation solution made up of 7 M urea, 2M thiourea, 4% CHAPS in 50 mM ammonium bicarbonate, pH 7.5–8.5. Thiourea must not be used if alkylating because the thiol groups will act as a scavenger for the alkylating reagent.
19. The 50 mM ammonium bicarbonate must be made and used within 1 week. As the ammonium ions evaporate the bicarbonate anion dissociates to a carbonate anion which shifts the pH upward over time. Urea also should be prepared fresh for best results.
20. Urea and heat all serve to unfold the proteins from their globular state and keep them open. This will allow trypsin (when added later) to access the cleavage sites at the carboxyl end of arginine and lysine.
21. Guanidine-HCl solution works well as a denaturant but it contains primary amine groups so if you are going to do any downstream work with reactions involving primary amine (e.g., biotinylations), you need to use the urea denaturant instead.
22. The Bond-Breaker TCEP solution is already neutralized so that you will not have to neutralize it yourself. It is convenient in that it is quite stable, is not reduced by air, works in the pH ranges of 4–9 and does not need to be made up fresh like dithiothreitol; so it's easier to use.
23. Increase volume of denaturation solution if needed for good solubilization. Scale up the addition of the following reagents accordingly.
24. Calcium chloride serves to conserve the correct conformation of trypsin in order for it to be active.
25. Iodoacetamide reacts with sulfhydryl groups (–SH left after the disulfide bonds were broken from the TCEP reduction). The –SH group on cysteines are replaced by $-SCH_2CONH_2$. Keep in mind that this will increase the cysteine mass by about 57 Da.

26. Ratios between 1:20 and 1:100 are fine as long as you are consistent with all samples that you want to compare.
27. Thiol reactive reagents like iodoacetamide are capable of reacting with amines as well as thiols, thus it is preferable to keep the pH at or below pH 7.0 during the alkylation step. This increases the likelihood that the amines are protonated and therefore nonreactive.
28. The alkylation reagent iodoacetamide is light sensitive and may form unwanted adducts in the presence of much light. Iodoacetamide blocks the reduced thiol groups on the cysteines from reforming disulfide bonds. Alkylation is optional since the proportion of cysteine containing peptides in prokaryotes is relatively small.
29. This step is making most peptides in the sample have a net positive charge which will allow them to stick to the cation exchange resin which is negatively charged.
30. At this step a lot of ppt may occur. This ppt material is primarily lipid materials and other cellular debris that becomes insoluble at a low pH. The soluble peptides from the integral membrane proteins will still be in solution.
31. This ppt will plug the SPE tube if you attempt to put it through so use only the supernatant.
32. This step is removing the CHAPS and salts from the peptides which are still stuck to the resin in the column.
33. Alternatively, you can speed vac the sample to ~100 μL. Treat all comparable samples the same. If you speed vac to dryness for one then do it for all. If you speed vac to dryness use sonication and vortexing vigorously to get peptides back into solution and off the walls of the microcentrifuge tube.
34. This protocol is for the alkylation of the cysteine residues using iodoacetamide. Alkylation of cysteines is necessary if identification of cysteine containing peptides is desired. If not alkylated, cysteines will randomly reform their disulfide bonds in time and these newly created species will not be identified. If cysteine containing peptide identification is not needed, alternatively you may use a denaturation solution made up of 7 M urea, 2 M thiourea, in 50 mM ammonium bicarbonate, pH 7.5–8.5. Thiourea must not be used if alkylating because the thiol groups will act as a scavenger for the alkylating reagent.
35. The Coomassie Blue assay used previously cannot be used here because it does not effectively detect peptides. Conversely, we do not use the BCA assay when determining the protein concentration in the lysates because the membranes

in the sample interfere with the assay to give falsely high concentration values. The BCA assay is sensitive to reducing agents or anything with a thiol group. Therefore the samples must be cleaned-up prior to its use.

36. This second BCA assay seems tedious and redundant but if the concentrations vary quite a bit from the first assay, the corrected concentrations may not be quite the same after dilution. The second assay aids in making the samples as close to the same concentration as possible. This is essential for label-free relative quantitative proteomics.
37. Hydrofluoric acid (HF) is extremely hazardous and corrosive. Extreme care must be taken to prevent exposure to HF liquid or vapor. HF solutions should be used in a ventilated hood, and appropriate protective equipment should be worn.
38. This tip is designed such that there is no taper (like with pulled tips) which lessens the chance of a plugged tip during electrospray.

References

1. Adkins, J.N. et al. (2006) Analysis of the *Salmonella typhimurium* proteome through environmental response toward infectious conditions. *Mol Cell Proteomics* 5, 1450–1461.
2. Logan, B.E. & Regan, J.M. (2006) Electricity-producing bacterial communities in microbial fuel cells. *Trends Microbiol* 14, 512–518.
3. Hunter, E.M., Mills, H.J. & Kostka, J.E. (2006) Microbial community diversity associated with carbon and nitrogen cycling in permeable shelf sediments. *Appl Environ Microbiol* 72, 5689–5701.
4. van der Heijden, M.G. et al. (2006) Symbiotic bacteria as a determinant of plant community structure and plant productivity in dune grassland. *FEMS Microbiol Ecol* 56, 178–187.
5. Zimmer, J.S., Monroe, M.E., Qian, W.J. & Smith, R.D. (2006) Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass Spectrom Rev* 25, 450–482.
6. Pasa-Tolic, L., Masselon, C., Barry, R.C., Shen, Y. & Smith, R.D. (2004) Proteomic analyses using an accurate mass and time tag strategy. *Biotechniques* 37, 621–624, 626–633, 636 passim.
7. Callister, S.J. et al. (2006) Application of the accurate mass and time tag approach to the proteome analysis of sub-cellular fractions obtained from *Rhodobacter sphaeroides* 2.4.1. Aerobic and photosynthetic cell cultures. *J Proteome Res* 5, 1940–1947.
8. Fang, R. et al. (2006) Differential label-free quantitative proteomic analysis of *Shewanella oneidensis* cultured under aerobic and suboxic conditions by accurate mass and time tag approach. *Mol Cell Proteomics* 5, 714–725.
9. Umar, A., Luider, T.M., Foekens, J.A. & Pasa-Tolic, L. (2006) NanoLC-FT-ICR MS improves proteome coverage attainable for 3000 laser-microdissected breast carcinoma cells. *Proteomics* 7, 323–329.
10. Resch, W., Hixson, K.K., Moore, R.J., Lipton, M.S. & Moss, B. (2006) Protein composition of the vaccinia virus mature virion. *Virology* 358, 233–247.
11. Shi, L. et al. (2006) Proteomic analysis of *Salmonella enterica* serovar typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar typhimurium inside macrophages. *J Biol Chem* 281, 29131–29140.
12. Qian, W.J. et al. (2005) Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using 16O/18O labeling and the accurate

mass and time tag approach. *Mol Cell Proteomics* 4, 700–709.

13. Bogdanov, B. & Smith, R.D. (2005) Proteomics by FTICR mass spectrometry: top down and bottom up. *Mass Spectrom Rev* 24, 168–200.
14. Eng, J.K., McCormack, A.L. & Yates, J.R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5, 976–989.
15. Perkins, D.N., Pappin, D.J., Creasy, D.M. & Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567.
16. Hixson, K.K. et al. (2006) Biomarker candidate identification in Yersinia pestis using organism-wide semiquantitative proteomics. *J Proteome Res* 5, 3008–3017.
17. Ding, Y.H. et al. (2006) The proteome of dissimilatory metal-reducing microorganism *Geobacter sulfurreducens* under various growth conditions. *Biochim Biophys Acta* 1764, 1198–1206.
18. Callister, S.J. et al. (2006) Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J Proteome Res* 5, 277–286.
19. Kelly, R.T. et al. (2006) Chemically etched open tubular and monolithic emitters for nanoelectrospray ionization mass spectrometry. *Anal Chem* 78, 7796–7801.
20. Shen, Y. et al. (2001) Packed capillary reversed-phase liquid chromatography with high-performance electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for proteomics. *Anal Chem* 73, 1766–1775.
21. Jaitly, N. et al. (2006) Robust algorithm for alignment of liquid chromatography-mass spectrometry analyses in an accurate mass and time tag data analysis pipeline. *Anal Chem* 78, 7397–7409.
22. Fujiki, Y., Hubbard, A.L., Fowler, S. & Lazarow, P.B. (1982) Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J Cell Biol* 93, 97–102.

Chapter 4

Classical Proteomics: Two-Dimensional Electrophoresis/MALDI Mass Spectrometry

Ursula Zimny-Arndt, Monika Schmid, Renate Ackermann, and Peter R. Jungblut

Summary

The rapid development in proteomics over the last 10 years has led to a series of new technologies and combinations of them designed to unravel as much as possible of the proteins of an organism or otherwise specified biological material. Despite being a little tricky at certain steps, 2-DE has a very high resolution power with more than 10,000 spots per gel and is able to separate one protein into its different protein species caused by posttranslational modifications, alternative splicing or genetic variability. This high-resolution separation is combined with a highly sensitive identification method using peptide mass fingerprinting combined with sequence information by MS/MS, which results in high sequence coverage: the key to elucidate protein species structures. The off-line measurement by MALDI-TOFTOF-MS allows the repeated measurement of each sample and therefore provides more complete structure information for each protein species. The presented protocols represent the basic technology consisting of 2-DE, two staining methods, tryptic digestion and MALDI-TOFTOF-MS.

Key words: Isoelectric focusing, Carrier ampholytes, SDS–polyacrylamide gel electrophoresis, Peptide mass fingerprint; Protein species, MS/MS ion search, TOFTOF-MALDI-MS, Proteome database.

1. Introduction

With the first two-dimensional electrophoresis gels by Macko and Stegemann, 1969 *(1)* and Kaltschmidt and Wittmann, 1970 *(2)* the idea of separating and analyzing complete organisms, or at least complete protein complexes such as the ribosome in screening-like experiments, was realized. An important step forward in resolution was reached by combining isoelectric focusing using carrier ampholytes *(3)* with SDS–polyacrylamide gel electrophoresis of

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_4

Laemmli *(4)* by O'Farrell in 1975 *(5)*. This combination was optimized in detail by Klose and Kobalz *(6)* to reach a resolution of up to 10,000 spots per gel in a reproducible procedure. In parallel a completely commercialized method combining immobilized ampholytes with SDS–PAGE was developed *(7)*.

For a long time identification of proteins from 2-DE gels was a difficult task. First antibodies *(8)*, then Edman degradation *(9)* and amino acid analysis *(10)* were applied. The development of soft ionization procedures in mass spectrometry such as matrix assisted laser desorption/ionization (MALDI) *(11)* and electrospray ionization (ESI) *(12)* was a breakthrough. Peptide mass fingerprints after enzymatic digestion are ideal for the identification of proteins *(13)*. Confirmation and complementation of identification was obtained by MS/MS fragmentation of the peptides, which was improved in MALDI-MS with the TOFTOF machines *(14)*. A flow schema of the methods presented here is shown in **Fig. 1**.

A large size gel system is used to reach a high resolution: 23 cm (IEF) × 30 cm (SDS–PAGE). The IEF dimension can be extended to 40 cm. In this case the IEF gel has to be divided into two parts for application onto the SDS–PAGE gels *(6, 15)*. Analytical gels are produced with 0.9-mm thick gels in IEF and 0.75-mm thick gels in SDS–PAGE. Preparative gels for identification by MS are produced with 1.5-mm thick gels in both dimensions or may even be extended to 2.5 mm in the first dimension to increase the number of identifiable spots. A prerequisite for optimal separation of protein species is the sample preparation (*see* **Note 1**), which depends on the nature of the sample

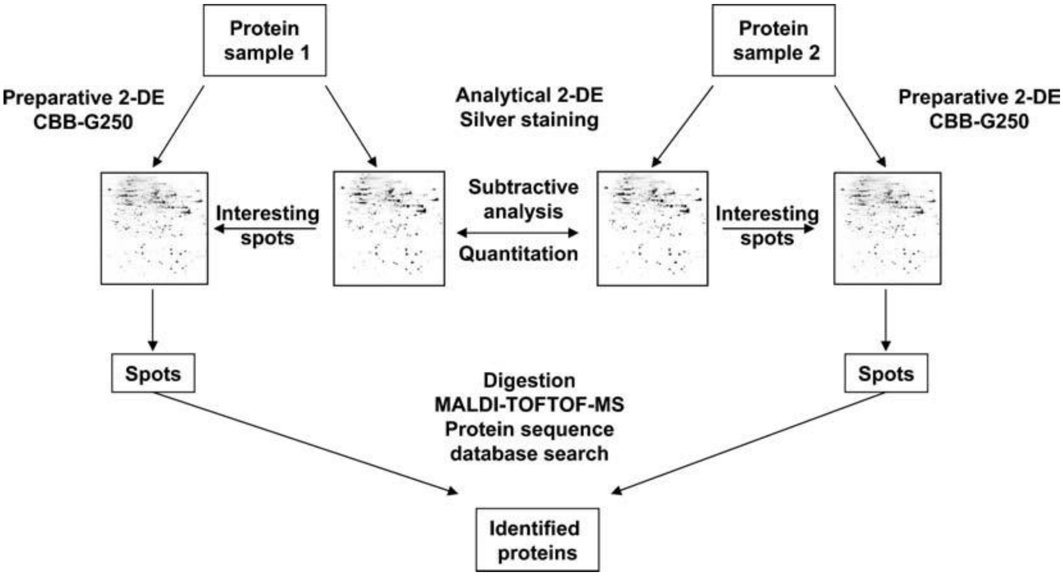

Protein sample 1
Protein sample 2
Preparative 2-DE CBB-G250
Analytical 2-DE Silver staining
Preparative 2-DE CBB-G250
Interesting spots
Subtractive analysis
Quantitation
Interesting spots
Spots
Digestion
MALDI-TOFTOF-MS
Protein sequence database search
Spots
Identified proteins

Fig. 1. Flow schema of the 2-DE-MALDI-TOFTOF-MS technology for identification of proteins and protein species.

and is therefore not described here. An outline of solubilization of proteins in 2-DE was published recently *(16)*. Analytical gels are stained with silver *(17)* to gain high sensitivity and preparative gels are stained with Coomassie Brilliant Blue *(18)* to have optimal compatibility with MALDI-MS *(19)*. The presented detection procedures may be complemented by a series of staining protocols reviewed recently *(20)*. With about 1,800 spots a representative amount of proteins and protein species can be detected for *Helicobacter pylori* cellular proteins *(21)* (**Fig. 2**).

The main advantage of the 2-DE/MALDI-MS approach is the separation of the proteins into their protein species in the top-down like separation step 2-DE (*see* **Note 2**). Each protein species may be analyzed separately, which allows the hypothesis-free identification of posttranslational modifications *(22)*. The bottom-up like identification by peptide mass fingerprinting benefits from the high sensitivity in the fmol range and high mass accuracy of peptide detection lower than 20 ppm by MALDI-MS (**Fig. 3**). TOFTOF-MS increases the security of identification and raises the chances of identifying posttranslational modifications. Several examples of proteome projects using the 2-DE/MALDI-MS approach are presented in the 2D-PAGE proteome database

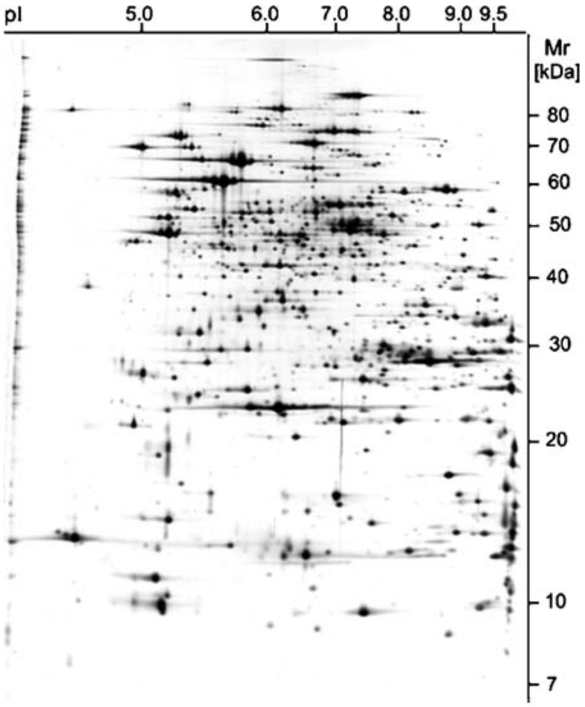


Fig. 2. 2-DE pattern of *Helicobacter pylori* 26695 cellular proteins. Sixty micrograms of protein was applied to analytical 2-DE and the protein species were stained with silver. Calibration of *Mr* and p*I* was performed with the calibration program in the 2D-PAGE database portal of MPIIB (http://www.mpiib-berlin.mpg.de/2D-PAGE/), where information about identity of most of the spots is also available.

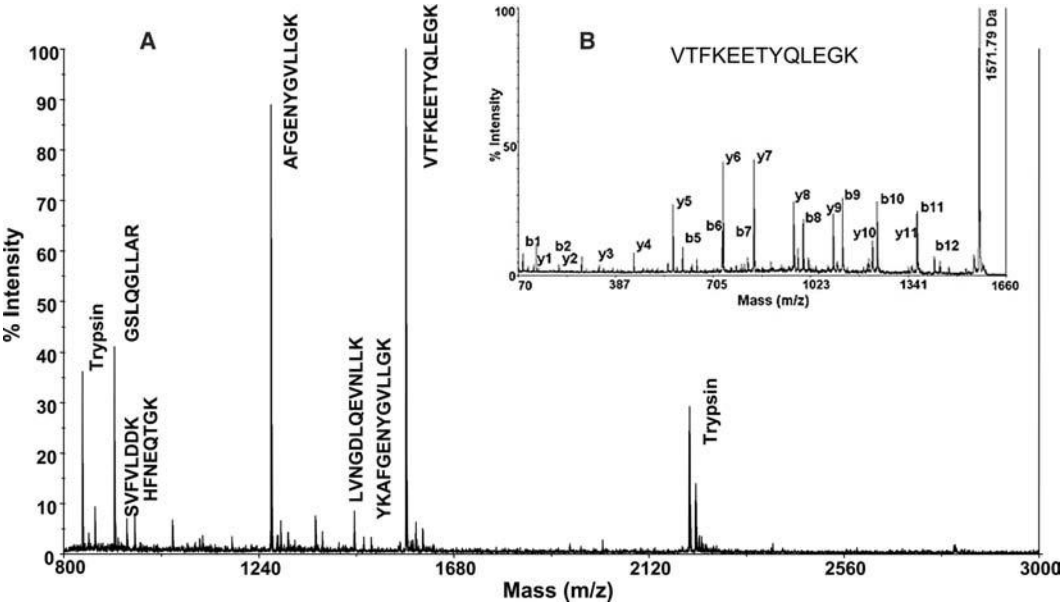


Fig. 3. MALDI-TOFTOF-MS analysis of spot 550 from *Helicobacter pylori I* 26695 cellular proteins, separated by 2-DE. (**a**) MS. The labeled peaks correspond to peptides of protein HP0390, adhesin-thiolperoxidase (TagD), resulting in a sequence coverage of 39%. (**b**) MS/MS of the peptide with the mass 1,571.79. The labeled peaks represent fragment ions (y1–yx and b1–bx) corresponding to the sequence VTFKEETYQLEGK.

portal (http://www.mpiib-berlin.mpg.de/2D-PAGE/), which is a first step towards standardization of procedures.

2. Materials

2.1. Equipment

Equipment for ampholyte 2-DE can be produced partially by a good laboratory shop or may be purchased from WITA (Teltow, Germany), a company which also proceeds in developing the system further.

2.1.1. IEF

1. Gel casting apparatus (WITA). The gel casting apparatus allows the preparation of eight IEF gels (**Fig. 4**).
2. Glass tubes for IEF gels: 0.9 or 1.5 mm inner diameter, length 23 cm, markings 3 and 9 mm from the bottom, 21, 24 and 30 mm from the top. The glass tubes may be purchased from Schott Glaswerke, D-55014 Mainz, Germany or already with markings from WITA. High quality tubes with exact inner diameters are a prerequisite for perfect gel casting and gel removal out of the tubes. Many attempts to substitute these high quality tubes with cheap plastic tubes did not succeed.

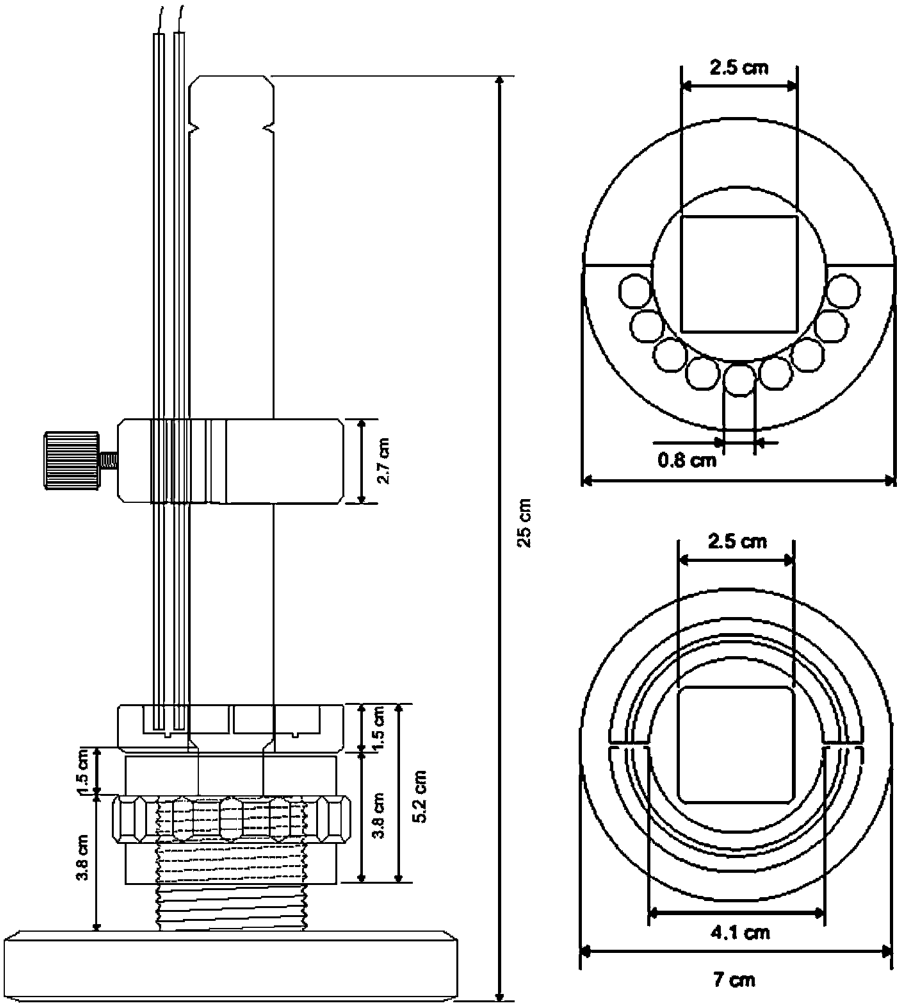


Fig. 4. Gel casting apparatus for IEF gels.

3. Polypropylene (PP) threads with 0.8 or 1.4 mm diameter and sealed edges to pour the gels and remove them from the tubes.
4. IEF gel chamber (WITA). Within the gel chamber eight gels can be run in parallel. For anodic sample application the upper chamber contains the anodic solution and the lower chamber the cathodic solution (**Fig. 5**).
5. Power supply for low currents and high voltages (2,000 V) and programmable (e.g., Electrophoresis Power Supply EPS 3501XL, GE Healthcare, Freiburg, Germany).
6. Instruments for sample loading: Gel loader tips or 10-µL Hamilton syringes (type 701 RN + Chaney, Hamilton, Bonaduz, Switzerland) with rounded tips, especially for thin gels.

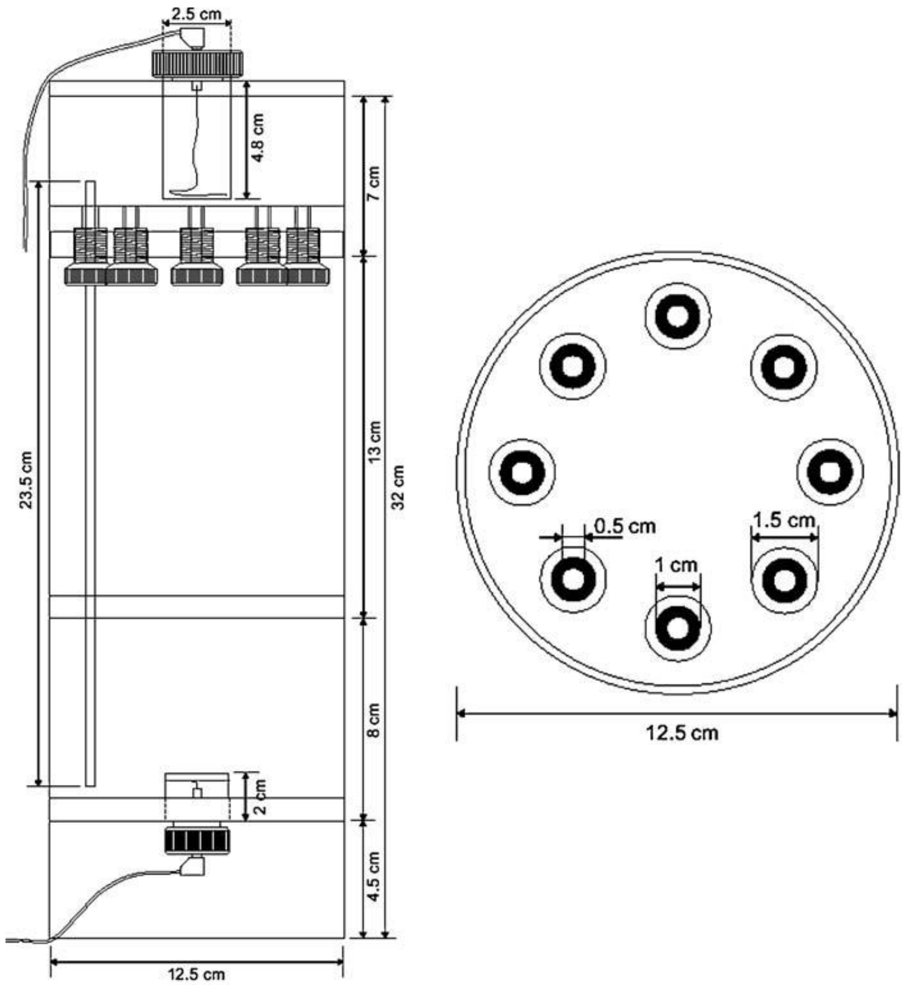


Fig. 5. IEF gel chamber. The intermediary cylinder between the upper and the lower chamber can have different lengths to adapt to smaller or larger gels.

7. For gel preparation the following materials have to be provided before starting: 2 × 975 μL separation gel solution (WITA or self made), 390 μL cap gel solution (WITA or self made), 0.8% ammonium persulfate solution (APS), deionized water, Gilson pipettes 0–20, 0–200, 0–1,000 μL, gel loader tips, pump for degassing of gel solution, timer, Parafilm cut into small squares, strips of filter paper (width < 1.5 mm) and ruler.
8. For sample application and running of the gels the following materials have to be provided before starting: anode solution, cathode solution, prepared sample, Sephadex solution, urea, mixture of ampholytes (pH 2–11), overlaying solution,

deionized water, 0–20- and 0–200-µL Gilson pipettes, Hamilton syringe for sample application, Vortex, filter paper strips, narrower than 1.5 mm.

2.1.2. SDS–page

1. SDS–PAGE chamber DESAPHOR VA 300 (DESAGA, Heidelberg, Germany) including clamp assemblies and casting stand with rubber gasket on the bottom end. The rubber gasket and the casting stand have to be perforated, if the glycerol underlaying procedure described here is to be applied.
2. Power supply (1,500 V; 150 mA, at least 200 W).
3. For IEF gel equilibration the following materials have to be provided before starting: equilibration solution, 87% glycerol, deionized water, Petri dishes, diameter 8 cm, 10-mL glass pipette with Peleus ball, polypropylene (PP) threads, diameter 0.8 or 1.4 mm.
4. For preparation of SDS–PAGE gels the following materials have to be provided before starting: 70% (v/v) ethanol, 40% (v/v) glycerol, deionized water, Kimwipes, four glass plates, four spacers 0.75 or 1.5 mm, four fitting rubber sealings or alternatively Silicone, ruler, Pasteur pipettes, agarose solution, for two gels: 1 (0.75-mm thick gels) or 2 (1.5-mm thick gels) × 135 mL of gel solution (WITA), 1 or 2 × 9 mL 1.28% APS, timer, water bath with thermometer.
5. For sample loading the following materials have to be provided before starting: agarose solution (40°C), electrode buffer (10 L), gels of first dimension (IEF gels), filter paper, and small spatula.
6. During SDS–PAGE the fixation solution has to be provided.

2.1.3. Coomassie Brilliant Blue G250 Staining

For this procedure, staining dishes for large gels, opaque white, without grooves 430 mm × 530 mm × 105 mm, ammonium sulfate, Coomassie Brilliant Blue G250 (CBB G250), deionized water, methanol and phosphoric acid 85% are required.

2.1.4. Silver Staining

The following materials are required for this procedure: staining dishes for large gels, opaque white, without grooves 430 mm × 530 mm × 105 mm, deionized water of highest quality, Milli Q Synthesis System, conductivity 18.2 MΩ, EDTA (Titriplex III) p.a., acetic acid 100% p. a., ethanol absolute 1% MEK denatured, formaldehyde 37% stabilized, acid-free, glutaraldehyde for electron microscopy 25% Grade II, store at 4°C, sodium acetate anhydrous p.a., sodium carbonate p.a. anhydrous, sodium hydrogencarbonate p.a. anhydrous, sodium thiosulfate – pentahydrate p.a., silver nitrate GR p.a. ACS ISO, Thimerosal USP, Merck-Schuchardt 8.17043.

2.1.5. Tryptic Digestion

1. Microconcentrator 5301 (Eppendorf).
2. Solutions: destaining buffer; digestion buffer; 0.1 μg/μL trypsin; 60% ACN, 0.5% TFA; 100% ACN; 33% ACN, 0.1% TFA; CHCA matrix.

2.1.6. MALIDI-MS

1. 4700 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA) or 4800 Proteomics Analyzer.
2. Zip Tips: Micro ZipTip U-C18 (Millipore) ZTC18M096 (0.2 μL bed volume) and ZipTip C18 (Millipore) ZTC18S096 (0.6 μL bed volume).

2.2. Buffers and Solutions

2.2.1. IEF

Stock solutions

1. Washed ion-exchange material Serdolit (for purification of acrylamide). Suspend 7 g Serdolit MB1 (Serva 40701) in 700 mL of deionized water and pour through a glass fiber filter (Whatman 1822047 Glass Microfibre Filters).
2. Acrylamide/PDA for separation gel: 17.5% (m/v) acrylamide, 1.5% (m/v) piperazine diacrylamide (PDA). Weigh out 17.5 g acrylamide and 1.5 g PDA. Fill up to 100 mL with deionized water. Add 5 g (wet weigh) of Serdolit suspension as prepared in 1. Stir with a blade agitator for 1 h avoiding light. Filtrate through glass fiber filter and freeze 16 mL aliquots for storage at −20°C.
3. Acrylamide/PDA for cap gel: 60% (m/v) acrylamide, 0.65% (m/v) PDA. Weigh out 18 g acrylamide and 0.195 g PDA. Fill up to 30 mL with deionized water. Add 1.5 g (wet weigh) of Serdolit suspension as prepared in 2. Stir with a blade agitator for 1 h avoiding light. Filtrate through glass fiber filter and freeze 7 mL aliquots for storage at −20°C.
4. Glycerol: 28.6% (m/v). Weigh out 14.3 g glycerol. Fill up to 50 mL with deionized water.
5. TEMED: 0.6% (v/v). Measure 0.12 mL TEMED and fill up to 20 mL with deionized water.
6. Ampholyte mixture pH 2–11, 40% (*see* **Notes 3** and **4**). Mix 8 mL Servalyt (R) 2–11; 8 mL Ampholine 3.5–10.0 broad range; 24 mL Pharmalyte 4.0–6.5 broad range; 16 mL Pharmalyte 5.0–8.0 broad range; and 8 mL Ampholine 7.0–9.0 narrow range. Freeze in aliquots of 8 mL at −80°C.
7. IEF separation gel: 9 M urea, 5% (m/v) glycerol, 3.5% (m/v) acrylamide, 0.3% (m/v) PDA, 0.06% (m/v) TEMED, 4% (m/v) ampholyte mixture 2–11, 0.02% (v/v) APS (**Table 1**). Add all solutions except APS to urea. Solubilize in water bath at 37°C. Filter through 0.45 μm surfactant-free cellulose acetate filter (Nalge Nunc International 191-2045) and freeze in 1 mL aliquots at −80°C.
8. IEF cap gel: 9 M urea, 5% (m/v) glycerol, 12% (m/v) acrylamide, 0.13% (m/v) PDA, 0.06% (m/v) TEMED, 4% (m/v)

Table 1
IEF separation gel

Amount		
1 mL	**75 mL**	**Substance**
0.54 g ≅0.4 mL	40.5 g ≅30.0	Urea
0.175 mL	13.13 mL	Glycerol 28.6%
0.2 mL	15.0 mL	Acrylamide 17.5%/PDA 1.5%
0.10 mL	7.5 mL	TEMED 0.6%
0.10 mL	7.5 mL	Ampholyte Mixture 2–11 40%
0.025 mL[a]	1.875 mL[a]	Ammoniumperoxodisulfate (APS) 0.8%

[a] To be added directly before gel preparation (only mentioned here to be included in the calculation of the final concentrations of the substances within the gel)

Table 2
IEF cap gel

Amount		
1 mL	**30 mL**	**Substance**
0.54 g ≅0.4 mL	16.20 g ≅ 6.0 mL	Urea
0.175 mL	5.25 mL	Glycerol 28.6%
0.20 mL	6.00 mL	Acrylamide 60%/PDA 0.65%
0.10 mL	3.00 mL	TEMED 0.6%
0.10 mL	3.00 mL	Ampholyte Mixture 2–11, 40%
0.025 mL[a]	0.750 mL[a]	Ammoniumperoxodisulfate (APS) 0.8%

[a]To be added directly before gel preparation (only mentioned here to be included in the calculation of the final concentrations of the substances within the gel)

ampholyte mixture pH 2–11, 0.02% (v/v) APS (**Table 2**). Add all solutions except APS to urea. Solubilize in water bath at 37°C. Filter through 0.45-μm cellulose acetate filter and freeze in 400 μL aliquots at −80°C.

9. Sephadex gel stock suspension. Suspend while softly stirring 4 g Sephadex G-200 superfine in 100 mL deionized water and swell for 5 h at 90°C. Chill and precipitate Sephadex and remove supernatant. Add 200 mL of 25% glycerol and stir for 1 h with a blade agitator. Remove supernatant and repeat the glycerol wash. Dry by filtrating in a filtration bottle (1,000 mL, Nalgene 513A3003) without vacuum to obtain a viscous

Sephadex suspension, which still can be pipetted. For storage freeze 9 mL aliquots of this suspension at –20°C. For immediate use, to the 9 mL of this viscous Sephadex suspension add 1 mL 1.4 M dithiothreitol (DTT) to reach a final concentration of 0.14 M DTT. Freeze in 100 μL aliquots at –20°C. To pipette 100 μL of this viscous suspension it is recommended that the 100 μL level at the pipette tip be marked; use 200 μL adjustment at the pipette and soak up to the marking at the pipette tip for optimal adjustment.

Final Solutions

1. APS: 0.8%. Weigh out 0.8 g of APS, fill up to 100 mL with deionized water and filtrate through 0.45-μm cellulose acetate filter. Divide into 100 μL aliquots; store frozen at –20°C.
2. IEF-separation gel (T = 3.8%, C = 7.9%): 3.5% (w/v) acrylamide, 0.3% (w/v) piperazine diacrylamide, 4% ampholyte mixture pH 2–11, 9.0 M urea, 5% (w/v) glycerol, 0.06% (v/v) TEMED, 0.02% (w/v) APS. Shortly before application, mix 4 × 975 μL/2 × 975 μL of *degassed* separation gel solution with 25 μL of 0.8% APS for each 975 μL aliquot.
3. IEF cap gel (T = 12.1%, C = 1.1%): 12% (w/v) acrylamide, 0.13% (w/v) piperazine diacrylamide, 2% ampholyte mixture pH 2–11, 9.0 M urea, 5% (w/v) glycerol, 0.06% (v/v) TEMED, 0.02% (w/v) APS. Shortly before application, mix 390 μL of *degassed* cap gel solution with 10 μL of 0.8% APS.
4. Anode solution: 7.27% (w/v) phosphoric acid, 3.0 M urea (use freshly prepared). Weigh out 72 g urea; fill in graduated cylinder with deionized water to 380 mL and solubilize by stirring. Before the application add 20 mL of 85% (m/v) phosphoric acid, then stir and degas.
5. Cathode solution: 5% (v/v) ethylenediamine, 9.0 M urea, 5% (w/v) glycerol (use freshly prepared). Weigh out 162 g urea and 15 g glycerol; fill up to 285 mL with deionized water and solubilize by heating (do not heat above 40°C) and stirring. Add 15 mL ethylenediamine immediately before the application, stir and degas.
6. Overlaying solution: 5% (w/v) glycerol, 5 M urea, 2% (w/v) Servalyte 2-4. Weigh out 6 g urea and 1 g glycerol; fill up to 19 mL with deionized water and filtrate through 0.45-μm cellulose acetate filter. Mix a 3.8 mL aliquot of this stock solution with 0.2 mL Servalyte 2-4. Divide into 200 μL aliquots and store at –20°C.
7. Sephadex gel suspension. Mix 100 μL Sephadex-gel stock suspension with 108 mg urea and 10 μL of ampholyte mixture pH 2–11 and thoroughly vortex for 10–15 min.
8. Equilibration solution: 125 mM Tris–phosphate, pH 6.8, 40% (m/v) glycerol, 65 mM DTT, 3% (m/v) SDS. Prepare equilibration buffer using 12.11 g Tris-base and 35 mL of

deionized water; adjust pH value to 6.8 using 1 M phosphoric acid and fill up to 100 mL with deionized water. Add 160 g glycerol and 12 g SDS (and 40 mg bromophenol blue, if you want to control the frontline during the run) to 50 mL of equilibration buffer. Fill up to 400 mL with deionized water and filtrate through paper filter. Freeze at –20°C in 40 mL aliquots. Before the application to each 40 mL of the solution, add 0.4 g DTT and solubilize.

2.2.2. SDS–PAGE Stock Solutions

It is recommended that 3 or 6 L of SDS–PAGE gel solution be prepared. Three liters are sufficient for 22 preparative gels (1.5 mm thick) or for 44 analytical gels (0.75 mm thick).

1. SDS–PAGE solution A, pH 8.8 at 20°C (Table 3).

Table 3
SDS-PAGE gel buffer: Solution A

Amount			
For 3,000 mL final gel solution (A + B)	For 6,000 mL final gel solution (A + B)	Substance	Final concentration in gel
110.41 g	220.82 g	Tris-base	0.375 M
45.43 g	110.86 g	Tris–HCl	
0.960 mL	1.92 mL	TEMED	0.03% (v/v)
Add 1,400 mL	Add 2,800 mL	Deionized water	

Add 1 L of deionized water to precisely weighed Tris-base and Tris–HCl. Add TEMED and fill up to 1,400 mL with deionized water.

2. SDS–PAGE Solution B (Table 4).

Table 4
SDS-PAGE gel buffer: Solution B

Amount			
For 3,000 mL final gel solution (A + B)	For 6,000 mL final gel solution (A + B)	Substance	Final concentration in gel
498 g	996 g	Acrylamide 4×	15% (m/v)
6.64 g	13.28 g	Bisacrylamide 2×	0.2% (m/v)
Add 1,660 mL	Add 3,320 mL	Deionized water	

Weigh out acrylamide and bisacrylamide in a 2-L beaker under a hood (wear a dusk mask!). Add about 1,200 mL deionized water, close the beaker with Parafilm. Stir on a magnetic stirrer within a water bath (40°C) to reach a temperature of 20°C.

Transfer to a 2-L graduated cylinder, wash the beaker with deionized water and use this water for filling up to 1,660 mL. Add 41.5 g Serva Serdolit MB1 (for 3 L gel solution) to a glass fiber filter on a 1,000-mL filtrate flask, add water (in total 2 L), stir, and apply vacuum until all of the water has moved through the ionic exchange material. Weigh out 41.5 g of the wet Serdolit and add it to the acrylamide/bisacrylamide solution within the 2-L beaker (2.5% (m/v) Serdolit in solution B). Stir with a blade agitator for 1 h protected from light. (This is the juncture at which solution A can be prepared.) After stirring Serdolit is removed by filtration in a filtrate flask through a glass fiber filter, the filtrate is transferred into a 2-L graduated cylinder.

3. SDS–PAGE gel solution (pH 8.8; 20°C); $T = 15.2\%$; $C = 1.3\%$ (**Table 5**)

Mix solutions A and B within a 5-L beaker. Filtrate through glass fiber filter into a 1-L filtrate flask. Degas filtrate for 30 min under vacuum and stir in a 3-L vacuum flask. During the process of degassing, weigh out SDS in a 100-mL beaker. Add 50 mL of the solutions A and B mixture, and solubilize SDS by careful pivoting. Carefully transfer these 50 mL into the mixture of solution A and B by a syringe with a 0.45-µm filter to remove particles. The mixture is stirred carefully to avoid foam development. Measure pH and temperature and add this information to the protocol. The pH should be 8.8 ± 0.05 at 20°C (*see* **Note 5**). Store gel solution in 135 mL aliquots at −20°C. Each aliquot is sufficient for one preparative gel or two analytical gels.

Final Solutions

1. APS: 1.28% (m/v). Weigh out 2.56 g APS; fill in graduated cylinder to 200 mL with deionized water and filtrate through 0.45-µm cellulose acetate filter. Divide into 9 mL aliquots for storage at −20°C.
2. Agarose: 1% (w/v) agarose (low melt preparative grade, BioRad); 125 mM, Tris–phosphate, pH 6.8; 0.1% (w/v) SDS. Add 1.25 mL of equilibration buffer (S. final solutions 8.) to 8.7 mL of 0.115%

Table 5
SDS-PAGE gel

Amount			
3,000 mL	**6,000 mL**	**Substance**	**Final concentration in gel**
1,400 mL	2,800 mL	Solution A Tris/TEMED	0.375 M/0.03% (m/v)
1,600 mL	3,200 mL	Solution B acrylamide/bisacrylamid	15%/0.2% (m/v)
3.2 g	6.4 g	Sodium dodecyl sulfate (SDS)	0.1% (m/v)
200 mL[a]	400 mL[a]	Ammonium peroxo disulfate[a] (APS) 2.56%	0.08%

[a] To be added directly before gel preparation (only mentioned here to be included in the calculation of the final concentrations of the substances within the gel)

(m/v) SDS solution. Add 0.1 g agarose and solubilize at 70°C. Filter and keep at 40°C until application, or store at 4°C.

3. SDS–PAGE gel (T = 15.2%, C = 1.3%): 15% (w/v) acrylamide; 0.2% (w/v) bisacrylamide; 375 mM Tris–HCl, pH 8.8 (20°C); 0.03% (v/v) TEMED; 0.1% (w/v) SDS; 0.08% (w/v) APS. Before application, add 9 mL of 1.28% APS to 135 mL of SDS–PAGE gel solution.
4. Electrode buffer for second dimension: 25 mM Tris; 192 mM glycine; 0.1% (w/v) SDS. Weigh out 30.0 g Tris-base, 144.0 g glycine and 10 g SDS; fill in graduated cylinder to 2,000 mL with deionized water and solubilize using a magnetic stirrer. Add 8 L deionized water. Use freshly prepared!
5. Underlay solution: 40% glycerol. 40 mL glycerol filled up to 100 mL with deionized water.
6. SDS-overlay solution: 375 mM Tris–HCl, pH 8.8 (20°C); 0.1% (w/v) SDS. Dissolve 17.251 g Tris-base, 7.099 g Tris–HCl and 0.5 g SDS in 500 mL and store 20 mL aliquots at –20°C.
7. Fixation solution (CBB G250 staining): 50% (v/v) methanol; 2% (v/v) phosphoric acid. Mix 500 mL methanol (extra pure) and 23.5 mL *ortho*-phosphoric acid (85%) and add deionized water to 1,000 mL.
8. Fixation solution (silver staining): 50% (v/v) ethanol; 10% (v/v) acetic acid. Mix 500 mL ethanol (absolute, denatured) and 100 mL acetic acid p.a. 100% and fill up to 1,000 mL with deionized water. This solution will suffice for 3×7 gel.

2.2.3. CBB G250 Staining

1. Incubation solution: 34% (v/v) methanol; 2% (v/v) phosphoric acid; 17% (w/v) ammonium sulfate. Measure 340 mL methanol (extra pure) and add 500 mL deionized water. Add 23.5 mL *ortho*-phosphoric acid (85%) and 170 g ammonium sulfate (extra pure) then fill up to 1,000 mL with deionized water.
2. Staining solution: 34% (v/v) methanol; 2% (v/v) phosphoric acid; 17% (w/v) ammonium sulfate; 0.066% (w/v) CBB G250. Add 0.660 g CBB G250 into 1,000 mL of incubation solution.
3. Washing solution: 25% (v/v) methanol. Measure 250 mL methanol (extra pure) and add deionized water to 1,000 mL.

2.2.4. Silver Stain

1. Incubation solution: 30% (v/v) ethanol; 0.5 M sodium acetate; 0.5% (v/v) glutaraldehyde; 0.2% (w/v) sodium thiosulfate. In 300 mL ethanol (absolute, denatured) add 41 g sodium acetate anhydrous p.a., 20 mL glutaraldehyde 25%, Grade II and 2 g sodium thiosulfate-pentahydrate p.a. then fill up to 1,000 mL with deionized water. Prepare this solution prior to use. Keep separately as hazardous waste.

2. Silver nitrate solution: 0.1% silver nitrate; 0.01% formaldehyde. Weigh out 1 g silver nitrate p.a., add 1,000 mL deionized water and 0.288 mL formaldehyde 37% (filtrated through filter paper). Prepare this solution prior to use.
3. Developer: 2.5% (w/v) sodium carbonate; 0.0012% (w/v) sodium thiosulfate; 0.01% (w/v) formaldehyde; 0.04% (w/v) sodium hydrogencarbonate. Weigh out 25 g sodium carbonate p.a. and 0.012 g sodium thiosulfate pentahydrate p.a. and fill up to 1000 mL with deionized water. Add 0.288 mL formaldehyde 37% (filtrated through filter paper) and approximately 0.4 g sodium hydrogencarbonate p.a. to adjust pH to 11.3. Prepare this solution directly prior to use.
4. Stop solution: 0.05 M EDTA; 0.02% (w/v) Thimerosal. Weigh out 18.612 g EDTA (Titriplex III) and fill up to 1,000 mL with deionized water. Add 0.200 g Thimerosal. Prepare this solution prior to use.

2.2.5. Tryptic Digestion

1. Destaining buffer: 200 mM ammoniumbicarbonate; 50% ACN (v/v). Weigh out 1.58 g ammoniumbicarbonate, add 50 mL ACN and fill up to 100 mL with deionized water.
2. Digestion buffer: 50 mM ammoniumbicarbonate; 5% ACN (v/v). Weigh out 0.39 g ammoniumbicarbonate, add 5 mL ACN and fill up to 100 mL with deionized water.
3. Trypsin solution: 0.1 μg/μL trypsin (Sequencing grade modified Trypsin, Promega, WI, USA); 50 mM acetic acid. Add 20 μg lyophilized trypsin into 200 μL resuspension buffer (Promega: 50 mM acetic acid).
4. ACN 60% (v/v); TFA 0.5% (v/v). Add 50 μL TFA into 6 mL ACN and fill up to 10 mL with deionized water.
5. ACN 100%.
6. ACN 33%/TFA 0.1%. Add 3.3 mL ACN into 10 μL TFA and fill up to 10 mL with deionized water.

2.2.6. MALDI-MS

1. CHCA matrix: 0.5% alpha-cyano-4-hydroxycinnamic acid (CHCA) solubilized in 60% ACN and 0.3% TFA. Weigh out 2.5 mg CHCA, add 250 μL deionized water, 250 μL ACN and 1.5 μL TCA.

3. Methods

3.1. IEF

3.1.1. Preparation of Gels

1. Pass PP threads from the lower end of the glass tubes to the upper end (sealed end of the thread at the lower end of the tube). Four glass tubes with cap gel end (two calibration rings 3 and 9 mm) below are inserted into the holder of each of the

two gel casting apparatuses. The glass tubes are adjusted to have contact with the "filling boat."

2. Thaw separation gel solution and APS solution in water bath at 37°C. Temperature of the gel solution has to be controlled, because the polymerization may be too fast at high temperatures (>37°C).
3. Degas separation gel solution for about 4 min; the glass tubes are gently dabbed at the edge of the table to precede air bubbles.
4. Thaw and degas cap gel solution and for the subsequent procedure, *see* **step 3**.
5. Add 25 μL of APS to 975 μL of separation gel solution using a Gilson pipette. The solution is carefully shaken without introduction of oxygen into the gel.
6. Fill separation gel solution into the "filling boat" (2,000 μL per four tubes). All the ends of the glass tubes have to be immersed by the gel solution.
7. Draw gel solution up to the 30 mm mark by carefully pulling out of PP threads.
8. Change compartment of "filling boat."
9. Add 10 μL 0.8% APS to 390 μL of cap gel solution using a Gilson pipette and mix the solution by gentle shaking.
10. Fill 200 μL cap gel solution into the free compartments of the "filling boat" of each of the two gel casting apparatus and pull out gel solution as far as to the 24 mm mark.
11. Remove "filling boat" and pull out gel solutions as far as 21 mm mark. As a result, the air is sucked into the tube between the lower end and the 3 mm mark.
12. After 30 min polymerization, pull out PP threads. Form a moist chamber on the upper side by adding of one big drop of deionized water on the aperture of capillary. An air bubble is formed between the gel and water drop. The moist chamber formed in this way and the cap gel side is subsequently closed by Parafilm. For further polymerization, the gels should be kept at room temperature. The prepared tubes can be used 2–4 days after their preparation.

3.1.2. Sample Application

1. Add ethylenediamine to the cathode solution and degas for 5 min. Fill the bottom chamber with cathode solution (*see* **Note 6**).
2. Thaw Sephadex solution, add urea and a mixture of ampholytes pH 2–11. After thawing, 108 mg urea and 10 μL of the ampholyte mixture is added to 100 μL of Sephadex gel. The mixture is thoroughly shaken for 10–15 min using a Vortex, for example.

3. Insert the tubes into the anode part of the focusing chamber. The 21 mm marking should be still visible. Cap gel side faces upwards.
4. Fill cathode solution into the cathodic side of the tubes without forming bubbles (remove water before adding the solution).
5. Mount the anode part of the focusing chamber on the bottom one. The tubes must dip into cathode solution.
6. Remove water and dry up loading side. Water is removed using a Hamilton syringe or a gel loader tip and the gel surface is subsequently dried by filter paper.
7. Thaw prepared sample and overlay solution.
8. Overlay the gel by approximately 2 mm of Sephadex solution using multiflex pipette tips or Hamilton syringe.
9. Load the sample (1–30 μL), and apply it onto the Sephadex surface. The gel loader tip has to have contact with the Sephadex layer at the moment you press out the sample from the gel loader tip. The sample is loaded carefully to prevent formation of air bubbles (*see* **Note 7**).
10. Load overlay solution and fill in with anode solution freshly prepared by addition of phosphoric acid (see solutions for 2-DE). About 5 μL of overlay solution is gently layered onto the sample using multiflex tips or Hamilton syringe (to prevent mixing overlay solution with the sample). The capillary is subsequently filled up with anode solution without forming air bubbles.
11. Use the remaining anode solution to fill up the upper chamber and cover all tubes with buffer.

3.1.3. Running IEF

Voltage is increased in a stepwise manner as given in Table 6 (*see* **Note 3**).

Table 6
IEF running conditions

Voltage (V)	Time
100V	60 min
200V	60 min
400V	17.5 h
600V	60 min
1,000V	30 min
1,500V	10 min
2,000V	5 min

3.2. SDS–PAGE

3.2.1. Incubation of IEF Gels

1. Thaw the equilibration solution about 1 h before the end of the IEF run. Add 0.4 g DTT to 40 mL of equilibration solution.
2. After the IEF ends, remove the cathode and anode solutions from the tubes.
3. Overlay the cap gel side with 87% glycerol solution.
4. Fill 10 mL of equilibration solution into the Petri dishes.
5. Extrude the gels using PP threads (*see* **Note 8**). PP threads are inserted into the tubes from the cap gel side. While the loading side of the tube is dipped into deionized water, the gel is carefully expelled a little down. At this time, the mixing of the overlay solution with the water can be observed. Pull the gel out about 5 mm and briefly wash the extruded end with water. Then the tube is held above the Petri plate filled with the equilibration solution and the whole gel is completely extruded by the PP thread into the plate. One gel is equilibrated within one plate.
6. Equilibrate gel for exactly 10 min by shaking at room temperature.
7. If the gels are not used immediately after IEF, the equilibration solutions are poured off and the gels are stored in Petri plates at –80°C. From our experience, storage at –20°C results in lower resolution or destruction of the gels, and storage in liquid nitrogen destroys the gels.

3.2.2. Preparation of SDS–PAGE Gels

1. Thaw gel solution and APS.
2. Clean glass plates and spacers. Glass plates and spacers are sprinkled with 70% ethanol and cleaned with Kimwipes.
3. The gel is prepared upside-down. To obtain an as even as possible surface of the cathodic side (IEF gel application), the cathodic surface is produced by a glycerol underlaying procedure. For this purpose, the rubber gasket is perforated and the glycerol introduced through a tube. A syringe is filled with 4 mL of 40% glycerol, connected with the tube and the tube is filled up to the end with 40% glycerol. The syringe should still contain at least ww1.2–2 mL for underlaying.
4. Assemble glass plates. Rubber sealings (on the outside) and spacers are placed on the right and the left margin of a glass plate. Now the second glass plate can be positioned precisely over the first one. Attach clamp assemblies and tighten the clamp screws.
5. Gel preparation: 2 × 135 mL (4 × 0.75 mm thick gels or 2 × 1.5 mm thick gels) of gel solution is mixed with 2 × 9 mL APS in a beaker without formation of foam. The solution must be homogenous and the uptake of oxygen should be kept down.

6. The separation gel is poured carefully avoiding air bubbles into the gel cell up to 1 cm below the top of the glass plate assembly.
7. Underlay now with 1.2 mL (for 0.75 mm thick gels) or 2 mL (for 1.5 mm thick gels) of 40% glycerol by pressing it with the syringe through the tube. The gel solution should reach the upper end of the glass plates.
8. Remove the eventually produced air bubbles with a platinum wire or by knocking carefully on the glass plates.
9. Close the upper gel surface with Parafilm and polymerize for 25 min.
10. Turn upside-down after polymerization and removal of Parafilm.
11. Replace 40% glycerol with SDS-overlay solution. First wash the surface thoroughly with overlay solution and then fill up with overlay solution. Close with Parafilm. After 1 h at room temperature and storage overnight at 4°C, the gels are ready.

3.2.3. Application of IEF Gels and Run of SDS–PAGE

1. Pre-chill the electrode buffer in the tank down to 15°C.
2. Liquefy agarose solution by heating it up to 70°C, then cool to 40°C (to keep it liquid).
3. Load the IEF gels on the surface of the slab gels. Remove the SDS-overlay solution from the slab gel using filter paper. Pull out the rod gel from the Petri dish and lay it lengthwise on the margin of one of the glass plates. Try to place the rod gels as near as possible to the edge of the filling groove. The rod gel is then pushed down gently from this end (with spatula) to slide carefully into the filling groove. The rod gel must lie on the surface of the SDS–PAGE gel and no air bubbles should be present.
4. Pouring agarose solution: 40°C heated-agarose is filled in the filling groove by a Pasteur pipette (without forming air bubbles) as far as the upper edge. Agarose is solidified after 2 min.
5. Insert glass plate assemblies into the electrode rack of the DESAGA chamber filled with 8.5 L electrode solution (in the bottom chamber, pre-chilled to 15°C): after the installation of the electrode assembly, about 1.5 L of electrode buffer is filled into the upper chamber in order to cover the gel surface with electrode buffer.
6. The SDS–PAGE separation is performed at 15°C. The gels should be cooled in the same way from both sides (*see* **Note 9**). The current is increased in a stepwise manner: 0.9 mm gels: 15 min 65 mA, about 6.5 h 85 mA; 1.5 mm gels: 15 min 120 mA, about 6.5 h 150 mA

After the SDS–PAGE separation, the electrode buffer in the upper chamber is poured off and the glass plate assemblies are removed from the electrode assembly. To remove gel, open the glass plates gently with a broad spatula so that one glass plate is pulled away. Now the gels can be transferred to fixation solution (*see* **Note 10**). Use talcum free gloves. Before the use, wash the gloves with water to remove potentially stuck talcum. After fixation, various staining techniques can be applied.

3.3. Cleaning of IEF Tubes

The tubes used for IEF must have an exact inner diameter or the casting and the removing of the gels using the polypropylene strings will not work. For these mechanical reasons, but also to avoid cathodic drift effects, the inner surface has to be free of any impurities. The following cleaning procedure is strongly recommended (*see* **Note 8**). The following materials are used: 2,000-mL glass beaker, stable wire from inert material, Deconex 12 PA (Borer Chemie AG, Zuchwil, Switzerland)), 0.1 M HCl, deionized water, tweezers, heated plate.

1. After the IEF ends, thoroughly wash the tubes with deionized water.
2. Heat 500 mL of a 6% Deconex solution i to 60 to –80°C.
3. Dip the glass tubes dipped into the Deconex solution and leave at 70°C for 3 × 10 min. After 10 min, take out the glass tubes individually (using tweezers), remove the cleaning solution and place the tubes again into the cleaning solution for 10 min at 70°C.
4. Thoroughly wash the tubes with deionized water.
5. Heat 500 mL of 0.1 M HCl solution to 95°C.
6. Overlay the tubes in the glass beaker with 0.1 M HCl until the vessel is full and tubes are covered. The tubes are left in HCl solution for 30 min at 95°C.
7. Thoroughly wash the tubes with deionized water and dry.

3.4. CBB G250 Staining (see Notes 11 and 12)

Staining is performed at room temperature under gentle shaking, 1 gel per dish and 1 L solution.

1. Fixation: gels are incubated in fixation solution overnight by shaking.
2. Wash: 3 × 30 min with tap water.
3. Incubation: 1 h in incubation solution.
4. Staining: after 1-h incubation, add 660 mg CBB G250 in 1 L solution and keep shaking up to 5 days without changing the staining solution.
5. Washing: after 5 days the gels are quickly (max. 1 min) washed with 25% methanol (to remove the CBB G250 colloids), then for 5 min with deionized water, and stored at 4°C sealed in a foil (*see* **Notes 11,13,14**).

3.5. Silver Staining (see Notes 14–23)

All steps should be performed under constant room temperature especially during the developing step. Complete all steps by gentle shaking on a horizontal shaker (*see* **Note 10**). In all steps, 1 L of the corresponding solution was used for each gel (*see* **Note 18**) unless specified differently.

1. Fixation: after electrophoresis, transfer gels into fixation solution and incubate overnight. Incubation: incubate gels in the incubation solution for 2 h (*see* **Note 20**).
2. Wash: wash gels 2× in 4 L of deionized water (20 min each step).
3. Incubation: incubate gels in silver nitrate for 30 min.
4. Rinse gels: rinse gels 2× in deionized water to remove silver ions from the surface.
5. Development: develop gels in developer under observation. Developing step is dependent on the amount of loaded proteins. Staining should begin after 2 min and should be finished within the next 5 min. More than 20 min development time is not recommended (*see* **Notes 13,14,23**).
6. Stop reaction: when the required color is reached, transfer to stop solution and incubate for 15 min.
7. Exchange stop solution. Store at 4°C in the dark. The gel may be stored for several months sealed in a foil.
8. Rinse in deionized water and dry the gel.

3.6. Tryptic In-Gel Digestion

Spots are cut manually with the help of a 1-mL pipette tip shortened by about 0.5 cm or with a spot cutter (*see* **Note 24–26**). Here we describe a manual digestion procedure.

1. Transfer the gel spot into a 0.5-mL Eppendorf tube (*see* **Note 27**).
2. Destaining. After destaining the gel pieces in 500 μL of destaining buffer for 30 min at 37°C, destaining buffer is removed.
3. Equilibration. Incubate for 30 min at 37°C in digestion buffer (without trypsin), then remove the digestion buffer.
4. Dry the gel pieces for 30 min at 37°C.
5. Digestion. Add 25 μL of digestion buffer containing 0.1 μg/μL trypsin and incubate overnight at 37°C.
6. Extraction of the peptides. After short centrifugation, transfer the supernatant in a new Eppendorf tube (S1). Add 25 μL 60% ACN and 0.5% TFA to the gel. Incubate at room temperature for 10 min and transfer supernatant S2 to S1. Add 25 μL 100% ACN to the gel. Incubate at room temperature for 10 min and transfer supernatant S3 to S1 + S2.
7. Dry the combined supernatants at 45°C for 60–90 min (*see* **Note 28**).

3.7. MALDI-MS

3.7.1. Sample application Without Desalting

The dried peptides are dissolved in 1 µL of 33% ACN and 0.1% TFA; mix 0.25 µL (*see* **Note 29**) of this solution with 0.5 µL of alpha-cyano-4-hydroxycinnamic acid (CHCA) solubilized in 60% ACN, 0.3% TFA on Parafilm by threefold up and down pipetting (*see* **Notes 30** and **31**). Transfer the resulting mixture to the MALDI-MS plate and analyze. Crystallization is accomplished at room temperature (*see* **Notes 32** and **33**).

3.7.2. Sample application with Zip Tip Desalting

Poor crystallization and associated low MALDI-MS sensitivity are often related to the presence of detergents or salts within the sample. These impurities can be removed by Zip Tips. The following procedure is adjusted to the Micro ZipTips.

1. Wetting: pipette 10 µL of 50% ACN, 0.1% TFA 3× up and down.
2. Equilibration: pipette 10 µL of 0.1% TFA 3× up and down.
3. Binding of peptides: peptides are dissolved in 25 µL of 0.1% TFA. Pipette peptide solution 20× up and down.
4. Wash: pipette 10 µL of 0.1% TFA 3× up and down.
5. Peptide elution: elute peptides by pipetting 5× up and down with 1–5 µL of 60% ACN, 0.1% TFA (*see* **Note 34**).
6. ZipTip cleaning: 100% ACN 3× up and down, 60% ACN, 0.1% TFA 3× up and down, 0.1% TFA 3× up and down.

3.8. PMF and MS/MS

1. The PMF MS spectra are obtained using the following parameters: reflectron mode, 20 kV accelerating voltage, a low mass gate of 800 Da and a mass range between 800 and 4,000 Da. A maximum of 1,000 shots/spectrum is accumulated.
2. Tandem mass spectra are obtained without collision gas by accumulation of a maximum of 10,000 shots per spectrum. The precursor mass window is adjusted to a relative resolution of 100 and the metastable suppressor is set on.
3. For each spot, a mass spectrum and three tandem mass spectra (for the three most intense MS peaks) are measured. An exclusion list, which specifies masses to be excluded from MS/MS measurement, should be added (e.g., trypsin auto digestion, matrix and keratin peptides). More specific exclusion list can be obtained using MS-Screener *(23)*, a software available at the 2D-PAGE proteome database portal (http://www.mpiib-berlin.mpg.de/2D-PAGE/). For internal calibration the trypsin auto digestion peptides with masses of 842.51 and 2,211.11 are used.
4. The peak detection parameters have to be optimized individually for each mass spectrometer. An example of the parameter setup used with the 4000 Series Explorer Software Version 3.6 is given below.

For MS: Minimum S/N Filter: 10, Local Noise Window Width (*m/z*): 50, Min Peak Width at Full Width Half Max (bins): 2.9, Cluster Area S/N Optimization: yes, S/N Threshold: 15; MS/MS: Subtract Baseline Peak Width: 50, Savitsky-Golay Points Across Peak: 3, Polynomial Order: 4, Minimum S/N Filter: 7, Local Noise Window Width (*m/z*): 50, Min Peak Width at Full Width Half Max (bins): 2.9, Cluster Area S/N Optimization: yes, S/N Threshold: 10. The resulting rough peak list may be exported as a text file using the software Peaks to Mascot. Alternatively, the GPS Explorer TM Software (Version 3.5) can be used to further process this fairly accurate peak list.

For the Peaks to Mascot software: MS peak filtering: Peak Density Filter: 10 peaks per 200 Da, Minimum S/N Filter 10, Min Area: 1,000, Max. No. of Peaks: 100; MS/MS peak filtering: Mass range (Da): 60 – (Precursor – 20 Da), Peak Density Filter: 50 peaks per 200 Da, Minimum S/N Filter: 5, Min Area: 0, Max. No. of Peaks: 100. The peak lists are in Mascot generic data format (.ppw files for combined MS and MS/MS files).

For GPS Explorer TM Software: MS peak filtering: Minimum S/N Filter: 10, Peak Density Filter: 10 peaks per 200 Da, Max. No. of Peaks: 100, Adduct List: 21.982 (=NaCl), Adduct Match Tol.: 0.03 Da; MS/MS peak filtering: Minimum S/N Filter: 5, Peak Density Filter: 50 peaks per 200 Da, Max. No. of Peaks: 100. The GPS Explorer TM Software performs the Mascot search using an in-house Mascot server without the option of exporting the peak lists.

3.9. Protein Identification by MASCOT Search

The database searches with data files containing combined MS and MS/MS peak lists are performed with Mascot (http://www.matrixscience.) using the MS/MS Ions Search form. An alternative is to install an in-house Mascot server, which allows the use of batch searching. An example of the standard search parameters is given below.

Enzyme: Trypsin/P, Variable modifications: *N*-acetyl (Protein), Oxidation (M), Propionamide (C), Pyro-glu (N-term Q), Mass values: Monoisotopic, Peptide charge: 1+, Protein mass: unrestricted, Peptide Mass Tolerance: ±30 ppm, Fragment Mass Tolerance: ±0.3 Da, Max Missed Cleavages: 1, Instrument type: MALDI-TOF-TOF.

For identification of a protein within a 2-DE spot, the following identification criteria are used:

(a) If a protein has at least 30% sequence coverage by the MS analysis, it is identified (class 1 identification). (b) If it has a sequence coverage between 15% and below 30%, at least one tandem mass spectrum with Mascot identity combined with mass loss information and hyper cleavage sites *(24)* is necessary to reach the identification status (class 2 identification). (c) If the sequence coverage is below 15%, two tandem mass spectra or a single MS/MS with additional information from 2-DE (e.g., from neighbor spots or p*I* or protein M_r) are required (class 3 identification). (d) For remaining candidates the analysis has to be repeated.

For proteins with M_r below 10 kDa or above 70 kDa these parameters are sometimes too strict and additional parameters should be considered to avoid incorrect results. Criteria only based on the Mascot Score are problematic, because the sizes of the sequence databases and the search criteria both have an influence on the score value. For certain applications, the use of inhouse versions of sequence databases may be useful. The quality of identification can be controlled by searches against reverse or shuffled databases, which should result in less than 5% identification (i.e. false positives).

4. Notes

1. Alkylation by, e.g., iodoacetamide may be helpful but is not strictly required. The proteins are denatured and reduced by the 2-DE sample preparation and alkylation occurs during electrophoresis by reaction with unpolymerized acrylamide leading to propionamide.
2. In 2006, it was possible to identify and document about 700 spots at the protein level within 3 months. The identification at the protein species level, which requires 100% sequence coverage, becomes possible by combining different digestion procedures and modifications of MALDI-MS measurements, but with much lower throughput at present *(25, 26)*.
3. If the spots are accumulated in a certain region of the gel, change the ampholyte composition or change the running time in IEF. Proteins are focused already very far away from their p*I* value and so, for example, very basic ribosomal proteins may be separated by reducing the running time to about one-fourth using the ampholyte mixture suggested here.
4. With the 2–11 ampholytes and the running conditions described here, an effective p*I* range between 4 and 10 is established.
5. The optimal pH value might be difficult to achieve. If proteins with a M_r below 20 kDa are not separated, but collected in the front line, the pH of the gel buffer is too low. While decreasing the pH sharpens the spots, increasing the pH improves the resolution of low M_r proteins. With the gels prepared here, 5-kDa proteins should be detectable.
6. Whereas for most of the proteins the anodic sample application in IEF has advantages, a cathodic sample application may be advantageous for very acidic proteins (because their solubility increases in the basic pH range).

7. Sample application onto the IEF gels is the most difficult step in the whole procedure. The needle has to have contact with the Sephadex at the moment the sample is applied, otherwise air bubbles, which are difficult to remove, may form. This step needs experience.
8. IEF gels break at their cathodic side: cathodic drift effect: remove salt out of the system, avoid salt in the sample preparation and clean the tubes more extensively.
9. The temperature within the electrode solution of SDS–PAGE is critical, because of the temperature dependency of the pH in the gel, which is buffered by Tris. Temperature should be controlled by a thermometer within the gel chamber and kept at 15°C throughout the run.
10. Thin 23 cm × 30 cm gels tend to break if treated without caution. Again here, some training is necessary.
11. The preparative gels should be protected from all sources of keratin to avoid keratin contamination, which will complicate MS analysis.
12. Most but not all silver stained proteins can be detected on CBB G250 stained gels if 10- to 50-fold amount is applied onto the gels. These high amounts of material need thick gels, which in turn reduce the resolution.
13. The most common reasons for “empty 2-DE gels” are wrong polarity of electrodes, air bubbles within the system and erroneous protein concentration.
14. 2-DE patterns exhibiting an increased number of low M_r proteins with many spot series, high background in the low M_r range (<40 kDa) and no high M_r proteins indicate protein degradation (by endogenous proteases or during sample preparation). Add protease inhibitors as early as possible, reduce temperature and time during sample preparation.
15. There are also examples of proteins that cannot be stained with silver but can be stained with CBB G250.
16. Depending on the complexity of the sample, the protein amount to be applied is 30–100 μg for gel size and silver staining discussed herein.
17. As the ratio between the gel and staining solution is important for the quality of the stained gels, all gels should float in the staining solutions. 23 cm × 30 cm gels need 1 L of solution. An increase of gels per volume leads to inferior staining results.
18. All solutions should be prepared in deionized water of high quality. If the quality is not high enough, as in mono-distilled water, precipitation may occur.

19. The dishes for the staining should be made of white plastic material, free of grooves. The white color is important for an optimal view of the development of the gels.
20. The most common reasons for insufficient sensitivity are quality or age of glutaraldehyde and formaldehyde.
21. The most common reasons for smear on gels are dirty dishes and incomplete or omitted wash step.
22. Point streaks are caused by dust particles on the surface of the glass plates of the second dimension. Clean thoroughly.
23. Protein spots containing 1–10 ng of protein are detectable with the silver staining procedure described herein.
24. Automated spot pickers and digesters are helpful for high-throughput proteomics. If only 2,000 spots per year are identified, spot picking and digestion are not the limiting steps of the procedure.
25. To avoid salts and contaminations, use water and chemical products with high-purity quality.
26. Protein spots that can be detected on CBB G250 stained gels, even if the spot staining is very faint, should result in identification.
27. Large number of samples can be analyzed in 96 well micro titer plates. Destaining time has to be doubled, because of the lower volume of the wells, and destaining is performed 2× with 200 μL, in each step. The proteins are digested overnight at 37°C in an incubator and the peptides are extracted as described for the procedure within Eppendorf tubes.
28. For complete removal of ammonium bicarbonate from the peptide sample, it is essential to dry the combined supernatants at 45°C.
29. In the case of low protein concentration, increase the volume of peptide mixtures from 0.25 to 0.5 μL. Alternatively, concentrate the remaining sample using the Zip Tip procedure (described in **Subheading 3.6**) by direct elution with 1 μL of CHCA solubilized in 60% ACN, 0.3% TFA onto the MALDI-MS plate.
30. CHCA is light sensitive and should be stored in the dark. Matrix solution can be stored in aliquots at −20°C. After thawing, do not freeze the matrix solution again.
31. Sodium adducts might be associated with impure matrix. Solubilize 5 g of CHCA (e.g., Aldrich 1302B1) with 50 mL of warm ethyl alcohol. Fill the supernatant in a new cup and add 450 mL of deionized water. Wait for precipitation and discard the supernatant. Wash twice with 500 mL of deionized water. Transfer the CHCA in 2-mL cups and lyophilize. Store at −20°C.

32. Possible reasons for no MS signal are inappropriate protein/matrix relation or salt content and low overall protein amount (i.e., below detection limit).
33. A high protein concentration may result in poor crystallization. Add 0.5 μL matrix solution to the sample on the MALDI-MS plate and measure again.
34. To increase sequence coverage, a step-wise elution from ZipTip is recommended (e.g., 30% ACN, 50% ACN and 80% ACN).

Acknowledgements

We thank Prof. Klose from The Institute of Human Genetics, Charité Berlin, who helped with developing the basis of the 2-DE technology presented here and coworkers from MPIIB for many useful discussions.

References

1. Macko, V. and Stegemann, H. (1969) Mapping of potato proteins by combined electrofocusing and electrophoresis identification of varieties. *Hoppe-Seyler's Z. Physiol. Chem.* 350, 917–919.
2. Kaltschmidt, E. and Wittmann, H.G. (1970) Ribosomal proteins. VII Two-dimensional polyacrylamide gel electrophoresis for fingerprinting of ribosomal proteins. *Anal. Biochem.* 36, 401–412.
3. Vesterberg, O. and Svensson, H. (1966) Isoelectric fractionation, analysis and characterization of ampholytes in natural pH gradients. IV. Further studies on the resolving power in connection with the separation of myoglobins. *Acta Chem. Scand.* 20, 820–834.
4. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
5. O'Farrell, P.H. (1975) High-resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007–4021.
6. Klose, J. and Kobalz, U. (1995) Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 16, 1034–1059.
7. Westermeier, R., Postel, W., Weser, J., and Goerg, A. (1983) High-resolution 2-DE with IEF in immobilized pH gradients. *J. Biochem. Biophys. Methods* 8, 321–330.
8. Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl Acad. Sci. U. S. A.* 76, 4350–4354.
9. Vandekerckhove, J., Bauw, G., Puype, M., Van Damme, J., and Van Montagu, M. (1985) Protein-blotting on Polybrene-coated glass-fiber sheets. A basis for acid hydrolysis and gas-phase sequencing of picomole quantities of protein previously separated on sodium dodecyl sulfate/polyacrylamide gel. *Eur. J. Biochem.* 152, 9–19.
10. Eckerskorn, C., Jungblut, P., Mewes, W., Klose, J., and Lottspeich, F. (1988) Identification of mouse brain proteins after two-dimensional electrophoresis and elec-

troblotting by microsequence analysis and amino acid composition analysis. *Electrophoresis* 9, 830–838.

11. Karas, M. and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* 60, 2299–2301.
12. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246(4926), 64–71.
13. Pappin, D.J., Hojrup, P., and Bleasby, A.J. (1993) Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* 3, 327–332.
14. Medzihradszky, K.F., Campbell, J.M., Baldwin, M.A., Falick, A.M., Juhasz, P., Vestal, M.L., and Burlingame, A.L. (2000) The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal. Chem.* 72, 552–558.
15. Klose, J. (1999) Large-gel 2-D electrophoresis, In *Methods in Molecular Biology 112: 2-D Proteome Analysis Protocols* (Link, A.J., (ed.), Humana Press, Totowa, NJ, pp 147–172.
16. Rabilloud, T. (1999) Solubilization of proteins in 2-D electrophoresis, In *Methods in Molecular Biology 112: 2-D Proteome Analysis Protocols* (Link, A.J., (ed.), Humana Press, Totowa, NJ, pp 9–19.
17. Jungblut, P.R., and Seifert, R. (1990) Analysis by high-resolution two-dimensional electrophoresis of differentiation-dependent alterations in cytosolic protein pattern of HL-60 leukemic cells. *J. Biochem. Biophys. Methods* 21, 47–58.
18. Doherty, N.S., Littman, B.H., Reilly, K., Swindell, A.C., Buss, J.M., and Anderson, N.L. (1998) Analysis of changes in acute-phase plasma proteins in an acute inflammatory response and in rheumatoid arthritis using two-dimensional gel electrophoresis. *Electrophoresis* 19, 355–363.
19. Scheler, C., Lamer, S., Pan, Z., Li, X.P., Salnikow, J., and Jungblut, P. (1998) Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser desorption/ionization-mass spectrometry (MALDI-MS). *Electrophoresis* 19, 918–927.
20. Miller, I., Crawford, J., and Gianazza, E. (2006) Protein stains for proteomic applications: which, when, why? *Proteomics* 6, 5385–5408.
21. Jungblut, P.R., Bumann, D., Haas, G., Zimny-Arndt, U., Holland, P., Lamer, S., Siejak, F., Aebischer, A., and Meyer, T.F. (2000) Comparative proteome analysis of *Helicobacter pylori. Mol. Microbiol.* 36, 710–725.
22. Schmidt, F., Donahoe, S., Hagens, K., Mattow, J., Schaible, U.E., Kaufmann, S.H., Aebersold, R., and Jungblut, P.R. (2004) Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. *Mol. Cell. Proteomics* 3, 24–42.
23. Schmidt, F., Schmid, M., Facius, A., Mattow, J., Pleissner, K.-P., and Jungblut, P.R. (2003) Iterative data analysis is the key for exhaustive analysis of peptide mass fingerprints from proteins separated by two-dimensional electrophoresis. *JASMS* 14, 943–956.
24. Schmidt, F., Krah, A., Schmid, M., Jungblut, P.R., and Thiede, B. (2006) Distinctive mass losses of tryptic peptides generated by matrix-assisted laser desorption/ionization time-of-flight/time-of-flight. *Rapid Commun. Mass Spectrom.* 20, 933–936.
25. Hoehenwarter, W., Ackermann, R., Zimny-Arndt, U., Kumar, N.M., and Jungblut, P.R. (2006) The necessity of functional proteomics: protein species and molecular function elucidation exemplified by *in vivo* alpha A crystallin N-terminal truncation. *Amino Acids* 31, 317–323.
26. Okkels, L.M., Müller, E.C., Schmid, M., Rosenkrands, I., Kaufmann, S.H., Andersen, P., Jungblut, P.R. (2004) CFP10 discriminates between nonacetylated and acetylated ESAT-6 of *Mycobacterium tuberculosis* by differential interaction. *Proteomics* 4, 2954–2960.

Chapter 5

The Use of Difference In-Gel Electrophoresis for Quantitation of Protein Expression

Rajat Sapra

Summary

Two-dimensional difference in-gel electrophoresis (2D-DIGE) is a modified 2D electrophoresis (2DE) technique that enables comparison of two (or three) proteomes simultaneously on the same gel. The different protein samples to be compared are covalently tagged with spectrally different fluorescent dyes that are designed to have no effect on the relative migration of proteins during isoelectric focusing or molecular mass separation during electrophoresis. The "spot maps" generated from the dye scans for each of the dyes are then superimposed to discern the expression pattern of the proteome samples being compared. Proteins that do not change in expression are seen as spots with a fixed ratio of fluorescent signals, whereas proteins that differ between the samples have different fluorescence ratios. The fluorescent dyes used in DIGE are cyanine flours and matched in respect of molecular weight. Two different dye chemistries are available enabling fluorescent tagging of as low as 5 μg of proteins to get the analysis of the regulation of the proteome. Furthermore, DIGE is a sensitive technique, capable of detecting as little as 0.5 fmol of protein, and this detection system is linear over a >10,000-fold concentration range.

Key words: Difference in-gel electrophoresis, Proteomics, Comparative proteomics, 2-D gel electrophoresis

1. Introduction

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), first described by O'Farrell in 1975 *(1)*, is the simplest, most popular and versatile method of protein separation among a rapidly growing array of comparative proteomic technologies. The technique is based on the orthogonal separation of proteins according to their isoelectric points through isoelectric focusing and molecular masses separation using SDS–PAGE. The major advantage of 2D-PAGE is that it allows for the separation of complex protein

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_5

samples with high resolution (e.g., several hundred proteins can be displayed simultaneously on a single gel). In addition, the large loading capacity of 2D 'protein spot maps' facilitates mass spectrometric characterization of various protein spots. However, in traditional 2D-PAGE different samples have to be analyzed on separate gels and resulting protein spot maps can be analyzed only after staining. As a result, gel-to-gel reproducibility and inter-gel comparison and matching becomes a challenge.

Difference in-gel electrophoresis (DIGE) was developed to overcome these challenges *(2)*. In DIGE, samples of interest are labeled with two (or three) spectrally distinct fluorescent dyes prior to running them on the same gel (**Fig. 1**). The fluorescent dyes used in DIGE are cyanine dyes that are available either as N-hydroxy succinimide (NHS) esters or have maleimide reactive group. The three NHS-ester dyes – Cy-2, Cy-3 and Cy-5 – are molecular weight matched, positively charged and react with primary amines in Lys residues. The maleimide reactive dyes, on the other hand, are available as Cy-2 and Cy-3 dyes, which have

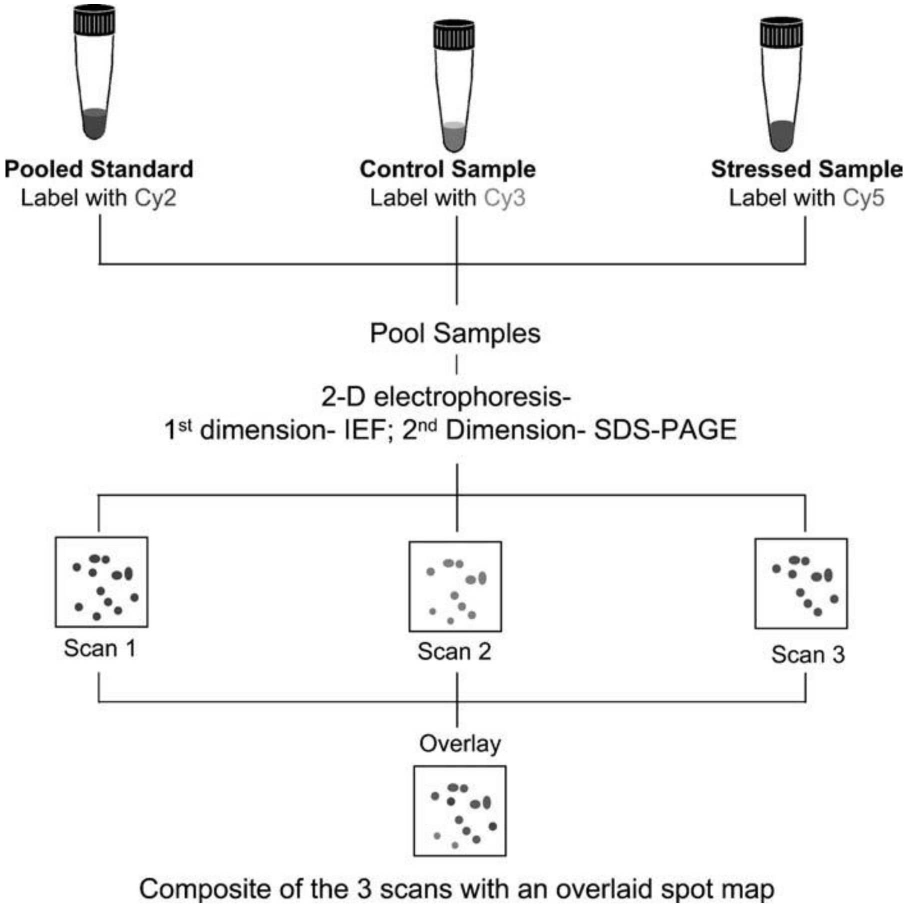

Pooled Standard
Label with Cy2
Control Sample
Label with Cy3
Stressed Sample
Label with Cy5
Pool Samples
2-D electrophoresis-
1st dimension- IEF; 2nd Dimension- SDS-PAGE
Scan 1
Scan 2
Scan 3
Overlay
Composite of the 3 scans with an overlaid spot map

Fig. 1. Overview of difference in-gel electrophoresis (DIGE). (*See Color Plates*)

similar molecular masses and react with Cys residues. These characteristics of the dyes result in no electrophoretic mobility shifts between the two proteins to be compared *(3)*.

In a simple experiment two cell extracts to be compared are separately labeled with Cy3-NHS and Cy5-NHS dyes, which covalently link to lysine residues in the protein. A third sample (pooled standard), which consists of equal amounts of the two samples being compared is labeled with Cy2-NHS *(4)*. The three sets of labeled proteins are then combined and co-electrophoresed using 2D-PAGE methodology. The gel is then imaged on a fluorescent gel imager at the Cy-2, Cy-3 and Cy-5 wavelengths. The three gel images corresponding to each of the Cy-Dye scans are then superimposed to generate a spot map where paired spots within the same gel can be accurately quantified. Since DIGE is an extremely sensitive method, in which a 15% change in protein abundance is more than two standard deviations (SDS) above the normal variation *(5)*, two most important considerations in performing DIGE are experimental design and sample preparation *(6)*.

2. Materials

2.1. Cell Lysis

1. Phosphate buffered saline (PBS). 137 mM sodium chloride (Sigma, St. Louis, MO), 2.7 mM potassium chloride (Sigma) and 10 mM phosphate buffer at pH 7.4 (*see* **Notes 1** and **2**).
2. Cell lysis solution I (for minimal labeling). Dissolve 7 M urea (BioRad, Hercules, CA), 2 M thiourea (GE Healthcare, Piscataway, NJ), 4% (w/v) CHAPS (Sigma) detergent in buffer containing 20 mM Na-HEPES (pH 8.5). The pH of the solution is very important for the labeling step. The solution can be prepared and stored in aliquots at –80°C for up to 3 months (*see* **Note 3**).
3. Cell lysis solution I (for saturation labeling). Dissolve 7 M urea, 2 M thiourea, 4% (w/v) CHAPS detergent in buffer containing 20 mM Na-HEPES (pH 8.0). The pH of the solution is very important for the labeling step. The solution can be prepared and stored in aliquots at –80°C for up to 3 months.
4. Cell lysis solution II. Any sample specific cell lysis solution containing protease inhibitor can be used. However, the proteins need to be precipitated and resuspended (in resuspension buffer) for the procedure described herein.

2.2. Protein Quantitation

1. 2D Quant kit (GE Healthcare) should be used to measure the concentration of the proteins if using cell lysis solution I, since it is compatible with all denaturants and detergents used in the procedure.

2. Bradford assay kit or Lowry assay kit (BioRad) should be used with any lysis solution other than cell lysis solution I.

2.3. Sample Cleanup

1. Acetone (ACN; MS Grade; Sigma). ACN is kept ice-cold at −20°C and used when needed for sample precipitation.

2.4. Sample Resuspension

1. Sample resuspension buffer for minimal labeling. Dissolve 7 M urea, 2 M thiourea, and 4% (w/v) CHAPS detergent in buffer containing 20 mM Na-HEPES (pH 8.5). The solution can be prepared and stored at −80°C for extended periods. Sample resuspension buffer for minimal labeling has identical composition as cell lysis solution for minimal labeling. It is important to check the pH of the solution before use as the buffer word pH is critical for the proper labeling of the Lys residues.
2. Sample resuspension buffer for saturation labeling. Dissolve 7 M urea, 2 M thiourea and 4% (w/v) CHAPS detergent in buffer containing 20 mM Na-HEPES (pH 8.0). The solution can be prepared and stored at −80°C for extended periods. The composition of the sample resuspension buffer for saturation labeling is identical to that of the cell lysis solution for saturation labeling. Again, it is important to check the pH of the solution before use as the pH is critical for the proper labeling of the Cys residues.
3. Dithiothreitol (DTT; Sigma) is purchased as a powder and stored at −20°C. Prepare fresh solution when needed.

2.5. Sample Labeling

2.5.1. DIGE Minimal Dyes

1. DIGE minimal dyes (GE Healthcare): cyanine NHS esters: Cy-2, Cy-3 and Cy-5 (GE Healthcare). Samples should be stored in the dark and at −20°C at all times. We recommend preparing and storing smaller aliquots (*see* **Notes 4** and **5**).
2. Dimethyl formamide (99.8% anhydrous DMF; Thermo Fisher Scientific/Pierce, Rockford, IL). DMF should be stored in a cool dry place. Only use DMF that is <3 months old.
3. Reducing agent: TCEP (“Bondbreaker” –500 mM TCEP neutral buffered solution from Thermo Fisher Scientific/Pierce) is the preferred reducing agent for the saturation labeling.
4. Quenching solution. 10 mM Lysine solution prepared in 10 mM Na-HEPES (pH 8.5). This solution can be stored in small aliquots for up to 3 months.

2.5.2. DIGE Saturation Dyes

1. DIGE saturation dyes (GE Healthcare): Cy-3 and Cy-5 (GE Healthcare). Samples should be stored in the dark and at −20°C at all times. We recommend preparing and storing smaller aliquots.
2. Dimethyl formamide (99.8% anhydrous DMF; Pierce).
3. Quenching solution: sample resuspension buffer for saturation labeling is used as the quenching solution.

2.6. First Dimension: IEF

1. IPG strips: IPG strips (GE Healthcare) can be purchased in five different sizes (7, 11, 13, 18 and 24 cm) and in different pH ranges depending on the survey required of the samples. The larger strips have a higher protein load and are used for preparative DIGE work (*see* **Note 6**).
2. IPG buffer: use IPG buffer (GE Healthcare) to match the pH range of the IPG strips. IPG buffer is used at the concentration of 0.5% (v/v; 12.5 μL IPG buffer in 2.5 mL of the sample).
3. Tris buffer.
4. Reducing agent for proteins during IEF: DeStreak Reagent (GE Healthcare) is used at the concentration of 1.25% (v/v) (12.5 μL of reagent in 1.0 mL of solution). Alternatively, DTT can be used at the concentration of 65 mM in the final solution.
5. 1% (w/v) bromophenol blue (tracking dye) in 25 mM Tris–HCl (pH 8.0).

2.7. Second Dimension: SDS–PAGE

2.7.1. IPG Strip Equilibration

1. Equilibration buffer stoc solution. 50 mM Tris–HCl (pH 8.8), 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS and 0.002% (w/v) bromophenol blue. Aliquot and store at –20°C for up to 6 weeks.
2. Reducing buffer. 1% (w/v) solution of DTT in equilibration buffer stock solution: add 100 mg of DTT (Sigma) to 10 mL of equilibration buffer stock solution to prepare the samples in IPG Strip for reduction. Always prepare fresh solution.
3. Acetylation buffer. 2.5% (w/v) iodoacetamide (IAA; Sigma) in equilibration buffer stock solution: add 250 mg of IAA to 10 mL of equilibration buffer stock solution to alkylate the proteins at Cys residues. Always prepare fresh solution.

2.7.2. SDS–PAGE

1. Bind-Silane (GE Healthcare) solution. Mix 8 mL ethanol (8 mL), 200 μL glacial acetic acid, 10 μL Bind-Silane and 1.8 mL of dH_2O. Always prepare fresh and from fresh Bind-Silane.
2. Separating buffer (4×): 1.5 M Tris–HCl, pH 8.8; 0.4% (w/v) SDS. Store at 4°C for up to 6 months.
3. 30% acrylamide/bisacrylamide (37.5:1 solution with 2.6% C; BioRad) and TEMED (BioRad) (*see* **Note 7**).
4. Ammonium persulfate (Sigma) solution (10%, w/v) in dH_2O. The solution should be frozen in small aliquots at –20°C immediately after resuspension.
5. Water saturated butanol. Shake equal volumes of water and *n*-butanol (Sigma) in a glass bottle and allow the two phases to separate. The top layer is water saturated *n*-butanol (*see* **Note 8**).
6. 10× SDS running buffer: 250 mM Tris, 1.92 M glycine (Sigma), 1.0% (w/v) SDS, pH 8.3.

7. Sodium azide (>99.5%, Sigma) is added at a concentration of 0.2 g/L of the 1× SDS electrophoresis buffer for the lower tank of the electrophoresis unit to prevent bacterial growth over longer term use and storage.
8. Molecular weight marker: kaleidoscope prestained MW marker (Invitrogen, Carlsbad, CA)
9. Agarose sealing solution: 0.5% (w/v) low melt agarose (Invitrogen) is dissolved in 1× SDS running buffer. The solution can be stored at RT for up to a year.
10. Fixing solution for the protein gels: 10% methanol (Sigma) and 5% acetic acid (Sigma) in dH_2O.
11. Low fluorescence glass plates (GE Healthcare) for SDS–PAGE.
12. Protein visualization stains: Sypro Ruby (Invitrogen), Deep Purple (GE Healthcare) or Krypton (Thermo Fisher Scientific/Pierce).

2.8. Protein Digestion

All reagents should be MS-grade and all solutions should be prepared with MS-grade reagents.

1. Buffer 1: 25 mM NH_4HCO_3 (100 mg/50 mL) in 5% acetonitrile (ACN).
2. Buffer 2: 25 mM NH_4HCO_3 in 50% can.
3. Trypsin buffer: 25 mM NH_4HCO_3 (100 mg/50 mL) in water.
4. Elution buffer: 50% ACN/5% formic acid (TFA or acetic acid may be substituted).

3. Methods

3.1. Sample Lysis

Make sure to use the lysis solution that corresponds to the Cy dye being used for (minimal or saturation) labeling. While the lysis procedure described below is applicable to any bacterial cell, it should be optimized for the type and the volume of the sample being analyzed.

1. Centrifuge the cell culture to pellet the cells. Wash the cells 2× with ice-cold PBS. Remove the remaining buffer from the tube using a pipettor.
2. Add ice-cold lysis buffer (use the proper buffer for labeling with minimal or saturation dyes) equal to ~4× the volume of the pellet to resuspend.
3. Using a 1-mL syringe with a 25G-5/8 needle, lyse the cells by repeatedly drawing and dispensing the liquid back into

the tube. Repeat the procedure 5× to shear the DNA. Do not dispense the liquid forcefully to minimize foaming. Alternatively, the cell suspension can be sonicated on ice 5× for 5 s, using a low power setting to minimize foaming and heat generation.

4. Centrifuge the resulting suspension for 5 min at 8,000 × *g* to pellet any remaining cell debris. Transfer the supernatant to another tube.
5. Determine the protein concentration of the solution using the 2D-Quant kit. Adjust the protein concentration to 5 mg/mL by either diluting the protein sample if it is more concentrated or using the precipitation procedure outlined in the next step if it is less concentrated.

3.2. Acetone Precipitation

This procedure is recommended for concentrating the protein sample and removing the ionic contaminants from the protein samples.

1. Determine the concentration of the protein solution(s) to be precipitated.
2. Cool the acetone to −20°C before adding to the protein solution.
3. Place protein sample in acetone compatible tube and add protein sample to the tube in a volume that is <1/5th of the maximum volume of the tube.
4. Add 4× the sample volume of cold (−20°C) acetone to the tube, vortex and incubate for at least 2 h at −20°C.
5. Centrifuge 10 min at 14,000 × *g*. Decant the supernatant without dislodging the pellet and allow the acetone to evaporate from the uncapped tube at room temperature for 30 min. Do not over-dry pellet using speed-vac or it may not dissolve properly.
6. Add rehydration buffer to the pellet to resuspend the pellet at a final concentration of 5 mg/mL.

3.3. Dye Reconstitution for Dye Stock Solution

3.3.1. DIGE Minimal Dyes

These instructions are for resuspending 10 nmol of the DIGE minimal dyes.

1. Label five small tubes for each of the Cy-2/3/5 dyes for storing aliquots.
2. The CyDyes should be stored at −20°C. Remove the dye containers and keep them on the benchtop for 5–10 min to bring the bottles to room temperature and prevent water condensation inside the dye tubes. Remove the tubes containing the dyes (the dyes are in powder form and may not be visible) and spin the tubes to pellet any dye powder that may be stuck on the walls of the tube (*see* **Note 4**).

3. Add 10 µL of the DMF to each dye tube (use a 10-µL syringe to accurately measure the DMF (*see* **Notes 9** and **10**) to get a stock concentration of 1 nmol/µL (1 mM). Close cap, vortex and spin again to pellet the dyes from the walls and place on ice in a covered ice bucket. Store 2 µL stock dye solution aliquots at −80°C for up to 3 months.

3.3.2. Preparation of Working Dye Solution for DIGE Minimal Dyes

1. Remove one 2 µL aliquot of each of the three dyes and add 3 µL DMF to each dye tube (kept on ice) to get a final volume of 5 µL and a final concentration of 400 pmol/µL (400 µM).
2. Any leftover dye should immediately be stored at −20°C and used within a week.

3.3.3. DIGE Saturation Dyes

Dye Resuspension for Analytical Samples

These instructions are for resuspending 100 nmol of DIGE saturation dyes.

1. Label two small tubes for each of the Cy-3 and Cy-5 dyes for storing aliquots.
2. The CyDyes should be stored at −20°C. Remove the dye containers and keep them on the benchtop for 5–10 min to bring the tubes containing the sye to room temperature before opening to prevent water condensation inside the dye tubes. Remove the tubes containing the dyes (the dyes are in powder form and may not be visible) and spin the tubes to pellet any dye powder that may be stuck on the walls of the tube.
3. Add 50 µL of the DMF to each dye tube (use a 50-µL syringe to accurately measure the DMF (*see* **Notes 9** and **10**) to get a working concentration of 2 mM for labeling analytical amount of protein (5 µg) Close cap, vortex and spin again to pellet the dyes from the walls and place on ice in a covered ice bucket. Store at −80°C for up to 3 months.

Dye Resuspension for Preparative Samples

The dye for preparative labeling is available as Cy-3 dye only. However, other comparative maleimide dyes can be used if needed.

1. Use the tube containing 400 nmol of the Cy-3 preparative dye to prepare the working dye solution for labeling.
2. The Cy-3 dye, like other Cy dyes for protein labeling, should be stored at −20°C. Remove the dye containers and keep them on the bench top for 5–10 min to bring the bottles to room temperature and prevent water condensation inside the dye tubes. Remove the tubes containing the dyes (the dyes are in powder form and may not be visible) and spin the tubes to pellet any dye powder that may be stuck on the walls of the tube.
3. Using a 50-µL syringe, add 20 µL of the DMF (*see* **Note 9**) to the tube containing 400 nmol of the Cy-3 dye to get a working concentration of 20 mM for labeling 500 µg of protein.

4. Close cap, vortex and spin again to pellet the dyes from the walls and place on ice in a covered ice bucket.

3.4. Protein Labeling

3.4.1. DIGE Minimal Labeling

The preferred protein concentration for labeling procedure is 5 mg/mL. For lower protein concentration, use acetone precipitation to concentrate the sample. The following example describes a comparison between control and perturbed proteome with a pooled internal control consisting of equal amounts of the two samples being compared.

1. To label Lys residues with NHS-ester dye, the pH of the protein solution should be between 8.0 and 9.0. Adjust the pH if necessary.
2. Label three microcentrifuge tubes as Cy-2 (Pool), Cy-3 (Control) and Cy-5 (Perturbed). Add 25 µg of each of the control and the perturbed sample to the tube labeled Cy-2. Add 50 µg of the control sample to the tube labeled Cy-3 and 50 µg of the perturbed proteome sample to the tube labeled Cy-5.
3. Add 1 µL each of Cy-2, Cy-3 and Cy-5 working dye solution to the appropriate tube. Vortex to mix the sample and pulse centrifuge using a tabletop centrifuge to spin the liquid from the walls of the tube. Incubate covered (to protect from light) on ice for 30 min.
4. Add 1 µL of 10 mM lysine solution and leave on ice for 30 min to quench the labeling reaction.
5. Labeled protein samples can be used immediately for IEF or frozen at −80°C for up to 3 months.

3.4.2. DIGE Saturation Labeling

Unlike minimal dyes saturation dyes are available as two different flours: Cy-3 and Cy-5. The experimental setup for saturation labeling differs from the setup used for the minimal dyes.

Sample Labeling for Analytical Gels

To accurately quantify the changes in the proteome, two gels should be analyzed: the gel with the labeled control sample and the pooled sample, and the gel with the labeled perturbed proteome sample and the pooled sample. The pooled sample is labeled with Cy-3, while both the control and the perturbed sample are labeled separately with Cy-5.

1. The protein solution to be labeled should be at pH 8.0. Adjust the pH if necessary.
2. Label two sets of two microcentrifuge tubes: Set 1: Cy-3 (Pool), Cy-5 (Control) and Set 2: Cy-3 (Pool) and Cy-5 (perturbed). Add 2.5 µg of each of the control and the perturbed sample to the tubes labeled Cy-3 (Pool), vortex to mix and pulse centrifuge to collect the sample at the bottom of the tube. Add 5 µg of the control sample to the tube labeled Cy-5

(Control) and 5 μg of the perturbed proteome sample to the tube labeled Cy-5 (Perturbed).

3. Prepare 2 mM solution from the concentrated 500 mM TCEP solution with sample resuspension buffer for saturation labeling.
4. Add 1.5 μL of 2 mM TCEP solution to each of the tubes containing 5 μg of the protein to be labeled, vortex to mix and pulse centrifuge to collect the sample at the bottom of the tube. Incubate in the dark at room temperature (RT) for 90 min.
5. Spin down the samples in the microcentrifuge. Add 3.0 μL of the Cy-3 dye to the tubes containing the pooled sample; add 3.0 μL of the Cy-5 dye to the tubes containing the control and the perturbed samples. Vortex to mix and centrifuge to collect the sample at the bottom of the tube. The dye:protein ratio has to be optimized for specific samples being analyzed. This example uses dye:protein ratio optimized for labeling bacterial samples.
6. Incubate in the dark at RT for 60 min.
7. Add 2× the volume of sample resuspension buffer for saturation labeling containing 130 mM DTT to stop the reaction. Vortex to mix the sample and the dye, centrifuge to collect the sample at the bottom of the tube.
8. Labeled protein samples can be used immediately for IEF or frozen at −80°C for up to 3 months.

Sample Labeling for Preparative Gels

The sample labeling procedure is slightly different for DIGE saturation dyes as compared to the DIGE minimal dyes. A larger amount of protein (≥500 μg) is labeled and only the labeled protein is loaded onto the gel. Since there is only one DIGE saturation dye commercially available, the pooled sample from all the comparative analyses is labeled. For preparative labeling, TCEP is added in a constant volume of 10 μL and the dye is added at a constant concentration of 20 mM for every 500 μg of the protein to be labeled.

1. The protein solution to be labeled should be at pH 8.0. Adjust the pH if necessary.
2. Prepare a 2 mM solution from the concentrated 500 mM TCEP solution with sample resuspension buffer for saturation labeling.
3. Add 10 μL of the 2 mM TCEP solution to each of the tubes containing 500 μg of the protein, vortex to mix and pulse centrifuge to collect the sample at the bottom of the tube. Incubate in the dark at RT for 90 min.
4. Spin down the samples in the microcentrifuge and add 20 μL (20 mM) of the Cy-3 dye to the tubes containing the

pooled sample. Use the same dye: protein ratio as for the analytical samples.

5. Incubate in the dark at RT for 60 min.
6. Add 2× the volume of sample resuspension buffer for saturation labeling containing 130 mM DTT to stop the reaction. Vortex to mix and centrifuge to collect the sample at the bottom of the tube.
7. Labeled protein samples can be used immediately for IEF or frozen at −80°C for up to 3 months.

3.5. Preparation for IEF

IPG strips are supplied as dehydrated polyacrylamide gels containing immobiline pH groups (IPG) on a plastic backing and must be rehydrated before use. The strips can be rehydrated with or without the protein in the rehydration solution. Here we describe the more widely applicable method of sample loading – rehydration with the protein in the rehydration solution. (If rehydrating the strips without the proteins, then the protein can be applied to the rehydrated strip using either cup loading or paper-bridge loading.)

IPG strips allow effective and reproducible IEF over a wide pH range. The strips are available in four different sizes: 7, 11, 13, 18 and 24 cm and in a wide variety of pH ranges. For instance, to overview a total protein distribution, use a broad range pH 3–11 NL strips (NL refers to nonlinear). IPG buffers are ampholyte containing buffer concentrates specifically formulated for use with IPG strips.

3.5.1. Rehydration of Immobiline Drystrips

This example describes IEF separation on an 18 cm pH3–10NL IPG Strip (GE Healthcare) for an analytical gel. Analytical gels have only the labeled proteins loaded for separation while preparative gels have unlabeled protein included in addition to the labeled samples (to facilitate protein spot 'picking' and protein identification).

1. Select the Strip Holder(s) corresponding to the IPG Strip length chosen for the experiment and place the ceramic strips on the plate with the strip electrodes making contact with the upper (+) and lower (−) electrode plate of the IEF unit (IPG-Phor). Level the unit and make sure that the surface is flat.
2. Prepare the sample solution for IEF:
    (a) Samples labeled with DIGE minimal dyes: Prepare the sample for IEF by adding the three labeled samples (12.8 μL × 3 = 38.4 μL), 1.75 μL of IPG buffer, 4.25 μL of DeStreak reducing agent, 0.30 μL of 1% bromophenol blue solution and rehydration buffer to make the final volume of up to 340 μL (*see* **Notes 11** and **12**).
    (b) Samples labeled with DIGE saturation dyes: Prepare two clean microcentrifuge tubes labeled Control-Pool and

Perturbed-Pool for IEF separation of the two sets of labeled samples. In the tube labeled Control-Pool, combine one tube of Pooled and the tube of Control labeled sample. In the tube labeled Peturbed-Pool, combine the other tube of labeled Pool sample and one tube of labeled Perturbed sample. Sequentially add the following: 280.95 μL of rehydration buffer, 1.75 μL of the IPG buffer and 0.30 μL of 1% bromophenol blue solution in each tube (*see* **Note 13**).

3. Apply the protein containing rehydration solution for IEF containing solution by pipetting the protein solution into each Strip Holder. The sample should be deposited slowly (to avoid generation of bubbles) between the two electrodes.
4. Remove the protective cover foil from the IPG strip starting at the acidic (+) end and with the gel side facing down and the acidic end of the strip facing the pointed end of the strip holder, glide the strip into the protein solution. Gently lift and lower the strip and slide the gel side of the strip against the ceramic with the lubricating effect of the sample solution to coat the entire strip with the protein solution. Be careful not to trap bubbles under the IPG strips.
5. Apply IPG strip cover fluid to minimize evaporation and urea crystallization. Pipette the fluid dropwise into one end of the strip holder until one half of the IPG strip is covered; pipette the fluid dropwise into the other end of the strip holder, adding fluid until the entire gel is covered (~2 mL needed for the 18-cm strip). Make sure the cover fluid does not overflow since it will impede the current between the electrode contacts between the strip holder and the instrument.
6. Place the cover on the strip holder and close the IPGPhor cover to ensure that the IPG strip maintains good contact with the electrodes as well as the gel swells.
7. Cover the IPGPhor with a dark cloth or an opaque fitted cover to limit photobleaching (*see* **Note 14**).
8. Allow the IPG strip to rehydrate at room temperature (25°C) on the IPGphor platform for a minimum of 12 h (maximum 16 h). Optional: The IPGphor can be programmed to apply low voltage (30–50 V) when used during the rehydration process. This process, called active rehydration, improves the uptake of large proteins into the IPG strip (*see* **Note 15**).

3.6. First Dimension: IEF

1. The following program is recommended for the 18 cm pH3–10NL IPG strip (being featured in our protocol).
    - (a) Active rehydration: 50 V, 700 Vh (1400 hours) (*see* **Note 15**)
    - (b) IEF:

(i) Step and hold – 500 V, 500 Vh (0100 hours)
(ii) Gradient – 1,000 V, 0.8 Vh (0100 hours)
(iii) Gradient – 8,000 V, 21 kVh (0300 hours)
(iv) Step and hold – 8,000 V, 32 kVh (0400 hours) (*see* **Notes 16–19**)

2. After the first dimension IEF run is done (~24 h), promptly take the strips out and either process these for the second dimension SDS–PAGE or store the strips at −80°C for later use (*see* **Note 20**). Running the second dimension immediately after the first dimension if the preferred method if sample through put is not an issue.

3.7. Second Dimension: SDS–PAGE

3.7.1. Silanization of Glass Plates for Preparative Gels

Spot picking of protein spots for mass spectrometric identification requires that the polyacrylamide gels are immobilized on the glass plate during casting so that the gel stays firmly attached to the glass during the staining and drying procedures. The most common method to accomplish this is silane treatment of glass plates with 3-methacryloxypropyltrimethoxysilane (trade name Bind-Silane). Treating the glass plates with Bind-silane requires a few hours of time but minimal labor. Only one plate is treated with Bind-silane so that the gel is immobilized to one plate and the sandwich can be opened for spot picking from the gel.

1. Thoroughly wash the plate to be treated with Decon detergent, rinse with dH_2O and dry the plate using a lint-free tissue.
2. Pipette 2–4 mL (depending on plate size) of the Bind-Silane (GE Healthcare) solution onto the plate and distribute equally over the plate with a lint-free tissue.
3. Cover the plate to prevent dust contamination and leave to air dry on the bench for >2 h.

3.7.2. Pouring Gels

1. Use low fluorescence gel plates for Ettan Dalt Six (GE Healthcare). If using the gels as analytical gels, the plates should be cleaned with detergent like Alconox (Alconox, NY) and rinsed with dH_2O. Clean all plates with 70% (v/v) ethanol and rub dry using lint free wipes (Kimwipes). Clean all components for pouring gels and wash with dH_2O. For preparative gels, use low fluorescence plates treated with Bind-silane as mentioned in **Subheading 3.7.1**.
2. Filter polyacrylamide solution (without APS and TEMED) through a 1-L disposable filter unit and degas for 1 h.
3. Assemble the gel rig and place the rig flat on table. Assemble the "gel sandwich" consisting of the two plates and a 1.5-mm spacer. Insert a plastic sheet in the back of the rig and place the gel sandwich on top of the plate. Use this arrangement of alternating gel paper and gel sandwich to assemble the

gels for pouring and continue in this fashion until the rig is full. Before placing the last plastic sheet in the rig, mark the sheet with a straight line to the level at which the polyacrylamide solution should be poured. Place the assembled rig in a secondary container in case of leaks.

4. Two hundred milliliters of the final solution is needed for pouring a large format gel. For a 10% SDS–PAGE gel, add 66 mL of 30% acrylamide solution, 50 mL of 4× buffer, 81 mL of dH_2O, 2 mL of 10% SDS solution, 1.0 mL of APS solution and 0.1 mL of TEMED (in the specified order). Mix the solution for about 30 s to 1 min and then begin transferring the solution into the back of the rig with a 50-mL pipette. When the desired solution level has been reached, overlay each gel with water saturated *n*-butanol. Place plastic wrap over the top of the rig to prevent evaporation of the *n*-butanol and allow plates to set for 3 h. After the gels have polymerized, wash the top of the gel with dH_2O. The prepared gels can be stored at 4°C covered with SDS running buffer soaked water towels and wrapped with saran wrap.

3.7.3. Equilibrating IEF Strips

Equilibration is carried out in a two-step process using tubes and volume of equilibration solution.

1. The second dimension SDS–PAGE gel should be prepared to time with the equilibration steps to accept the IPG Strip after the second equilibration (total time for the two equilibration steps is ~40 min).
2. Place the IPG strip in an individual tube (50-mL pipette tubes can be used for equilibrating larger size IPG strips) with the plastic side of the gel towards the tube wall.
3. Prepare 10 mL reducing buffer solution containing 100 mg of DTT and pour it into the tube containing the IPG strip. Plug and close the tube and gently rock, covered from light, on a shaker for 15 min. Carefully decant the solution taking precautions not to damage the IPG Strip.
4. Add 10 mL of freshly prepared alkylating buffer containing 250 mg of iodoacetamide to the tube containing the IPG strip. Plug and close the tube and gently rock, covered from light, on a shaker for 15 min. Carefully decant the solution and rinse with dH_2O. *Omit this step for IPG strips containing proteins labeled with DIGE saturation dyes.*
5. Equilibrated IEF strips should be run on the second dimension immediately after the second equilibration step.

3.7.4. SDS–PAGE

1. Melt agarose-sealing solution in a microwave and place in a beaker of hot water to keep it molten.

2. Remove the IEF strip from the reduction and alkylation step using forceps and place the IPG strip on top of the second dimension gel with the plastic touching the back glass. Gently push the IPG strip down using the blunt end of a spatula until the strip makes contact with the top of the gel. Make sure not to damage the IPG strip gel by either stabbing the gel or from contact with the glass plate. The standard orientation of the strip is with the acidic end of the strip towards the left.
3. Gently cover the strip with melted agarose starting from one end of the IPG strip until it evenly covers the entire strip. Take care to avoid any bubbles between the contact surface of the bottom of the IPG Strip and the top of the polyacrylamide gel; any existing bubbles should be removed after the agarose solidifies.
4. After the agarose solidifies, mount the gel in the electrophoresis unit and fill the upper buffer chamber with 2× SDS running buffer and the lower buffer chamber with 1× SDS running buffer using the volume specified by manufacturer. The gels should be run in the dark (to prevent dye-bleaching) by either covering the apparatus with opaque protective covering or in a darkened room.
5. During the electrophoresis, the power through the gel should be limited at 1.5 W and should proceed for 14 h or until the dye front reaches slightly above the bottom edge of the gel. At the completion of electrophoresis, remove the gels from the glass plates.
6. The gels can be scanned immediately or kept in the SDS buffer overnight in the dark.

3.8. Scanning Gels

This protocol is written specifically for scanning the DIGE gels using Typhoon (GE Healthcare) gel imager.

1. Allow the Typhoon scanner to warm up for 15 min prior to scanning the gels.
2. Use the gel guides for placing the gel on top of the scanner plate. Orient the gel on the Typhoon scanner with the short plate down and the notched end to the left (This will ensure the same scanning and analysis orientation as if using the gel to pick protein spots with ETTAN Spot Picker from GE Healthcare).
3. Change the settings on the scanner in order specified below:
    (a) Fluorescence acquisition mode
        (i) Click on settings to change the lasers and Photo Multiplier Tube (PMT) voltages. Start with a general prescan with the PMT setting at 600 V for all three lasers. Typically, Cy-2 and Cy-3 scans need a higher PMT voltage to saturate the spot in comparison to Cy-5 dye scans.

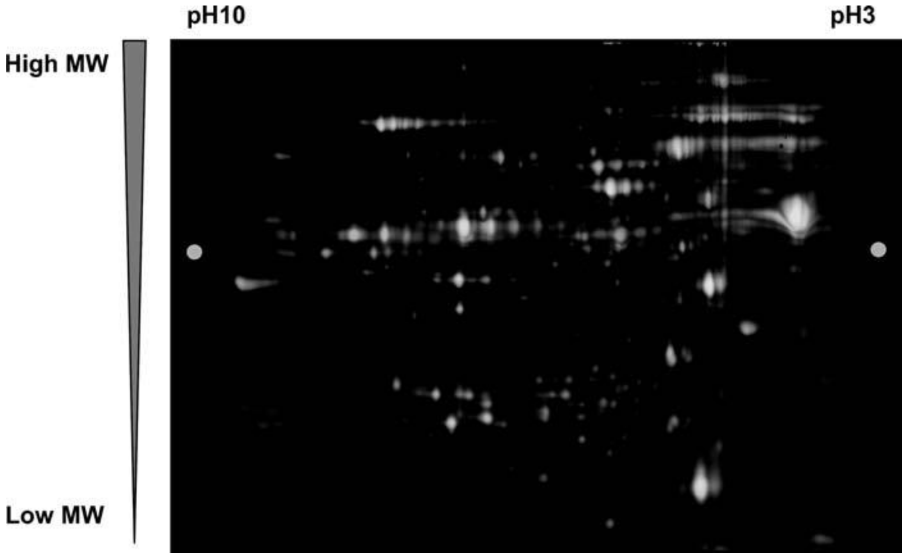


Fig. 2. Scan overlay of a 2-dye (Cy-3 and Cy-5) DIGE analysis. (*See Color Plates*)

   (ii) Choose Cy-2/FAM scan (blue laser)
   (iii) Choose Cy-3 scan (green laser)
   (iv) Choose Cy-5 laser (red laser)

4. Set the scan area for the size of the gel. The large gels are DALT™ gels (Ge Healthcare, Piscataway, NJ) and have preset area for scanning. Choose one gel option if scanning one large format gel. Use "User Select" option to mark the area for scanning if using smaller gels or custom size gels.
5. Set low resolution for the prescan to determine the PMT voltages for the saturation of spot signal on the gels. Once the PMT voltages have been optimized for spot saturation, the gels should be rescanned using a higher resolution setting.
6. The three dye scans can be overlaid (Fig. 2) using any commercial software for DIGE analysis. **Figure 2** shows the overlay of a 2-dye DIGE proteomic analysis where the control is labeled with Cy-3 (green spots) dye and the perturbed sample is labeled with Cy-5 dye (red spots). CY-2 scan has been omitted for visual comparsion of gel spots in a traditional 2-dye DNA microarray color analysis with red, green and yellow spots.

3.9. Fixing the Gels for Sypro Staining

1. After scanning, the gel can be stored in ~400 mL of gel fixing solution (10% methanol and 5% acetic acid in water solution). Do not use concentrations higher then specified to prevent stripping of the gel from the glass plate.
2. Pry open the gel with the spatula like tool. Keep the smaller plate down, pry lightly from the top and bottom edges. For

precast gels, the plates can be thrown away in the broken glass container; otherwise, rinse the gel and put it in the cleaning solution in the ice box.

3. Cover gels with SyproRuby and stain overnight. Polypropylene dishes are most appropriate for Sypro staining (glass dishes are not recommended). Since all fluorescent proteins are light sensitive, gel containers should be covered with an opaque cloth or aluminum foil or placed in the dark. The Sypro stain solution can be used twice (with virtually no loss of the staining activity).
4. Remove the gel from Sypro stain solution and wash in 10% (v/v) ethanol, 7% (v/v) acetic acid in water for 30 min.
5. Decant the solution and wash in dH_2O for 15 min.
6. Scan and analyze the gel.

3.10. Spot Picking

The gels spots of interest (e.g., showing change in intensity between perturbed and control sample) can be picked into a 96-well plate using a robotic gel picker.

3.11. Digestion of Protein Gel Plugs

Wipe down all surfaces in the hood with methanol/water moistened lint-free cloth, including the outside of all your tubes (make sure not to wipe off the labeling!), the outside and inside of the Speed Vac and centrifuge, tube racks, bottles etc. Wipe razor blades with methanol-soaked lint-free cloth. This protocol can be used with automated gel-digestion stations like ProPrep from Genomic Solutions (*see* **Note 21**).

1. Remove the water from the wells carefully so as not to remove the small gel plug.
2. Add 100 µL of buffer 1 and shake on a 96-well plate shaker for 10 min.
3. Remove the buffer from the wells carefully so as not to remove the small gel plugs.
4. *First wash:* Add 100 µL of buffer 2 and shake on a 96-well plate shaker for 10 min.
5. Incubate at RT for 20 min.
6. Remove the buffer from the wells carefully so as not to remove the small gel plugs.
7. *Second wash*: Add 100 µL of buffer 2 and shake on a 96-well plate shaker for 10 min.
8. Incubate at RT for 20 min.
9. Remove the buffer from the wells carefully so as not to remove the small gel plugs. The gel plugs should be completely destained by now even if using coomassie stain; if not, then repeat the buffer 2 wash one more time.

10. *ACN wash:* Add 200 µL of 100% ACN and incubate at RT for 10 min.
11. Remove the buffer from the wells carefully so as not to remove the small gel plugs. (The gel plugs are now whitish small dry pieces of polyacrylamide).
12. Speed Vac the gel pieces to complete dryness (~5 min).
13. Prepare the trypsin solution on ice (one Promega vial contains 20 µg of protein; add 2,500 µL of trypsin rehydration buffer for a final concentration of 8 ng/µL).
14. Add 15 µL of trypsin solution to the dehydrated gel pieces.
15. Rehydrate the gel pieces on ice or at 4°C for 10 min. Cover with a 96-well plate cover and put in a Ziplock bag.
16. Incubate at 37°C for a period ranging from 4 h to overnight.

3.12. Extraction of Peptides

1. If necessary, briefly spin the 96-well plate to pull the buffer from the wall to the bottom of the wells.
2. Transfer the digest solution (aqueous extraction) into a clean 96-well plate.
3. Add 30 µL (enough to cover) of 50% ACN/5% formic acid to gel pieces, and shake on the 96-well plate shaker for 20–30 min.
4. Spin and transfer the solution to the collection plate.
5. Repeat **steps 3** and **4**.
6. Speed Vac to completely lyophilize the peptides.

3.13. MS Analysis

The peptides are resuspended in the appropriate buffer for either MALDI (using α-cyano-4-hydroxycinnamic acid (CHCA) matrix in 50% ACN/1% TFA at a concentration of 10 mg/mL) or LC ESI-MS analysis (2% ACN, 0.2% formic acid in MS-grade water).

4. Notes

1. All chemicals should be of the highest purity (electrophoresis grade or better). Water should be double distilled or MS-grade. Consult the MSDS for proper material handling instructions and precautions.
2. Use proper protective equipment, especially lab coats and gloves for protection from chemicals and to prevent keratin contamination.
3. Do not heat any solutions containing urea above 30°C as urea is degraded to isocyanate, which will carbamylate the proteins in the sample, thus changing their isoelectric points.

4. Limit exposure of the dyes and dye labeled proteins to light at all times to prevent photobleaching of the dyes.
5. Avoid water at all steps during dye resuspension as that will degrade the NHS esters.
6. Always handle the strips with forceps, taking care not to stab the gel.
7. Extreme precaution should be taken when handling acrylamide/bisacrylamide as it is a neurotoxin.
8. The solution can be prepared in advance and stored at RT. Take precautions when handling butanol as it is harmful if inhaled or ingested.
9. Use a new bottle of DMF (or one that is <3 months old and that has been properly sealed to prevent exposure and leaks).
10. The solutions will appear colored after adding the DMF.
11. Measure the final volume of the solution in the tube to ensure that the volume of the protein solution does not exceed the maximum for the size of the IEF strip being used.
12. Bromophenol blue is used as a visual tracking dye.
13. Measure the final volume of the solution in the tube to ensure that the volume of the protein solution does not exceed the maximum for the size of the IEF strip being used.
14. Do not put a heavy object on top of the IEF unit as it may lead to arcing.
15. The current should <50 μA per strip to prevent overheating and burning of the strip, which may also damage the instrument. Make sure that the strip holder electrodes make contact with the + and – plates.
16. Increase the focusing step when using larger amounts of the protein sample in preparative gels use larger size IPG strips so as not to exceed the maximum permissible protein load for the size of the IPG Strip being used.
17. Focusing for substantially longer than recommended will cause horizontal streaking and loss of proteins (so-called “over-focusing” phenomenon).
18. Program the IEF instrument according to volt × hours and not according to total time of the protocol step. Since the voltage does not immediately ramp up to the desired maximum, programming by time for each step will stop each step before the proteins are properly focused as soon as the maximum time for that step is reached. For example, programming the focusing step for 4 h at 8,000 V would stop the protocol step as soon as 4 h is reached irrespective of the final voltage. Instead, programming 32,000 Vh (=4 h × 8,000 V) for the same protocol step ensures that the step ends only after the focusing from the voltage application has been achieved.

19. Salt and buffer ions, large quantities of protein and higher concentrations of IPG buffer in the sample will increase the time needed for effective IEF or even prevent focusing.
20. Clean each strip holder to remove residual protein with strip holder cleaning solution.
21. This procedure should never be carried out on a bench that is in close proximity to a high traffic area to avoid keratin contamination; it is in fact recommended that this procedure be performed in a laminar flow hood to minimize keratin contamination.

Acknowledgments

I would like to thank Dr. Anup Singh and Dr. Joe Schoeniger for enabling DIGE research at Sandia National Laboratories (in Livermore, CA), Carrie Kozina, Jaime Lachmann and George Buffleben for their efforts in effectively establishing the DIGE pipeline at Sandia National Laboratories. Financial support for this research was provided by the Sandia National Laboratories Laboratory Directed Research and Development Program, award #641610. Sandia is a multi-program laboratory operated by the Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

References

1. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250, 4007–21.
2. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 18, 2071–7.
3. Righetti, P. G., Castagna, A., Antonucci, F., Piubelli, C., Cecconi, D., Campostrini, N., Antonioli, P., Astner, H., and Hamdan, M. (2004) Critical survey of quantitative proteomics in two-dimensional electrophoretic approaches. *J Chromatogr A* 1051, 3–17.
4. Karp, N. A., Kreil, D. P., and Lilley, K. S. (2004) Determining a significant change in protein expression with DeCyder during a pair-wise comparison using two-dimensional difference gel electrophoresis. *Proteomics* 4, 1421–32.
5. Gong, L., Puri, M., Unlu, M., Young, M., Robertson, K., Viswanathan, S., Krishnaswamy, A., Dowd, S. R., and Minden, J. S. (2004) Drosophila ventral furrow morphogenesis: a proteomic analysis. *Development* 131, 643–56.
6. Shaw, M. M., and Riederer, B. M. (2003) Sample preparation for two-dimensional gel electrophoresis. *Proteomics* 3, 1408–17.

Chapter 6

Liquid-Chromatography-Mass Spectrometry of Thylakoid Membrane Proteins

Christian G. Huber, Anna-Maria Timperio, Hansjörg Toll, and Lello Zolla

Summary

The present chapter describes methods for the separation and identification of photosynthetic proteins of thylakoid membranes present in chloroplasts by using different detergents, high-performance liquid chromatography and mass spectrometry. Thylakoid membranes represent a good model for setting up analytical methods suitable for membrane protein characterization.The first step in the procedure is the preparation of purified membrane fractions from plant tissues, followed by the fractionation of membrane proteins by differential solubilization using different detergents. Thus, several protein complexes can be isolated, collected, separated by ion-pair reversed-phase chromatography and detected online by UV-absorption and/or mass spectrometry. Finally, identification of the eluting proteins is accomplished by comparing the molecular mass determined in silico with the molecular mass measured by mass spectrometry.

Key words: Ion-pair reversed-phase high-performance liquid chromatography, Electrospray ionization mass spectrometry, Thylakoid membranes, Photosystems, Differential solubilization, Hydrophobic proteins, Detergents.

1. Introduction

Membrane proteins can be classified into two broad categories – integral (intrinsic) and peripheral (extrinsic) – based on the nature of the membrane–protein interactions. Most biomembranes contain both types of membrane proteins. Some proteins are bound only to the surface of the membrane, whereas others have one region buried within the membrane and domains protruding on one or both sides of it. Through the use of various reagents in different sample preparation steps, the membrane proteins can be isolated according to their respective solubility and hydrophobic characteristics. Methods

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_6

for the separation and characterization of membrane proteins differ considerably, depending on the type of interaction of the protein with the membrane: proteins are either embedded in a lipid bilayer or associated with membrane structures. In the latter case, ionic interactions or hydrogen bonds are usually involved in the anchorage and therefore the protein may be removed by gentle solubilization with buffered salt solution and then analyzed. In the case of embedded proteins, they must first be extracted from the membrane structures and then isolated. In practice, solubilization of the membrane is the most suitable approach to recover the membrane proteins *(1)*. To this end, chaotropic reagents such as urea and guanidine hydrochloride are now used less frequently, while solubilization using detergents is more practical. If mass spectrometry is to be employed as the detection and characterization method, it is also advisable to use less denaturing detergents such as 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), or non-ionic detergents such as *n*-octyl-β-D-glucopyranoside (OD) or *n*-dodecyl-β-D-maltoside (β-DM) for solubilization *(1)*.

The effort necessary for solubilization of the protein increases with the growing complexity of the membrane structure. We shall discuss below different approaches for the complete solubilization of spinach thylakoid membranes of chloroplasts, which may serve as a general model for protein solubilization from membranes so that the methods can be generalized to many other species.

Thylakoid membranes contain the photosynthetic apparatus consisting of a large number of proteins. Approximately 50 different protein subunits are assembled into two main complexes: photosystem I (PSI) and photosystem II (PSII), having different molecular masses, different sizes and hydrophobicities *(2, 3)*, which makes them good models for setting up analytical methods suitable for membrane proteins. It will be shown that a preparation of membrane proteins according to their hydrophobic characteristics can be achieved by selective extraction. Once the membrane proteins have been pre-treated, several chromatographic and mass spectrometric methods may be used for the characterization of the isolated proteins *(4–6)*.

In many cases, membrane proteins are organized as multimeric units, therefore no single technique is likely to be sufficient for the complete characterization of the different proteins. It is important to select a battery of complementary techniques which will ensure that the necessary level of separation and characterization is obtained. With regard to the fractionation of membrane proteins, several chromatographic or electrophoretic separation strategies including reversed-phase- or ion-pair reversed-phase high-performance liquid chromatography (RP-HPLC, IP-RP-HPLC) *(7, 8)*, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) *(9)*, blue-native gel electrophoresis (BN-GE) *(10)*, affinity chromatography (AC), or ion-exchange high-performance liquid chromatography (IEX-HPLC) are applicable in principle. Nevertheless, due to the hydrophobic nature of the membrane proteins, they tend to adsorb, sometimes irreversibly, onto the surface of sample containers, connecting tubing, and column packings. Hence, methods involving strong denaturing conditions and/or detergents, such as IP-RP-HPLC or SDS–PAGE are usually preferred.

Mass spectrometry (MS) utilizing electrospray ionization (ESI) *(11, 12)* or matrix-assisted laser desorption/ionization (MALDI) *(13, 14)* as ionization techniques has been shown to be applicable to the mass spectrometric investigation of proteins and even protein complexes ranging in molecular mass from a few thousand to well over 200.000 Da. However, due to ion suppression and competitive ionization, MS of membrane proteins is impaired by the presence of detergents or high concentrations of chaotropes or salt necessary to keep them in solution. In this case, it is advisable to interface MS online to liquid chromatography as the purification method to remove the interfering reagents from the protein sample immediately before transfer to the mass spectrometer *(15)*. Protein identification is frequently based on intact molecular mass measurements *(5, 6)*. This is mainly because the hydrophobic transmembrane domains contain relatively few recognition sites for the commonly utilized proteases, such as trypsin or chymotrypsin, resulting in few relatively large fragments, which are not very suitable for identification by peptide mass fingerprinting (PMF) or peptide fragment fingerprinting (PFF).

2. Materials (*See* Notes 1 and 2)

2.1. Hydroponic Culture of Spinach Plants

1. 1% (w/v) solution NaOCl in deionized water.
2. 30% (w/v) solution of poly(ethyleneglycol) 3,000 in deionized water.
3. Solution 1 SB: 0.5 M KNO_3, 0.2 M $MgSO_4 \cdot 7H_2O$, 0.2 M KH_2PO_4, 0.1 M NaCl.
4. Solution 2 SB: 1 M $Ca(NO_3)_2 \cdot 4H_2O$.
5. Solution 3 SB: 9.14 M $MnCl_2 \cdot 4H_2O$, 4.3 M H_3BO_3, 0.12 M Na_2MoO_4, 2.40 M $ZnSO_4 \cdot 7H_2O$, 0.37 M $CuSO_4$.

2.2. Isolation of Intact Chloroplasts and of Thylakoid Membranes

1. Buffer B1: 20 mM *N*-(2-hydroxy-1,1-bis(hydroxymethyl) ethyl)glycine (TRICINE) pH 7.8, 0.30 M sucrose, 5.0 mM $MgCl_2$.

2. Buffer B2: 20 mM TRICINE pH 7.8, 70 mM sucrose, 5.0 mM $MgCl_2$.
3. Buffer B3: 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.3, 1.5 mM NaCl, 5.0 mM $MgCl_2$.
4. Buffer B4: 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.3, 15 mM NaCl, 5.0 mM $MgCl_2$.
5. Acetone.
6. Protease inhibitors: 100 mM phenylmethylsulfonyl fluoride (PMFS) in B1 buffer, 100 mM benzamidine in B1 buffer, 500 mM a-aminocaproic acid in B1 buffer.

2.3. Fractionation of PSII and PSI Membrane Proteins Using Digitonin as Detergent

1. Re-suspension buffer: 15 mM TRICINE, pH 7.8, 0.10 M sorbitol, 10 mM NaCl, 5 mM $MgCl_2$.
2. 1.0% (w/v) digitonin in water (*see* **Note 3**).

2.4. Fractionation of PSII and PSI Membrane Proteins Using a Dextran–Polyethyleneglycol Biphasic System

1. Re-suspension buffer: 10 mM sodium phosphate, pH 7.4, 20 mM sucrose, 3 mM NaCl, 1 mM $MgCl_2$.
2. 50% (w/w) dextran T500 (Pharmacia, Uppsala, Sweden) solution in deionized water.
3. 25% (w/w) polyethyleneglycol (PEG) 4000 (Carbowax PEG 3350, Union Carbide, New York, NY) solution in deionized water.
4. Vibra-cell ultrasonic processor Model VC 500 (Sonics and Materials, Danbury, CT).

2.5. Ion-Pair Reversed-Phase High-Performance Liquid Chromatography and Electrospray Ionization Mass Spectrometry

1. Solution of 10 ng/mL lysozyme (from chicken egg white) in water–acetonitrile (80:20, v/v) containing 0.05% trifluoroacetic acid.
2. Eluent A: 0.05% (v/v) trifluoroacetic acid in water.
3. Eluent B: 0.05% (v/v) trifluoroacetic acid in acetonitrile.
4. Monolithic PS-DVB column 50 × 0.1 mm i.d. (LC Packings – A Dionex Company, Amsterdam, the Netherlands) (*see* **Note 4**).
5. Vydac Protein C4 column, 250 × 4.6 mm i.d. or 250 × 1 mm i.d. (Grace Vydac, Hesperia, CA) (*see* **Note 4**).
6. High-performance liquid chromatograph: pumping system capable of high-pressure or low pressure gradient mixing at flow rates of 0.05–1.5 mL/min for 1 or 4.6 mm i.d. columns or 1–3 mL/min for 0.2 mm i.d. columns, manual or automatic injector capable of injecting volumes of 0.5–100 mL, column thermostat for operation of the columns at elevated temperature (advisable to improve separation efficiency, but not mandatory), UV-spectrometer set at 220 or 280 nm, UV-detection cell of 8–12 mL volume for 4.6 mm

i.d. columns, 1–2 μL/min volume for 1 mm i.d. columns and 3 nL for 0.2 mm i.d. columns.

7. Mass spectrometer with any type of analyzer (linear quadrupole, quadrupole or linear ion trap, linear or reflectron time-of-flight mass analyzer, ion-cyclotron resonance mass analyzer) suitable for interfacing via electrospray ionization, mass range at least *m/z* 500–1,500, preferably *m/z* 500–3,000.

3. Methods (*See* Note 5)

3.1. Hydroponic Culture of Spinach Plants

1. Hydroponic culture is carried out in a growth room with a 14-h light period (260–350 mE/m^2/s) at a temperature of 28°C (day) and 20°C (night).
2. Seeds of spinach (*S. oleracea* L.) are disinfected in 1% (w/v) of NaOCl solution for 10 min followed by thorough washing in deionized water.
3. Seeds are transferred into a beaker containing 30% (w/v) PEG 8000 and kept for 72 h, then washed in 1% (w/v) of NaOCl solution and transferred into deionized water for 5 h for seed rehydration with continuous areation.
4. Seeds are then germinated on moist filter paper for 2–3 days.
5. Germinated seeds are sown in moist vermiculite and grown for about 10 days to the four-leaf stage.
6. After the initial growth period in the seedbed, seeds are carefully removed from the seedbed and thoroughly washed free of any adhering particles under tap water.
7. Seedlings are subsequently transferred to PVC pots (8 cm in diameter and 17 cm in height) containing 500 mL of nutrient solution (0.5% (v/v) 1 SB solution, 0.25% (v/v) 2 SB solution, 0.05% (v/v) 3 SB solution, diluted with deionized water) with continuous aeration pH 5.8.

3.2. Isolation of Thylakoid Membranes from Intact Chloroplasts

1. Plants are dark-adapted for 24 h at 4°C before harvesting the leaves. Subsequently, 50 g of leaves (*see* Note 6) are harvested and powdered in a mortar under liquid nitrogen and homogenized in 200 mL ice-cold B1 buffer containing 1 mL each of the protease inhibitor solutions.
2. The homogenization is followed by filtration through one layer of Miracloth (Calbiochem, San Diego, CA) and centrifugation of the filtrate at 4,000 × *g* for 10 min at 4°C.

The pellet is re-suspended in B1 buffer and centrifuged as above. This second pellet is re-suspended in B2 buffer and centrifuged at 4,500 × g for 10 min.

3. Pellets containing the thylakoid membranes are then re-suspended in B3 buffer. The chlorophyll concentration is determined in a 5.0-μL aliquot upon dilution with 195 μL bidistilled water and 800 μL acetone (1,000 μL total volume) *(16)*. After centrifugation at 10,000 g for 5 min, the UV absorption, the UV absorption of the supernatant is measured at 710, 663, and 645 nm. The chlorophyll concentration is estimated according to the following equation (A_{xxx}, UV absorbance at the respective wavelength):

$$c(\mathit{chlorophyll}) = \frac{\text{total volume}}{\text{volume aliquot}}\left[8.02(\mathrm{A}_{663} - A_{710}) + 20.02(\mathrm{A}_{645} - A_{710})\right]\ \mu\text{g/ml}$$

4. This sample is suitable for direct chromatographic analysis of the whole membrane protein fraction of thylakoids. The determined chlorophyll concentration is used later to estimate the amount of sample for injection onto the chromatographic column.
5. For storage of the sample, the pellet is re-suspended in 50 mM TRICINE, pH 7.8, and 100 mM sorbitol stored at –20°C. Alternatively, the pellet can be frozen in liquid nitrogen and stored at –80°C for later experiments (*see* **Note 7**).

3.3. Fractionation of PSII and PSI Membrane Proteins Using Digitonin as Detergent

1. The primary thylakoid membrane suspension is diluted to a concentration of 200 mg chlorophyll per mL with re-suspension buffer.
2. Recrystallized digitonin (1.0% (w/v) in water) is added at 0°C to the stirred thylakoid membrane suspension to give a final concentration of 0.40% (*see* **Note 8**).
3. The digitonin treatment is continued for 2 min.
4. Lysis is terminated upon 70-fold dilution of the sample at 0°C with the resuspension buffer (*see* **Note 9**).
5. Differential centrifugation is performed according to the method of Anderson and Boardman *(17)*. Briefly, the suspension is centrifuged at 1,000 × *g* for 10 min, 10,000 × *g* for 30 min and 40,000 × *g* for 30 min. The pellets, containing the grana fraction of thylakoid membranes are collected. The supernatants containing the stroma fraction of thylakoid membranes are combined and centrifuged at 144,000 × *g* for 60 min (stoma in the pellet) to retrieve the stroma fraction in the pellet. Grana lamellae are rich in PSII, while stroma lamelle mainly contain PSI.

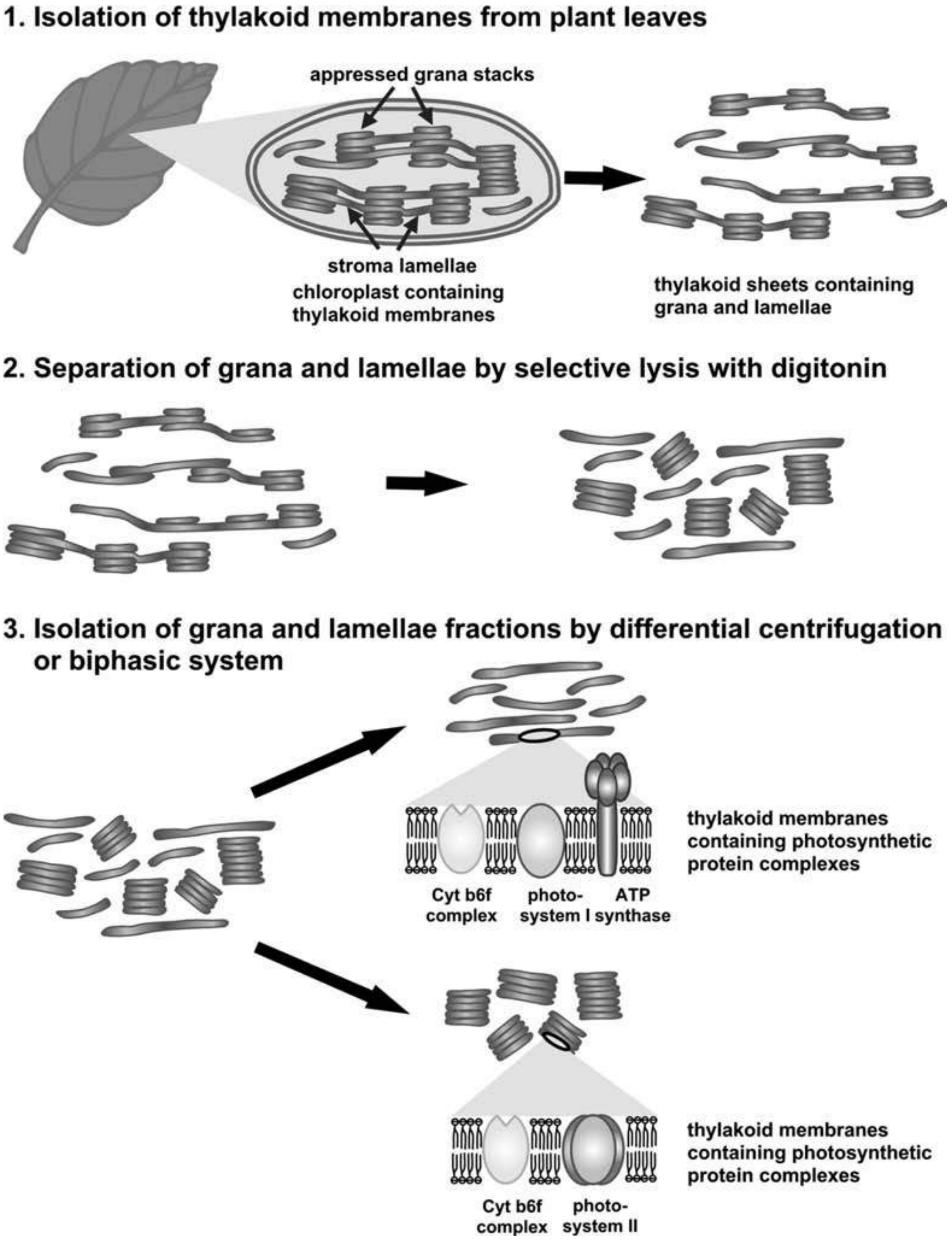

1. Isolation of thylakoid membranes from plant leaves
appressed grana stacks
stroma lamellae
chloroplast containing
thylakoid membranes
thylakoid sheets containing
grana and lamellae
2. Separation of grana and lamellae by selective lysis with digitonin
3. Isolation of grana and lamellae fractions by differential centrifugation
or biphasic system
thylakoid membranes
containing photosynthetic
protein complexes
Cyt b6f
complex
photo-
system I
ATP
synthase
thylakoid membranes
containing photosynthetic
protein complexes
Cyt b6f
complex
photo-
system II

Fig. 1. Scheme of sample preparation and separation of PSI and PSII by differential solubilization. (Reproduced from *(18)* by permission of Elsevier B. V.).

6. The total scheme of sample preparation for the separation of PSI and PSII membrane protein fractions is illustrated in **Fig. 1**. The obtained grana and stroma fractions are re-suspended in re-suspension buffer and ready for analysis by HPLC-MS.

3.4. Separation of PSII from PSI Membrane Proteins Using a Dextran–Polyethyleneglycol Biphasic System

1. Three hundred sixty-six milligrams of 30% dextran solution are added to 268.6 mg (about 267.5 μL) of re-suspension buffer to obtain a final volume of 1.0 mL. One hundred sixty milligrams of 25% PEG solution are added to this solution, mixed, and centrifuged at 3,000 × *g*. Due to the incompatibility of the polymer solutions, the system forms an aqueous, biphasic system, where dextran forms the more hydrophilic, denser, lower phase and polyethylene glycol the more hydrophobic, less dense, upper phase. The final volume of the biphasic system is about 1.2 mL.
2. One hundred microliters of thylakoid membranes suspended in re-suspension buffer containing 2.95 mg/mL chlorophyll are put on top of the biphasic system. The sample system is mixed at 4°C and sonicated using a Vibra-cell ultrasonic processor equipped with a 1/2-in. horn. Sonication must be performed in six 30-s bursts, alternating with 60-s resting intervals, in a cylindrical aluminium tube immersed in ice and water (*see* **Note 10**). The ultrasonic intensity output setting is 7, with 20% duty pulses. Centrifugation at 2,000 × *g* for 3 min is performed to separate the phases.
3. The upper and lower phases, enriched in thylakoid fragments from stroma and grana lamellae, respectively, are divided and washed (*see* **Fig. 2**). Grana lamellae are rich in PSII, while stroma lamellae mainly contain PSI.

3.5. Setup for IP-RP-HPLC–ESI-MS Analysis

1. The separation column is equilibrated with water/0.05% TFA-acetonitrile/0.05% TFA 95:5 and two blank gradients of 5–95% acetonitrile are run to remove strongly adsorbed compounds from the column. Subsequently, the column is equilibrated with the starting eluents employed for gradient elution, typically water/0.05% TFA-acetonitrile/0.05% TFA 95:5 for the monolithic column and water/0.05% TFA-acetonitrile/0.05% TFA 90:10 for the butyl-silica column.
2. The mass spectrometer is calibrated and tuned according to the manufacturer's recommendations.
3. Fine tuning of the instrument is performed by direct infusion of a solution of 10 ng/μL lysozyme in water–acetonitrile (80:20, v/v) containing 0.05% trifluoroacetic acid at the flow rate that is later used to transfer the proteins into the mass spectrometer, typically 2–10 μL/min. The 7+ charge state at *m/z* 2,044.590 or the 10+ charge state at *m/z* 1,431.5 is utilized to maximize both total ion current stability and signal-to-noise ratio of the mass spectrometric signal. If the mass spectrometer allows a second signal, e.g., at *m/z* 1,590.461, this may be monitored to optimize ion transmission over a broader mass range (*see* **Note 11**).

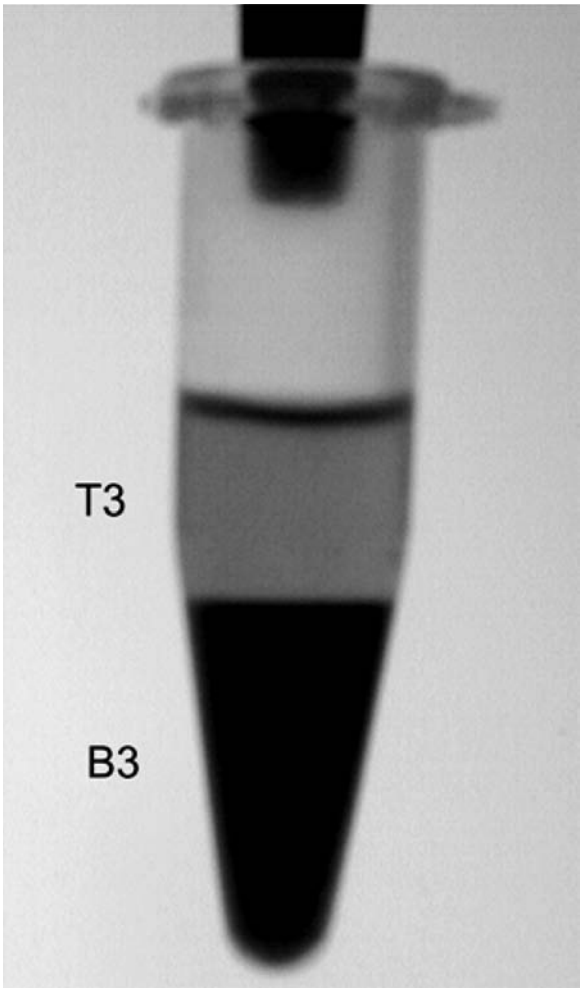


Fig. 2. Grana and stroma phases obtained upon lysis with dextran and PEG followed by ultracentrifugation.

4. Monolithic columns of 0.2 mm i.d. are directly interfaced to the electrospray ion source of the mass spectrometer and operated at a flow rate of 1–3 μL/min (*see* **Fig. 3a**) (*see* **Note 12**). Columns of 4.6 mm i.d. are eluted with a flow rate of 1 mL/min and the column effluent is usually split to introduce only a flow of about 3–10 mL/min into the mass spectrometer. Splitting is accomplished with a metal Tee-piece and a restriction capillary prepared from a 100- to 150-cm long piece of 75 μm i.d. fused silica tubing (*see* **Fig. 3b**). The primary flow delivered by the pump as well as the length of the fused silica capillary are adjusted to obtain the desired flow rate through the capillary column. The flow delivered to the mass spectrometer is measured by connecting a graduated 10-μL syringe from which the plunger

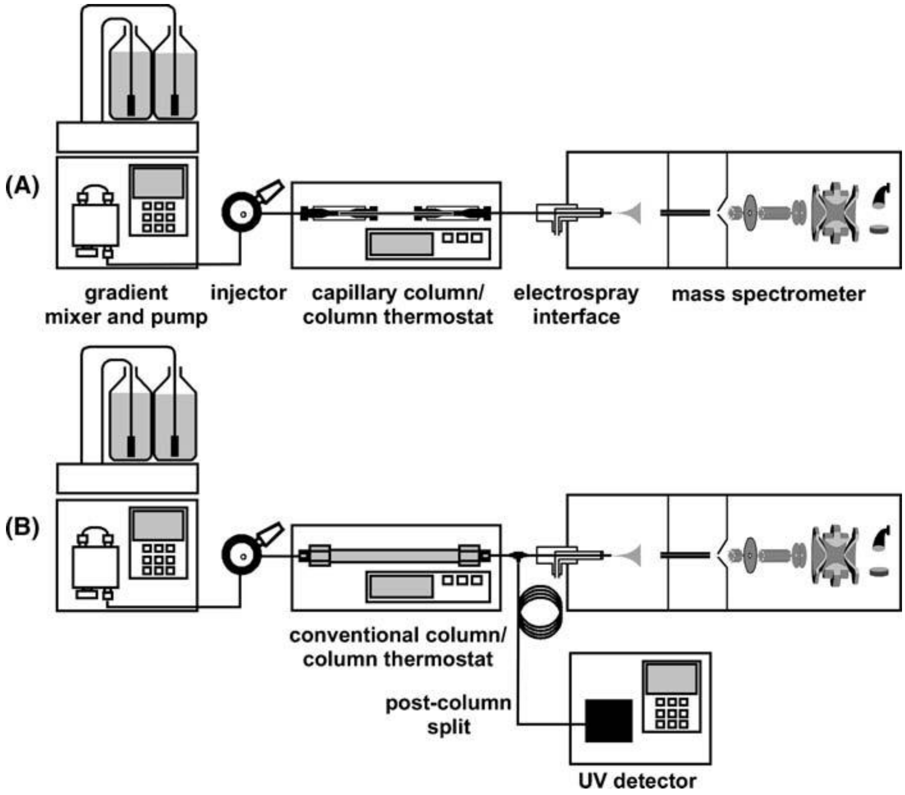

(A)
gradient mixer and pump
injector
capillary column/ column thermostat
electrospray interface
mass spectrometer
(B)
conventional column/ column thermostat
post-column split
UV detector

Fig. 3. Instrumental setups for the analysis of membrane proteins by conventional, microbore, or capillary IP-RP-HPLC-UV-ESI-MS.

has been removed and using a stopwatch to measure the time to deliver a volume of about 5 mL.

3.6. Analysis of a Total Thylakoid Protein Extract by Capillary Ion-Pair Reversed-Phase High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry

1. Five hundred nanoliters of membrane protein extract from thylakoids containing approximately 5–10 ng chlorophyll are injected directly onto the 50 × 0.2 mm i.d. monolithic column and eluted with a 10-min gradient of 5–25% acetonitrile in 0.05% TFA, followed by 25–75% acetonitrile in 0.05% TFA in 80 min at a column temperature of 70°C (*see* Note 13). Mass spectra are continuously recorded in a mass range of *m/z* 500–1,500 (2,000). A representative reconstructed total ion current chromatogram of a membrane protein preparation from spinach thylakoids is illustrated in Fig. 4. Although complete separation of the mixture is generally impossible due to the high complexity of the sample, the extracted ion chromatograms for selected protein *m/z* values (Fig. 4) clearly show that membrane proteins elute from polystyrene-divinylbenzene monolithic columns as sharp and well-defined peaks.

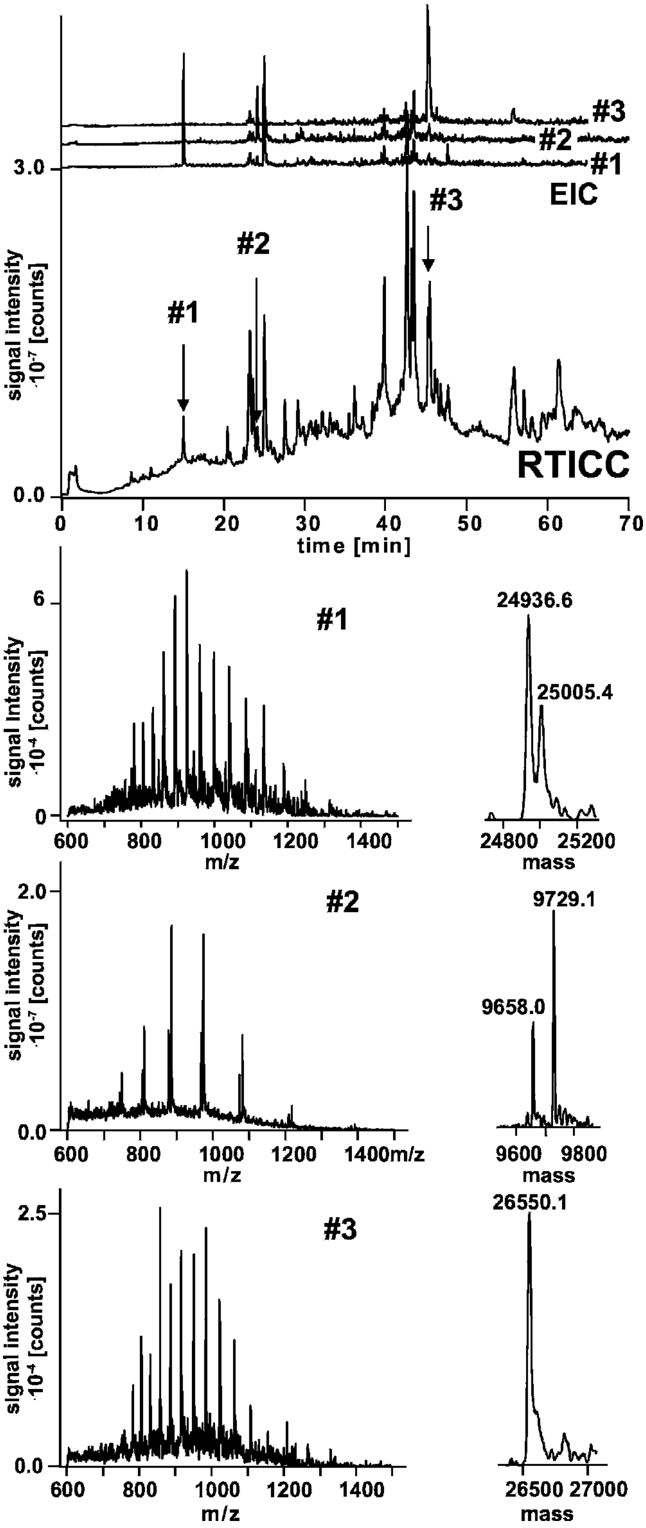


Fig. 4. Reconstructed total ion current chromatogram (RTICC), extracted ion chromatograms (EIC) and raw and deconvoluted mass spectra obtained upon IP-RP-HPLC-ESI-MS analysis of a whole membrane protein extract from thylakoids.

2. After protein elution, the column is washed for 10 min with 95% acetonitrile in 0.05% TFA. The column is subsequently re-equilibrated with the starting eluent for 15 min before injection of a new sample (*see* **Notes 14** and **15**).
3. For protein detection and identification, mass spectra are extracted from all major and minor peaks in the reconstructed total ion current chromatogram by averaging about 5–10 mass spectra. Proteins are detected by characteristic series of multiply charged ion species that can be deconvoluted into the intact molecular mass of the uncharged protein by well known deconvolution algorithms *(19)* (*see* **Fig. 4** and **Note 16**). In many cases, more than one protein mass is obtained after deconvolution, indicating that either different proteins or different isoforms of one protein are coeluting from the column (*see* **Fig. 4**).
4. Protein identification is performed based on comparison with elution patterns and intact molecular masses published in the literature *(5–8, 12, 18, 20, 21)*. Intact molecular masses can be calculated by extracting the protein sequence from the publicly available protein databases and using the "peptide mass" tool from the ExPASY Proteomics Server of the Swiss Institute of Bioinformatics http://www.expasy.org.. The sequence of the signal peptide must be deleted and the masses of eventual posttranslational modifications such as acetylation or phosphorylation need to be added in order to obtain the correct protein mass. Measured masses may significantly deviate from the theoretical masses, either because of unknown exact cleavage position of the signal peptide, sequence variants *(22)*, or unknown posttranslational modifications *(20)*. Nevertheless, elution position and approximate mass range in many cases suffice to identify the protein correctly. If sequence information is not available for the species under investigation, the mass range for the different membrane proteins may be deduced from the known protein sequences of other species.

3.7. Analysis of Grana and Stroma Protein Fractions by Microbore or Analytical IP-RP-HPLC-UV-ESI-MS

1. Microbore HPLC-ESI-MS: 20–50 mL of membrane protein extract from thylakoids containing approximately 125–250 ng chlorophyll are injected directly onto the 250 × 1 mm i.d. Vydac column and eluted at a flow rate of 50 mL/min with a 15-min gradient of 10–40% acetonitrile in 0.05% TFA, followed by 40–90% acetonitrile in 0.05% TFA in 60 min at a column temperature of 25°C (*see* **Note 12**). The column effluent is monitored by UV-absorbance detection at 214 nm. **Figure 5** depicts the chromatograms of a total thylakoid preparation (**Fig. 5a**) and the corresponding

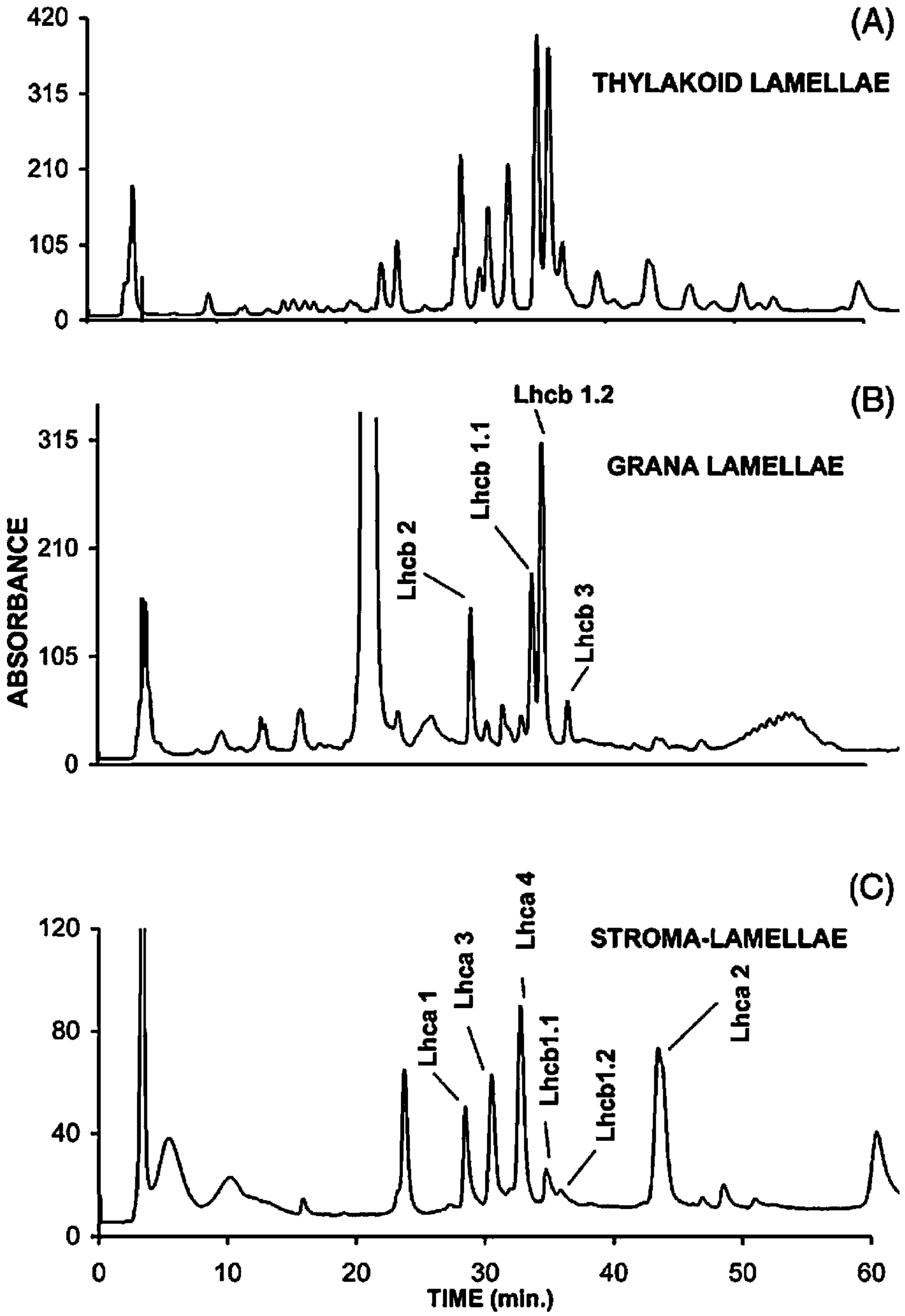


Fig. 5. UV chromatograms of membrane proteins isolated from thylakoid membranes. (**a**) Total thylakoid membrane preparation, (**b**) grana lamellae subfraction, (**c**) stroma lamellae subfraction. (Reproduced from *(18)* by permission of Elsevier B. V.).

grana (Fig. 5b) and stroma (Fig. 5c) fractions. In parallel, mass spectra are continuously recorded in a mass range of *m/z* 500–1,500 (2,000) (*see* Note 17).

2. Analytical HPLC-ESI-MS: 100 µL of membrane protein extract from thylakoids are injected onto the 250 × 4.6 mm i.d. Vydac column and separated at a flow rate of 1 µL/min under otherwise identical conditions as described above.
3. After protein elution, the column is washed for 10 min with 95% acetonitrile in 0.05% TFA. The column is subsequently re-equilibrated with the starting eluent for 15 min before injection of a new sample (*see* Notes 14 and 15).
4. For protein detection and identification, mass spectra are extracted from all major and minor peaks in the reconstructed total ion current chromatogram by averaging about 5–20 mass spectra. Proteins are detected by characteristic series of multiply charged ion species that can be deconvoluted to the molecular mass of the intact protein by well known de-convolution algorithms (*see* Note 16).
5. Protein identification is performed as described above, based on comparison with elution patterns and intact molecular masses published in the literature (Table 1) *(5–8, 12, 18, 20, 21)* (*see* Note 16).

Table 1
Retention times, molecular masses, and identification of proteins analyzed in spinach thylakoid grana and stroma lamellae

Thylakoid membrane fraction	HPLC elution times	M_r measured ± SD[a]	M_r expected (reference)	Accession no.	Protein identification
Grana lamellae	27.89	24,761 ± 0.9	24,788 *(23)*	P14279	Lhcb2
	32.41	24,936 ± 1.4	24,879 *(23)*	P07369	Lhcb1.1
	33.39	25,014 ± 0.5	25,015 *(24)*	P12333	Lhcb1.2
	35.06	24,321 ± 1.3	24,308 *(25)*	P27489	Lhcb3
Stroma lamellae	28.79	21,673 ± 1.5	21,982 *(25)*	P12360	Lhca1
	30.38	25,298 ± 1.0	25,340 *(25)*	P27522	Lhca3
	32.15	22,353 ± 0.5	22,336 *(25)*	S14305	Lhca4
	34.19	24,936 ± 0.5	24,943 *(25)*	P07369	Lhcb1.1
	35.08	25,014 ± 0.2	25,015 *(24)*	P12333	Lhcb1.2
	43.08	23,241 ± 1.6	23,079 *(25)*	P10708	Lhca2

[a)] Mean ± SD of three experiments are given

4. Notes

1. Unless otherwise specified, all components are of analytical reagent grade and available from Sigma-Aldrich/Fluka (St. Louis, MO) or Merck (Darmstadt, Germany). This protocol can be adapted for many other plants, such us cucumber, pea, Arabidopsis, and barley.
2. All solutions should be prepared in bidistilled water having a resistivity of 18.2 MΩ cm and total organic content of less than 5 ppb.
3. This detergent is hard to dissolve in water. Best results are obtained by heating the mixture at 80°C for 15 min and bringing the evaporation under control by the addition of water at the end of reaction, when necessary. (10 mL final volume 1%, w/v digitonin).
4. In order to protect the column from a particular or strongly adsorbed material, the use of 10-mm long precolumns is strongly recommended.
5. During all sample preparation steps it is very important to work in a cold room (4°C). All protocol steps should be performed in the dark. Use pre-cooled buffers, blender and centrifuge tubes (50 mL). Tube surfaces should be kept scrupulously clean. All preparations containing thylakoid membranes and membrane proteins are light sensitive and sensitive to degradation and/or oxidation. Hence, all procedures should be performed under light protection and cooling on ice.
6. For different amounts scale the quantities accordingly.
7. At this stage samples can be stored using different buffers, depending on their future use. For short term storage, the thylakoid preparation can be suspended in 50 mM TRICINE pH 7.8, 100 mM sorbitol and stored at 4°C. If the samples have to be stored for a long time (1–3 months), add glycerol 20%, v/v and store at –20°C. Remember to remove glycerol by centrifugation at 18,000 rpm (rotor JA20 Beckman Coulter supercentrifuge). This is necessary in the event of samples being loaded onto a sucrose gradient. For periods of 1 year and more, thylakoids may be stored under liquid nitrogen.
8. Add the digitonin to thylakoid membranes at 0°C in order to slow down its effect.
9. After incubation for different time periods, the digitonin effect can be stopped by diluting the sample with buffer by a factor of 70 and separating the membrane fractions

by ultracentrifugation. In this way the heavier grana membranes are separated from the lighter stroma membranes, which can be recovered from the supernatant.

10. This sonication procedure essentially breaks apart the two main compartments, grana and stroma lamellae *(26)*.
11. The tuning procedure has been shown to be essential to be able to detect the membrane proteins with good signal-to-noise ratios *(27)*. Optimal tuning conditions generally differ significantly between instruments. Sometimes, tuning parameters need to be set in a significantly different manner from those suggested by automatic tuning routines in order to achieve high sensitivity and spectrum quality, especially for high-molecular mass proteins.
12. The inner diameter of the transfer capillary to the ion source in combination with a 0.2 mm i.d. monolithic column should not be larger than 25 μm in order to avoid unacceptable band dispersion in the connecting tubing.
13. Gradient steepness may be adjusted according to the complexity of the sample. Very shallow gradients of 1–2 h are required to separate and detect the maximum number of proteins and to cover a broad dynamic range of protein concentrations.
14. The gradient dwell volume, i.e., the volume between the connection, where the gradient is formed and the column inlet may be considerably large for some pumping systems so that there is a significant time delay between programmed changes in eluent composition and actual changes realized in the column. Make sure that enough time passes to account for gradient delay.
15. Particular material or pieces of membranes getting trapped at the column inlet may lead to an increase in the column backpressure and, eventually, complete blockage, especially immediately after sample injection. In order to protect the analytical column from clogging, the utilization of a 5- to 20-mm long precolumn is recommended. The permeability of the precolumn may be restored by washing in back-flush direction. Columns may be washed by flushing repeated injections of 500 mL formic acid followed by washing with acetonitrile–water 90:10.
16. Deconvolution algorithms are usually implemented in the data analysis software components of the mass spectrometer.
17. At a flow-rate of 50 μL/min, UV and mass spectrometric detection may also be performed in series without split configuration.

References

1. Zolla, L. Chromatography of membrane proteins and lipoproteins. In Meyers, R.A. (ed.) Encylopedia of Analytical Chemistry. Wiley, Chichester (2000).
2. Hankamer, B., Barber, J., and Boekema, E.J. (1997) Structure and membrane organization of photosystem II in green plants. *Annu. Rev. Plant Mol. Biol.* 48, 641-671.
3. Boekema, E.J. et al. (2001) Green plant phtosystemI binds light-harvesting complex I on one side of the complex. *Biochemistry* 40, 1029-1036.
4. Zolla, L., Rinalducci, S., Timperio, A.M., and Huber, C.G. (2004) Separation and identification of photosynthetic antenna membrane proteins by high-performance liquid chromatography electrospray ionization mass spectrometry. *Eur. J. Mass Spectrom.* 10, 321-333.
5. Zolla, L., Rinalducci, S., Timperio, A.-M., and Huber, C.G. (2002) Proteomics of light-harvesting proteins in different plant species: analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. Photosystem I. *Plant Physiol.* 130, 1938-1950.
6. Zolla, L., Timperio, A.-M., Walcher, W., and Huber, C.G. (2003) Proteomics of light-harvesting proteins in different plant species: analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. Photosystem II. *Plant Physiol.* 131, 198-214.
7. Zolla, L., Bianchetti, M., Timperio, A.M., and Corradini, D. (1997) Rapid resolution by reverse-phase high performance liquid chromatography of the thylakoid membrane proteins of the photosystem II light-harvesting complex. *J. Chromatogr. A* 779, 131-138.
8. Zolla, L. et al. (1999) Isolation and characterization of chloroplast photosystem II antenna of spinach by reversed-phase liquid chromatography. *Photosynth. Res* 61, 281-290.
9. Peltier, J.B. et al. (2000) Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. *Plant Cell Physiol.* 12, 319-341.
10. Kügler, M., Jänsch, L., Kruft, V., Schmitz, U.K., and Braun, H.-P. (1997) Analysis of the chloroplast protein complexes by blue-native polyacrylamide gel electrophoresis (BN-PAGE). *J. Photosynth. Res.* 54, 35-44.
11. Cole, R.B. Electrospray Mass Spectrometry: Fundamentals, Instrumentation & Applications. Wiley, New York (1997).
12. Gomez, S.M., Nishio, J.N., Faull, K.F., and Whitelegge, J.P. (2002) The chloroplast grana proteome defined by intact mass measurements from liquid chromatography mass spectrometry. *Mol. Cell. Proteomics* 1, 46-59.
13. Karas, M. and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* 60, 2299-2301.
14. Szabò, I., Seraglia, R., Rigoni, F., Traldi, P., and Giacometti, G.M. (2001) Determination of photosystem II subunits by matrix-assisted laser desorption/ionization mass spectrometry. *J. Biol. Chem.* 276, 13784-13790.
15. Corradini, D., Huber, C.G., Timperio, A.-M., and Zolla, L. (2000) Resolution and identification of the protein components of the photosystem II antenna system of higher plants by reversed-phase liquid chromatography with electrospray mass spectrometric detection. *J. Chromatogr. A* 886, 111-121.
16. Porra, R.J., Thompson, W.A., and Kriedemann, P.E. (1989) Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophyll *a* and *b* extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta* 975, 384-394.
17. Anderson, J.M. and Boardman, N.K. (1966) Fractionation of the photochemical systems of photosynthesis. I. Chlorophyll contents and photochemical activities of particles isolated from spinach chloroplasts. *Bibl. Laeger* 112, 403-421.
18. Timperio, A.M., Huber, C.G., and Zolla, L. (2004) Separation and identification of the light harvesting proteins contained in grana and stroma thylakoid membrane fractions. *J. Chromatogr. A* 1040, 73-81.
19. Mann, M., Meng, C.K., and Fenn, J.B. (1989) Interpreting mass spectra of multiply charged ions. *Anal. Chem.* 61, 1702-1708.
20. Zolla, L., Rinalducci, S., Timperio, A.M., Huber, C.G., and Righetti, P.G. (2004) Intact mass measurements for unequivocal identification of hydrophobic photosynthetic photosystems I and II antenna proteins. *Electrophoresis* 25, 1353-1366.
21. Huber, C.G., Walcher, W.T.A.M., Troiani, S., Porceddu, E., and Zolla, L. (2004) Multidimensional proteomic analysis of photosynthetic membrane proteins by liquid extraction-ultracentrifugation-liquid chromatography-mass spectrometry. *Proteomics* 4, 3909-3920.
22. Huber, C.G., Timperio, A.-M., and Zolla, L. (2001) Isoforms of photosystem II antenna

proteins in different plant species revealed by liquid chromatography-electrospray ionization mass spectrometry. *J. Biol. Chem.* 276, 45755-45761.

23. Schwartz, E. and Pichersky, E. (1990) Sequence of two tomato nuclear genes encoding chlorophyll a/b-binding proteins of CP24, a PSII antenna component. *Plant Mol. Biol.* 15, 157-160.
24. Mason, J.G. (1989) Nucleotide sequence of a cDNA encoding the light-harvesting chlorophyll a/b binding protein from spinach. *Nucleic Acids Res.* 17, 5387-5394.
25. Schwartz, E. et al. (1991) Nucleotide sequence and chromosomal location of Cab11 and Cab12, the genes for the fourth polypeptide of the photosystem I light-harvesting antenna (LHCI). *FEBS Lett.* 280, 229-234.
26. Andreasson, E., Svensson, P., Weibull, C., and Albertsson, P.-A. (1988) Separation and characterization of stroma and grana membranes: evidence for heterogeneity in antenna size of both photosystem I and photosystem II. *Biochim. Biophys. Acta* 936, 339-350.
27. Oberacher, H., Walcher, W., and Huber, C.G. (2003) Effect of instrument tuning on detectability of biopolymers in electrospray ionization mass spectrometry. *J. Mass Spectrom.* 38, 108-116.

Chapter 7

High Accuracy Mass Spectrometry in Large-Scale Analysis of Protein Phosphorylation

Jesper V. Olsen and Boris Macek

Summary

With the appearance of a new generation of high-performance hybrid mass spectrometers, high accuracy (sub-parts-per-million) mass spectrometry is becoming increasingly available to a wider scientific community. Here we discuss the advantages of such mass spectrometric instrumentation in the global analysis of protein phosphorylation. We describe a detailed workflow for fractionation and enrichment of phosphopeptides from digests of whole cell/tissue lysates by strong cation exchange and TiO_2 chromatography and their subsequent measurement on an LTQ-Orbitrap mass spectrometer under several acquisition regimes.

Key words: Protein phosphorylation, Phosphoproteomics, Phosphopeptide enrichment, TiO_2 chromatography, Mass spectrometry, High mass accuracy, Orbitrap, FT MS.

1. Introduction

In a typical mass spectrometry-based proteomics experiment *(1)*, average absolute measurement mass accuracy (MMA) of less than one parts-per-million (ppm) enables the use of narrow (10 ppm) precursor ion maximum mass deviation (MMD) in the database search and therefore significantly reduces the search space *(2)*. This is of particular importance in the analysis of dynamic protein modifications, such as phosphorylation, where considerations of variable modifications lead to a “combinatorial explosion” during database search *(3)*. High MMA of fragment ions produced by commonly employed fragmentation techniques, such as collision-induced dissociation (CID) or electron-transfer dissociation (ETD), further increases the confidence of peptide

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_7

identification, phosphate group localization and reporter ion identification, and is therefore of great importance in reliable determination of previously unknown peptide sequences and their modifications (de novo sequencing). In addition, high MMA is connected with high-resolution MS which leads to a better peak definition and is therefore especially important for mass spectrometry-based quantitation. Finally, high resolution and MMA are crucial in the analysis of the whole proteins ("top-down" proteomics), where they are needed to resolve complex isotope clusters of multiply charged protein ions and to determine their accurate masses, as well as the masses of their modifications *(4)*.

While the advantages of the high measurement mass accuracy in proteomics are overwhelming, the limited sensitivity of the mass analyzers capable of performing sub-ppm MMA (typically Fourier transform ion cyclotron resonance analyzers, FT-ICR), and the underlying physical principles of mass measurement require relatively long (>1 s) scan times that are often not compatible with the liquid chromatography (LC) time scale. This is especially pronounced if both the precursor and fragment ions are measured at high resolution and it significantly decreases the sequencing capacity and reduces the number of identifications in a complex peptide mixture. As a consequence, less accurate but fast-scanning and more sensitive mass analyzers (e.g., ion traps) have been predominantly used in global studies of protein expression and phosphorylation *(5, 6)*.

A new generation of hybrid mass spectrometers has largely circumvented this issue by enabling the high sensitivity but low accuracy MS/MS events to be performed almost simultaneously with the high MMA of the precursor ions in the full scan. At the forefront of this development is the recently introduced LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) *(7, 8)* (**Fig. 1**). Along with routinely achieved low-ppm MMA in the

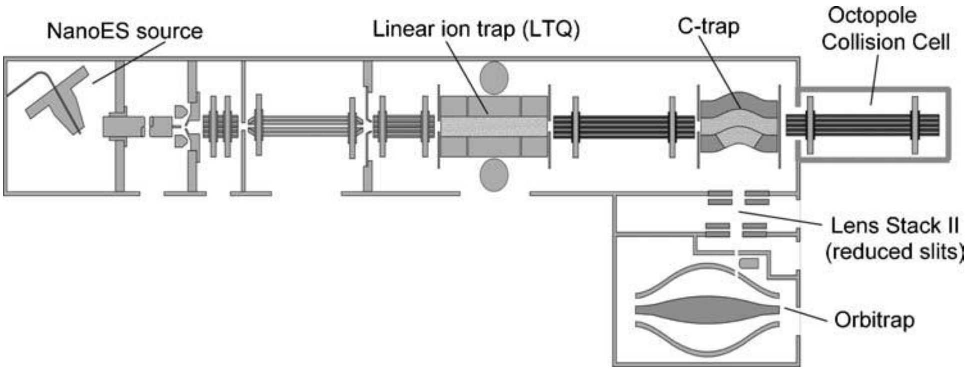


Fig. 1. Schematic of the LTQ-Orbitrap hybrid mass spectrometer. The instrument consists of a low-resolution but fast-scanning and sensitive linear ion trap (LTQ) mass analyzer connected via an intermediate storage device (C-trap) to a high-resolution but less sensitive orbitrap mass analyzer.

orbitrap mass analyzer *(9)*, especially important for the analysis of protein phosphorylation are the versatile fragmentation techniques available on this instrument. Ions can be fragmented by (1) low energy (~35 eV) CID in combination with multistage activation in the LTQ *(10)*; (2) higher energy (~60 eV) CID (HCD) in the C-trap or in an external octopole collision cell *(11)*; and (3) electron transfer dissociation (ETD), which will be available in the LTQ mass analyzer shortly *(12)*. Note that no controlled fragmentation is possible in the commercially available orbitrap analyzer itself; however, fragment ions produced in all instrument modules (LTQ, C-trap, octopole collision cell), can be efficiently transferred and measured at high MMA in the orbitrap.

Protein phosphorylation is often a low abundance and substoichiometric modification. Therefore, an enrichment of phosphopeptides or phosphoproteins from a complex mixture of their unmodified and usually highly abundant counterparts is at the core of every global mass spectrometry-based phosphoproteomics experiment *(13)*. This enrichment can be achieved by several different workflows that may include immunoprecipitation with antibodies targeting the phosphate group, especially on tyrosine *(14)*, and/or several stages of chromatography (e.g., strong cation exchange (SCX) or hydrophilic interaction chromatography) in combination with Fe- or Ga-based immobilized metal affinity (IMAC) *(15)* or TiO_2 chromatography *(16)*.

Here we present a generic phosphoproteomics workflow based on a two-stage enrichment of phosphopeptides from digests of whole cell/tissue lysates. Phosphopeptides are enriched using strong cation exchange and TiO_2 chromatography, and their analysis is performed in the LTQ-Orbitrap mass spectrometer. This strategy was successfully used in the previously reported global qualitative and quantitative studies of eukaryotic *(17)* and prokaryotic *(18)* protein phosphorylation.

2. Materials

2.1. In-Solution Protein Digestion

1. Denaturation buffer: 6 M urea (Sigma, Cat. No. U5128), 2 M thiourea (Invitrogen, Cat. No. Z10002), 1% *n*-octylglucoside (w/v) (Roche, Cat. No. 11359088001) in 10 mM HEPES buffer (Sigma, Cat. No. H3375), pH 8.0 (*see* **Note 1**).
2. Reduction buffer: 1 M dithiothreitol (Sigma, Cat. No. D0632) in 50 mM ammonium bicarbonate (Sigma, Cat. No. A6141).
3. Alkylation buffer: 550 mM iodoacetamide (Sigma, Cat. No. I6125) in 50 mM ammonium bicarbonate.

4. Lysyl Endopeptidase LysC (Waco, Cat. No. 129-02541).
5. Trypsin, sequencing grade, modified (Promega, Cat. No. V5113).

2.2. Strong Cation Exchange Chromatography

1. SCX solvent "A": 5 mM potassium dihydrogen phosphate (Merck, Cat. No. 1.04873.1000), 30% acetonitrile (Merck, Cat. No. 1.00030.2500); acidify with trifluoroacetic acid (Merck, Cat. No. 8.08260.0100) to pH 2.7.
2. SCX solvent "B": 5 mM potassium dihydrogen phosphate, 30% acetonitrile, 350 mM potassium chloride (Merck, Cat. No. 1.04936.0500), acidify with trifluoroacetic acid to pH 2.7.

2.3. Titanium Oxide Chromatography

1. Loading solution: 30 mg/mL 2,5 dihydrobenzoic acid (Fluka, Cat. No. 85707), 80% acetonitrile in water.
2. Washing solution I: 30% acetonitrile/3% trifluoroacetic acid.
3. Washing solution II: 80% acetonitrile/0.1% trifluoroacetic acid.
4. Elution solution I: 1% of NH_4OH (aq, 25% NH_3, Fluka, Cat. No. 09860) in 20% acetonitrile (pH > 10.5).
5. Elution solution II: 1% of NH_4OH (aq, 25% NH_3) in 40% acetonitrile (pH > 10.5).
6. Titansphere TiO_2 beads (GL Sciences, 10 µm, Cat. No. 5020-75010).

2.4. Liquid Chromatography-Mass Spectrometry

1. HPLC solvent "A": 0.5% acetic acid (Fluka, Cat. No. 45731) in water.
2. HPLC solvent "B": 0.5% acetic acid, 80% acetonitrile in water.
3. HPLC loading solvent: 1% trifluoroacetic acid, 2% acetonitrile in water.
4. Stage tips *(19, 20)*: Empore C8 Disk (Varian, Cat. No. 12145002).
5. Reversed phase material for nano-HPLC column *(21)*: Reprosil-Pur C18-AQ, 3 µm (Dr. Maisch, Cat. No. r13.aq).

3. Methods

3.1. In-Solution Protein Digestion

1. Dissolve the protein sample in denaturation buffer (*see* Note 2).
2. Add reduction buffer to the sample to a final concentration of 1 mM DTT; incubate for 1 h at room temperature (RT) (*see* Note 3).
3. Add alkylation buffer to the sample to a final concentration of 5.5 mM IAA; incubate for 1 h at RT in the dark.

4. Check pH (should be 8.0); adjust if necessary (*see* **Note 4**).
5. Add 1 µg of lysyl endopeptidase LysC per 100 µg protein and incubate for 3 h at room temperature.
6. Dilute sample with 4 volumes of water (*see* **Note 5**).
7. Check pH (should be 8.0); adjust if necessary.
8. Add 1 µg trypsin per100 µg sample protein and incubate overnight at RT.

3.2. SCX Chromatography

1. Acidify the protein digest to pH 2.7 with trifluoroacetic acid and centrifuge the sample to remove any precipitate that may form.
2. Load the protein digest onto an equilibrated SCX column (GE Healthcare Resource S, 1 mL) in SCX solvent "A" at a flow rate of 1 mL/min (*see* **Note 6**); collect the flow-through (*see* **Note 7**).
3. Elute the bound peptides with a linear gradient of 0–30% of SCX solvent "B" in 30 min at a flow rate of 1 mL/min (*see* **Note 8**); collect 2 mL fractions. A typical SCX chromatogram is shown in **Fig. 2**.
4. Wash the SCX column with five column volumes of 100% SCX solvent "B"

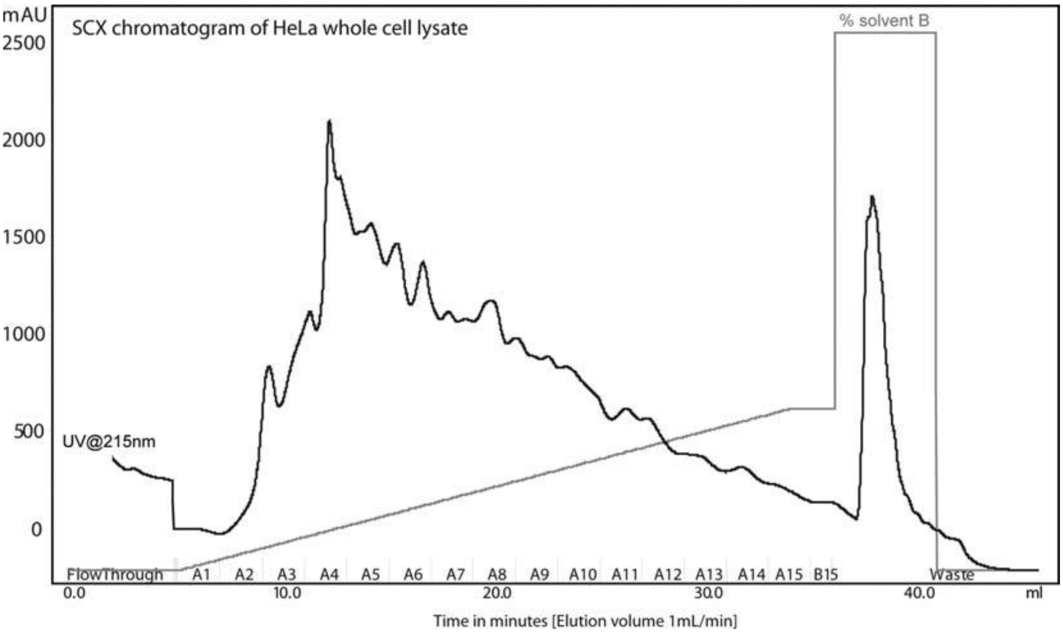


Fig. 2. Typical SCX chromatogram resulting from separation of 2 mg tryptic digest (in this case HeLa cell lysate) under the conditions described here. The highest enrichment of phosphopeptides is in the flow-through (not shown) and early fractions (A3–A8).

3.3. TiO_2 Chromatography

1. Each SCX fraction, including the flow-through, is incubated with the TiO_2 beads separately, and is referred to as "sample" in the text.
2. Weigh the TiO_2 beads into a dedicated tube; for each sample to be analyzed weigh 5 mg of beads (e.g., weigh a total of 50 mg if you want to process 10 SCX fractions).
3. Add the loading solution to the sample (1:6) (*see* **Note 9**).
4. Add 100 μL of loading solution to the TiO_2 beads; shake for 10 min at RT.
5. Add aliquots containing 5 mg of the TiO_2 beads slurry into each sample.
6. Incubate for at least 30 min (end-over-end rotation) at RT; centrifuge and discard supernatant.
7. Wash with 1.5 mL of washing solution I; shake vigorously; centrifuge and discard supernatant.
8. Wash with 1.5 mL of washing solution II; shake vigorously; centrifuge and discard supernatant.
9. Prepare one C8 microcolumn (stage-tip) for each sample by placing a ~1 mm^2 piece of Empore C8 material into a 200 μL (yellow) pipette tip, as described previously *(19, 20)* (*see* **Note 10**).
10. Transfer the TiO_2 beads into dedicated C8 microcolumns.
11. Elute each sample 2× with 100 μL of elution solution I; collect the eluates from each sample separately (*see* **Note 11**).
12. Elute each sample 1× with 100 μL of elution solution II (check pH of the last eluate – it should be >10; if not, elute with additional 100 μL of elution solution II).
13. Dry the eluates to ~5 μL (*see* **Note 12**).
14. Add HPLC loading solvent to 10 μL (final concentration of 1% acetonitrile and ~0.5% trifluoroacetic acid); the samples are now ready for nano-LC-MS.

3.4. Lc-ms

3.4.1. Liquid Chromatography

1. The liquid chromatography (LC) part of the analytical LC-MS system described here consists of an Agilent 1200 Series nanoflow LC system (Agilent Technologies) comprising a solvent degasser, a nanoflow pump, and a thermostated micro-autosampler with an 8-μL injection loop.
2. Pack an analytical column in a 20-cm fused silica emitter (Proxeon Biosystems, 75-μm inner diameter with a 5-μm laser-pulled tip), with a methanol slurry of reverse-phase C18 resin at a constant helium pressure (50 bar) using a bomb-loader device (Proxeon Biosystems), as described previously *(21)* (*see* **Note 13**).

3. Connect the packed emitter (C18 RP HPLC column) directly to the outlet of the 6-port valve of the HPLC autosampler through a 20-cm long, 25-μm inner diameter fused silica transfer line (Composite Metals) and a micro Tee-connector (Upchurch) ("liquid junction" connection) (*see* **Note 14**).
4. Load 6 μL of the tryptic phosphopeptide mixture using the HPLC autosampler onto the packed emitter at a flow of 500 nL/min (typical back-pressure of 130–200 bar) for 20 min using 2% of HPLC solvent B.
5. After loading, by-pass the sample loop, reduce the flow-rate to 200 nL/min and increase the HPLC solvent B content to 10%.
6. Separate and elute the bound peptides with a 90 min linear gradient from 10 to 30% of HPLC solvent B. Wash-out hydrophobic peptides by linearly increasing the HPLC solvent B content to 80% over 15 min.

3.4.2. Mass Spectrometry

1. All mass spectrometric experiments discussed here are performed on an LTQ orbitrap "XL" or "Classic" mass spectrometer (Thermo Fisher Scientific) connected to an Agilent 1200 Series nanoflow LC system (Agilent Technologies) via a nanoelectrospray LC-MS interface (Proxeon Biosystems) (*see* **Note 15**).
2. Operate the mass spectrometer in the data-dependent mode to automatically switch between MS and MS/MS using the Tune and Xcalibur 2.4 software package.
3. Use the following settings in the "Tune" acquisition software:
   (a) FT full scan: accumulation target value 1E6; max. fill time 1,000 ms
   (b) FT MS^n: accumulation target value 5E4; max. fill time 500 ms
   (c) IT MS^n: accumulation target value 5E3; max. fill time 150 ms
4. For accurate mass measurements enable the "Lock mass" option in both MS and MS/MS mode in the Xcalibur software *(9)*. Use the background polydimethylcyclosiloxane (PCM) ions generated from ambient air (e.g., $m/z = 445.120025$) for internal recalibration in real time. For single SIM scan injections of the lock mass ion into the C-trap, set the lock mass "ion gain" at 10% of the target value of the full mass spectrum. If the fragment ion measurements are performed in the orbitrap, use the PCM ion at $m/z = 429.088735$ (PCM with neutral methane loss).
5. In the "Xcalibur Instrument Setup" create a data-dependent acquisition method in which full scan MS spectra, typically in the m/z range from 300 to 1,800, are acquired by the orbitrap detector with resolution $R = 60{,}000$ (*see* **Note 16**).

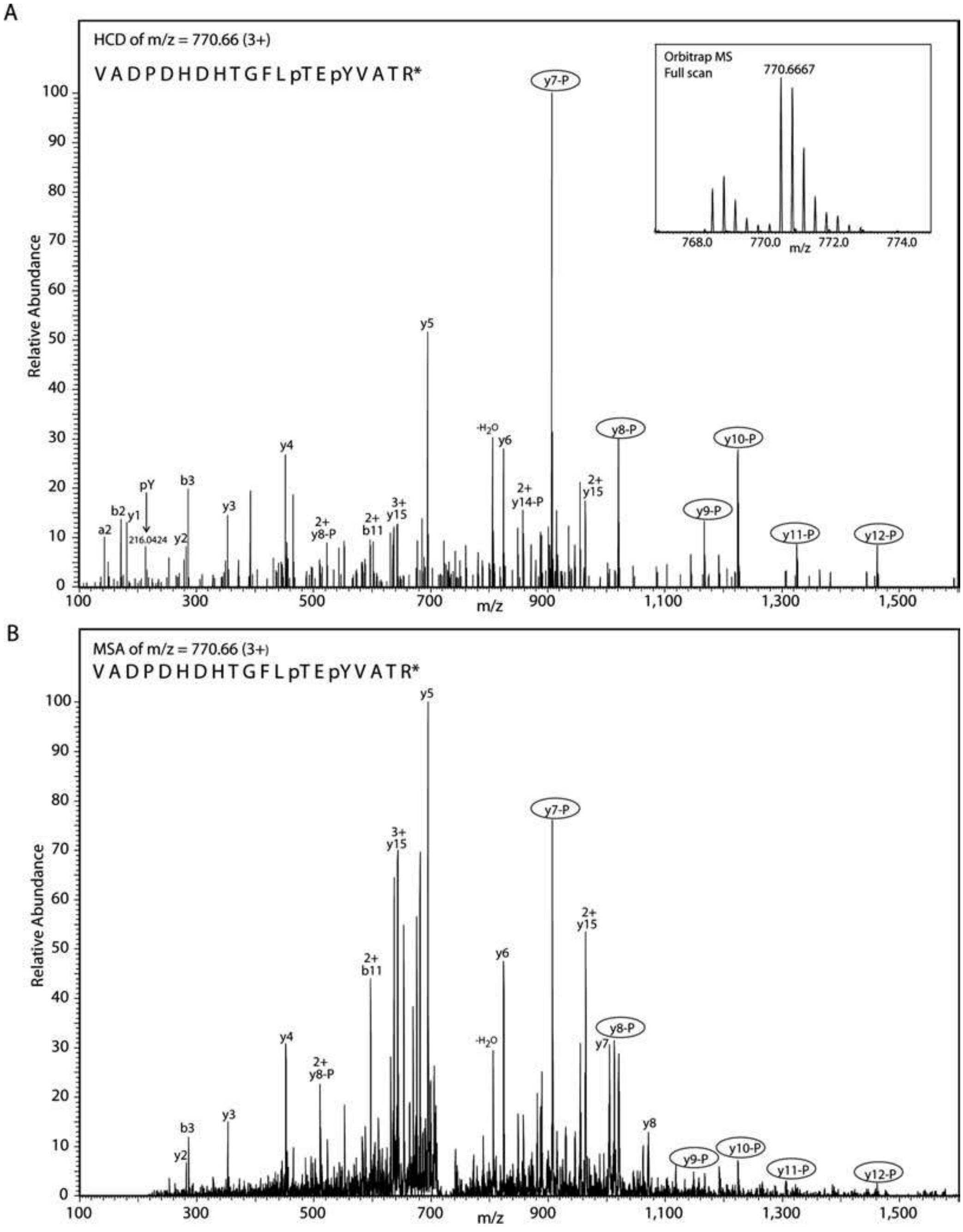
(a) For high-accuracy and full mass range measurements of fragment ions, set the data-dependent MS^2 of the three most intense multiply charged ions to be measured in the FT analyzer (orbitrap) at the resolution of $R = 15{,}000$ and enable the HCD option (Higher-energy C-trap dissociation) (*see* **Note 17**). Set the first mass in mass range to $m/z = 80$. This will allow for low mass reporter ions such as immonium ions to be identified *(11)*. Typical features of a HCD MS^2 spectrum are shown in **Fig. 3a**.

(b) For fast-scanning and high-sensitivity but low resolution measurements of fragment ions, set the data-dependent MS^2 of the five most intense multiply charged ions to be measured in the linear ion trap. Enable the preview mode for FTMS master scans to perform data-dependent MS^2 *in-parallel* with the full scan in the orbitrap. Set the fragmentation mode to CID (*see* **Note 18**) and enable the multi-stage activation (MSA) fragmentation option by which the neutral loss species at 97.97, 48.99, or 32.66 m/z below the precursor ion will be successively activated for 30 ms each *(10)* (*see* **Note 19**). Typical features of a MS^2 spectrum acquired in the linear ion trap using multistage activation are shown in **Fig. 3b**.

6. Standard acquisition method settings:
   (a) Electrospray voltage, 2.4 kV (*see* **Note 20**).
   (b) No sheath and auxiliary gas flow.
   (c) Ion transfer (heated) capillary temperature, 150°C.
   (d) Collision gas pressure, 1.3 mtorr.
   (e) Dynamic exclusion of up to 500 precursor ions for 60 s upon MS/MS; exclusion mass width of 10 ppm.
   (f) Normalized collision energy using wide-band activation mode; 35% for both CID and HCD.
   (g) Ion selection thresholds: 1,000 counts for CID and 10,000 counts for HCD.
   (h) Activation $q = 0.25$; Activation time = 30 ms.

→

Fig. 3. (**a**) HCD spectrum of a doubly phosphorylated Erk2 peptide VADPDHDTGFLpTEpYVATR. HCD is performed in the C-trap or in the external octopole collision cell and results in a high resolution, full range spectrum where the phosphotyrosine immonium ion (pY) at $m/z = 216.0424$ can be easily detected. Note that the majority of the fragment ions containing the phosphorylated Thr13 residue lose the phosphate group (–P), a typical feature of Ser/Thr-phosphorylated peptides. Average absolute mass deviation of the fragment ions is <2 ppm. (**b**) CID spectrum with multistage activation of the same Erk2 peptide. CID is performed in the LTQ and fragment ions can be detected either in the LTQ (shown here) or in the orbitrap mass analyzer. Due to higher sensitivity, more fragment ions are detected in the LTQ, albeit at lower resolution and MMA (~300 ppm). Due to the multistage activation, most fragment ions containing the phosphorylated Thr13 residue are present in phosphorylated and unphosphorylated forms. Note that the ions below $m/z = 250$ (or ~30% of the precursor ion mass) cannot be detected; hence the phosphotyrosine immonium ion is missing.



4. Notes

1. All solvents in this protocol should be prepared with high purity deionized Milli-Q water of the resistivity 18.2 MΩ cm (Millipore Q-Gard 2 cartridge). This deionized water is

referred to as "water" in the text. Denaturation, reduction and alkylation buffers can be frozen as stock solutions at –20°C.

2. The entire protocol is optimized for protein amounts from 2 to 20 mg. During cell lysis and protein extraction, keep the salt content at a minimum since it may interfere with the SCX chromatography. If needed, precipitate proteins with acetone or chloroform/methanol prior to analysis.
3. Do not heat-up the samples during protein solubilization and digestion; high concentration of urea will lead to carbamylation of free amino groups.
4. If necessary, adjust the pH with a very low volume of 1 M Tris–HCl, pH 8.0; note that the total salt concentration should not exceed 10 mM since it may interfere with the SCX chromatography.
5. To keep the salt concentration low, use pure water rather than ammonium bicarbonate buffer to dilute the sample prior to trypsin incubation.
6. During sample loading onto the SCX column it is important to monitor conductivity; under the conditions employed, conductivity higher than 4 mS/cm will lead to decreased binding of peptides – in that case the flow-through should be diluted with water and re-loaded onto the column.
7. Collection and subsequent separate analysis of the flow-through is extremely important since multiply phosphorylated peptides will not bind to the SCX column; this fraction is usually the richest in the number of identified phosphorylation sites, especially in the analysis of eukaryotic phosphorylation.
8. It is recommended that the conductivity be monitored throughout the SCX run, as it is the best measure for the gradient stability; under conditions employed, at 30% of the SCX solvent "B" the conductivity typically reaches 13 mS/cm.
9. This step is optional and should be used to increase the binding specificity in the samples with very little phosphorylation (e.g., prokaryotic cell lysates); otherwise it should be avoided because it increases the probability of DHB contamination during LC-MS. Alternatively, it is possible to use lactic acid as a competitive binder instead of DHB *(22)*.
10. C8 microcolumns are also commercially available (Proxeon Biosystems, Order Code SP121).
11. To quickly neutralize the high pH upon elution, elute the samples into a small volume (~20 μL) of the HPLC loading solvent.
12. Do not heat the samples during vacuum concentration and make sure they do not run dry.

13. The advantage of the resin used is the fact that it is active at low organic buffer content (e.g., less than 2% acetonitrile). Thereby lower amount of hydrophobic ion-pairing reagents such as heptafluorobutyric acid (HFBA), which interfere with the electrospray ionization, can be used in the HPLC solvents.
14. Note that there is no pre-column or split in this LC-MS setup.
15. Control the timing between the MS and the LC system with a standard double contact closure cable.
16. Resolution is defined at $m/z = 400$.
17. This gives a total scan cycle time (full scan + 3 MS^2 events) of up to 5 s.
18. This gives a total scan cycle time (full scan + 5 MS^2 events) of up to 3 s.
19. Multistage activation produces information-rich spectra, where many of the fragment ions show pronounced neutral loss of phosphoric acid (e.g., –97.97, –48.99, or –32.66 for singly, doubly or triply charged fragment ions, respectively). This information is very useful in validation of the phosphopeptide spectra.
20. Source settings have to be optimized for the emitter and nano-LC-MS setup.

Acknowledgements

The authors thank Prof. Matthias Mann and Dr. Chunaram Choudhary for critical reading of the manuscript and members of the Department for Proteomics and Signal Transduction at the MPI for Biochemistry for numerous discussions that led to the optimized protocols for phosphopeptide enrichment and MS analysis presented here.

References

1. Aebersold, R. & Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207.
2 Zubarev, R. & Mann, M. (2007) On the proper use of mass accuracy in proteomics. *Mol Cell Proteomics* 6, 377-381.
3. Haas, W., Faherty, B. K., Gerber, S. A., Elias, J. E., Beausoleil, S. A., Bakalarski, C. E., Li, X., Villen, J. & Gygi, S. P. (2006) Optimization and use of peptide mass measurement accuracy in shotgun proteomics. *Mol Cell Proteomics* 5, 1326-1337.
4. Macek, B., Waanders, L. F., Olsen, J. V. & Mann, M. (2006) Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. *Mol Cell Proteomics* 5, 949-958.
5. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F. & White, F. M. (2002)

Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 20, 301-305.

6. Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villen, J., Li, J., Cohn, M. A., Cantley, L. C. & Gygi, S. P. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A* 101, 12130-12135.
7. Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K. & Horning, S. (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal Chem* 78, 2113-2120.
8. Scigelova, M. & Makarov, A. (2006) Orbitrap mass analyzer - overview and applications in proteomics. *Proteomics* 6 (Suppl 2), 16-21.
9. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S. & Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics* 4, 2010-2021.
10. Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F. & Coon, J. J. (2004) A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal Chem* 76, 3590-3598.
11. Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S. & Mann, M. (2007) Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 4, 709-712.
12. McAlister, G. C., Phanstiel, D., Good, D. M., Berggren, W. T. & Coon, J. J. (2007) Implementation of electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. *Anal Chem* 79, 3525-3534.
13. Schmelzle, K. & White, F. M. (2006) Phosphoproteomic approaches to elucidate cellular signaling networks. *Curr Opin Biotechnol* 17, 406-414.
14. Blagoev, B., Ong, S. E., Kratchmarova, I. & Mann, M. (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat Biotechnol* 22, 1139-1145.
15. Posewitz, M. C. & Tempst, P. (1999) Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal Chem* 71, 2883-2892.
16. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P. & Jorgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* 4, 873-886.
17. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P. & Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635-648.
18. Macek, B., Mijakovic, I., Olsen, J. V., Gnad, F., Kumar, C., Jensen, P. R. & Mann, M. (2007) The serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis. *Mol Cell Proteomics* 6, 697-707.
19. Rappsilber, J., Ishihama, Y. & Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* 75, 663-670.
20. Rappsilber, J., Mann, M. & Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2, 1896-1906.
21. Ishihama, Y., Rappsilber, J., Andersen, J. S. & Mann, M. (2002) Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A* 979, 233-239.
22. Sugiyama, N., Masuda, T., Shinoda, K., Nakamura, A., Tomita, M. & Ishihama, Y. (2007) Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics* 6, 1103-1109.

Chapter 8

Manual Validation of Peptide Sequence and Sites of Tyrosine Phosphorylation from MS/MS Spectra

Amy M. Nichols and Forest M. White

Summary

Mass spectrometry-based analysis of protein phosphorylation has become increasingly powerful over the past decade and has been applied to many different biological systems. One of the most significant concerns facing the phosphoproteomics community and the proteomics field as a whole is the quality and accuracy of the data generated in these large scale efforts. For protein identification in a given sample, the solution has been to require multiple peptides per protein, eliminating "one-hit-wonders" (proteins identified on the basis of a single peptide assignment) which may increase false positives in the data set. Unfortunately, most of the phosphoproteomics data fall into the latter category, as each phosphorylation site will most likely be represented by a single tryptic peptide. Here we provide a detailed protocol describing our manual validation efforts to assure accurate peptide and phosphorylation site assignment for individual MS/MS spectra. In this procedure we use a combination of tools to assign b-, y-, neutral loss, and internal fragment ions, with the goal of assigning all significant ions in the MS/MS spectrum. Confident peptide and phosphorylation site assignment requires good coverage of the peptide with minimal unassigned fragment ions. Using this approach it is possible to maximize the quality of the phosphoproteomics data while minimizing database contamination associated with false positive identifications.

Key words: MS/MS, Mass Spectrometry, Spectra validation, Phosphorylation, Post-translational modification.

1. Introduction

Mass spectrometry has emerged as one of the primary techniques for large scale identification of proteins and post translational modifications such as phosphorylation. Analysis of proteome-wide protein phosphorylation by mass spectrometry can have a substantial impact in advancing our understanding of cell signaling networks if phosphorylation sites are correctly identified while minimizing false positive results. Phosphorylation site identification

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_8

is complicated because phosphorylation occurs only on specific residues in a protein sequence and typically, only one tryptic peptide will contain the phosphorylation site of interest. Therefore, the challenges in using mass spectrometry for phosphoproteomics are to extract and analyze the few peptides of interest and to definitively validate the sequence of the peptide and the location of the phosphorylation site from a single MS/MS spectrum. To accomplish these goals, phosphopeptide enrichment can be achieved with selective sample purification *(1)* and stringent manual data validation can be used to confidently identify peptide sequence and phosphorylation sites. The goal of manual sequence validation is to assign all peaks in the MS/MS spectra above m/z = 300 as fragments of the peptide sequence. Even when all the MS/MS peaks are assigned, if the fragments are not distributed over the full sequence, then the peptide and/or phosphorylation assignments can remain ambiguous.

2. Materials

2.1. Phosphopeptide Sample Preparation

1. Protocols for sample preparation have been documented in peer reviewed papers from our group *(1–6)* and in addition, a detailed version of the protocol is documented in *Methods in Molecular Biology (7)*.

2.2. Spectra Acquisition

1. Protocols for spectral acquisition have been documented in peer reviewed papers from our group *(1–6)* and in addition, a detailed version of the protocol is documented in *Methods in Molecular Biology (7)*.

2.3. Spectra Validation

1. Searching access on a MASCOT database from Matrix Science Ltd (*see* **Note 1**).
2. Analyst QS Software 1.0 from Applied Biosystems | MDS SCIEX.
3. A list of the monoisotopic masses of each amino acid, **Table 1**.
4. A list of common immonium ions, **Table 2**.

3. Methods

3.1. Phosphopeptide Sample Preparation

As mentioned above, general protocols for sample preparation have been published *(1–6)* and a detailed protocol is documented in *Methods in Molecular Biology (7)*. A brief synopsis of the protocol follows to help the reader understand the type of data being validated.

Table 1
Monoisotopic masses of the amino acids can be combined along with the charge state to calculate the theoretical *m/z* peak for a specific peptide or fragment. Note that leucine and isoleucine are isobaric isomers (same mass, different structures) and cannot be distinguished by mass spectrometry. In addition, the mass differential between amino acids such as glutamine (128.05858) and lysine (128.09496) may be below the resolution of some mass spectrometry instruments. Chemical modification during sample preparation may alter the monoisotopic masses, such as reduction and alkylation of cysteine residues

Monoisotopic masses		
Amino acid	**Single letter code**	**Monoisotopic mass**
Alanine	A	71.03711
Arginine	R	156.10111
Asparagine	N	114.04293
Aspartate	D	115.02694
Cysteine	C	103.00919
Glutamate	E	129.04259
Glutamine	Q	128.05858
Glycine	G	57.02146
Histidine	H	137.05891
Isoleucine	I	113.08406
Leucine	L	113.08406
Lysine	K	128.09496
Methionine	M	131.04049
Phenylalanine	F	147.06841
Proline	P	97.05276
Serine	S	87.03203
Threonine	T	101.04768
Tryptophan	W	186.07931
Tyrosine	Y	163.06333
Valine	V	99.06841

A cell line of interest is grown to confluence on 10-cm culture plates. The cells are cultured in serum-free culture media 12 h prior to cytokine stimulation. Following the chosen stimulation

Table 2
Immonium ions can be generated during fragmentation of a peptide and result from the loss of CO from the C terminus and the addition of hydrogen to the N terminus. Although immonium ions will indicate the presence of selected amino acids in a peptide, the absence of these ions is insufficient to exclude an amino acid

Common immonium ions

Amino acid	Single letter code	Immonium ion mass
Alanine	A	44.0495
Arginine	R	129.1135
Asparagine	N	87.0553
Aspartate	D	88.0393
Cysteine	C	76.0215
Glutamate	E	102.055
Glutamine	Q	101.0709
Glycine	G	30.0338
Histidine	H	110.0713
Isoleucine	I	86.0964
Leucine	L	86.0964
Lysine	K	101.1073
Methionine	M	104.0528
Phenylalanine	F	120.0808
Proline	P	70.0651
Serine	S	60.0444
Threonine	T	74.06
Tryptophan	W	159.0917
Tyrosine	Y	136.0757
Valine	V	72.0808

conditions, cells are lysed on ice with a solution comprised of 8 M urea and 1 mM sodium orthovanadate to inhibit phosphatase activity. Following reduction (10 mM dithiothreitol at 56°C for 1 h) and alkylation (55 mM iodoacetamide in the dark at room temperature for 1 h), proteins are enzymatically digested for at least 14 h in 100 mM ammonium acetate with 40 μg of trypsin (Promega). Tryptic peptides are desalted in a C18 Sep-Pak Plus cartridge (Waters) and eluted with 25% acetonitrile, 0.1% acetic acid; the eluted fraction is then lyophilized and stored at –80°C. For immunoprecipitation, lyophilized samples were dissolved in 400 µL of 100 mM Tris, 100 mM NaCl, 1% NP-40 at pH 7.4 (note: following solubilization of the sample, it may be necessary to re-adjust the pH to 7.4 for optimal immunoprecipitation). Samples are incubated overnight with a monoclonal anti-phosphotyrosine antibody (e.g., PT66, Sigma, Product #P5872) to immunoprecipitate the tyrosine phosphorylated peptides. These peptides are eluted with 50 µL of 100 mM glycine pH 2.5 for 30 min at room temperature (*see* **Note 2**).

3.2. Spectral Acquisition

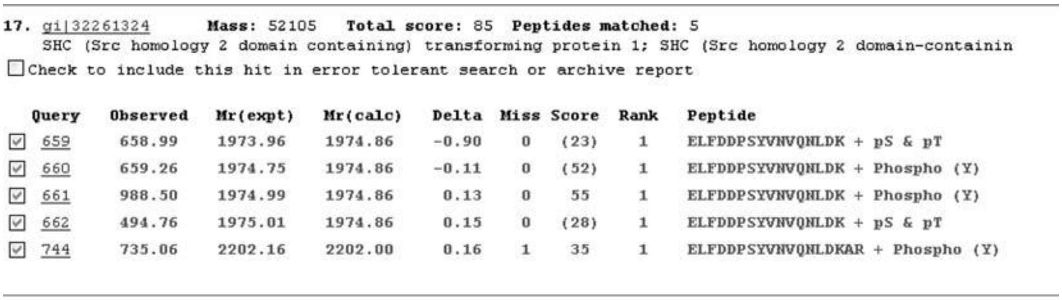
Peptides are loaded into an immobilized metal affinity column and nonspecifically bound peptides are removed with 25% acetonitrile, 1% acetic acid, 100 mM NaCl. Retained peptides are eluted into a reverse-phase C18 pre-column with 50 µL of 250 mM Na_2HPO_4, pH 8.0. The C18 pre-column is attached to a reverse phase analytical column and a high pressure liquid chromatography system. The peptides are slowly eluted with a 140 min gradient as described previously *(1)* and electrosprayed into a quadrupole time-of-flight mass spectrometer (QSTAR XL Pro, Applied Biosystems) for MS and MS/MS spectra acquisition. All the spectra from the sample are saved to a *.wiff file.

3.3. Initial MS/MS Results

1. Convert the data *.wiff file to a MASCOT readable *.msm or *.mgf file using MASCOT.dll in Analyst.
2. Search the data on MASCOT with the MS/MS Ion search form. Select settings relevant to the experimental sample.
3. The MASCOT search algorithm matches the MS/MS data with peptides from specific proteins (*see* **Note 3**) and provides a confidence score for each assignment. Although a high confidence score is generally indicative of a good assignment, manual validation is necessary to verify the peptide and phosphorylation site assignment.
4. To manually validate a peptide, find the MASCOT assignment with the highest confidence score and click on the query link to open the peptide view window, *see* **Fig. 1**.
5. From the peptide view window, record the mass-to-charge ratio (*m/z*) of the precursor ion and its elution time frame displayed at the top of the window, *see* **Fig. 2**.

3.4. Identifying and Annotating MS/MS Spectra

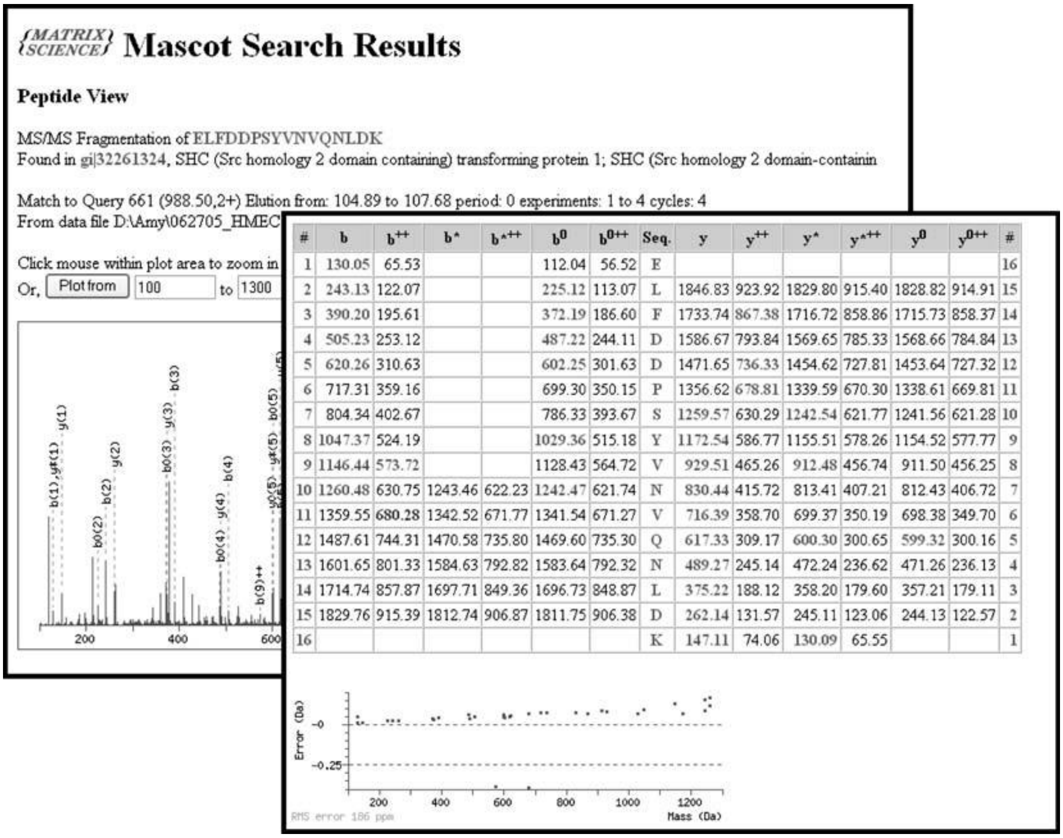
1. Open the data *.wiff file in Analyst 1.0.
2. Double click anywhere on the total ion chromatogram (TIC) to display the MS and MS/MS spectra taken at that time.
3. Use the elution time frame and *m/z* for the precursor ion of interest to find the correct full scan mass spectrum (MS scan) and the corresponding MS/MS spectrum from which the assignment was made.



```
17. gi|32261324      Mass: 52105   Total score: 85   Peptides matched: 5
    SHC (Src homology 2 domain containing) transforming protein 1; SHC (Src homology 2 domain-containin
☐Check to include this hit in error tolerant search or archive report

    Query   Observed   Mr(expt)   Mr(calc)    Delta  Miss  Score  Rank   Peptide
☑   659      658.99    1973.96    1974.86    -0.90    0    (23)    1     ELFDDPSYVNVQNLDK + pS & pT
☑   660      659.26    1974.75    1974.86    -0.11    0    (52)    1     ELFDDPSYVNVQNLDK + Phospho (Y)
☑   661      988.50    1974.99    1974.86     0.13    0     55     1     ELFDDPSYVNVQNLDK + Phospho (Y)
☑   662      494.76    1975.01    1974.86     0.15    0    (28)    1     ELFDDPSYVNVQNLDK + pS & pT
☑   744      735.06    2202.16    2202.00     0.16    1     35     1     ELFDDPSYVNVQNLDKAR + Phospho (Y)
```

Fig. 1. MASCOT search result displayed as a function of proteins and the associated peptides found throughout the experimental sample. Each entry following a protein is a summation of the MS/MS spectra of a given precursor ion over a limited elution time; repeated peptides may be listed if the precursor's elution time is sufficiently long. Each query in the list will be given a score indicating the probability of correctly assigning the MS/MS spectra to a given peptide. In this example, five sets of MS/MS spectra were found to correlate to the same peptide (and one missed cleavage) in multiple charge states (+2, +3, +4). The peptides appear to vary in their phosphorylation, but this variation is actually due to incorrect assignment of the site of phosphorylation; upon manual validation these assignments will be clarified.



{MATRIX SCIENCE} Mascot Search Results

Peptide View

MS/MS Fragmentation of ELFDDPSYVNVQNLDK
Found in gi|32261324, SHC (Src homology 2 domain containing) transforming protein 1; SHC (Src homology 2 domain-containin

Match to Query 661 (988.50,2+) Elution from: 104.89 to 107.68 period: 0 experiments: 1 to 4 cycles: 4
From data file D:\Amy\062705_HMEC

Click mouse within plot area to zoom in
Or, Plot from 100 to 1300

#	b	b⁺⁺	b*	b*⁺⁺	b⁰	b⁰⁺⁺	Seq.	y	y⁺⁺	y*	y*⁺⁺	y⁰	y⁰⁺⁺	#
1	130.05	65.53			112.04	56.52	E							16
2	243.13	122.07			225.12	113.07	L	1846.83	923.92	1829.80	915.40	1828.82	914.91	15
3	390.20	195.61			372.19	186.60	F	1733.74	867.38	1716.72	858.86	1715.73	858.37	14
4	505.23	253.12			487.22	244.11	D	1586.67	793.84	1569.65	785.33	1568.66	784.84	13
5	620.26	310.63			602.25	301.63	D	1471.65	736.33	1454.62	727.81	1453.64	727.32	12
6	717.31	359.16			699.30	350.15	P	1356.62	678.81	1339.59	670.30	1338.61	669.81	11
7	804.34	402.67			786.33	393.67	S	1259.57	630.29	1242.54	621.77	1241.56	621.28	10
8	1047.37	524.19			1029.36	515.18	Y	1172.54	586.77	1155.51	578.26	1154.52	577.77	9
9	1146.44	573.72			1128.43	564.72	V	929.51	465.26	912.48	456.74	911.50	456.25	8
10	1260.48	630.75	1243.46	622.23	1242.47	621.74	N	830.44	415.72	813.41	407.21	812.43	406.72	7
11	1359.55	680.28	1342.52	671.77	1341.54	671.27	V	716.39	358.70	699.37	350.19	698.38	349.70	6
12	1487.61	744.31	1470.58	735.80	1469.60	735.30	Q	617.33	309.17	600.30	300.65	599.32	300.16	5
13	1601.65	801.33	1584.63	792.82	1583.64	792.32	N	489.27	245.14	472.24	236.62	471.26	236.13	4
14	1714.74	857.87	1697.71	849.36	1696.73	848.87	L	375.22	188.12	358.20	179.60	357.21	179.11	3
15	1829.76	915.39	1812.74	906.87	1811.75	906.38	D	262.14	131.57	245.11	123.06	244.13	122.57	2
16							K	147.11	74.06	130.09	65.55			1

Fig. 2. Peptide view window displaying the totaled MS/MS spectra for a given peptide assignment in the MASCOT search results. MASCOT displays the b and y ions it is able to identify. In the example shown, the b and y ions overlap and appear to cover most of the sequence, indicating that this spectrum is appropriate for manual validation. This table of values can be used as an initial reference for manual validation and peak labeling, but the sequencing tool in Analyst 1.0 has an adaptable user interface. Note that the elution time is listed between 104.89 and 107.68; this time window can be used to find individual MS/MS spectra for the precursor ion in the Analyst 1.0 *.wiff file. The error plot shows a general trend away from zero as a function of *m/z*, an error trend commonly found in our data sets. The two points with substantial negative deviation are incorrectly assigned by MASCOT. (*See Color Plates*)

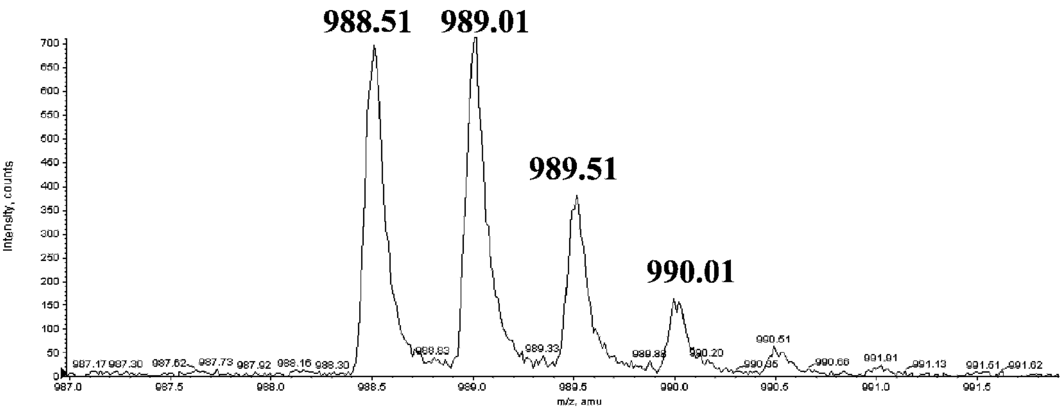


Fig. 3. Precursor ion isolation window in the full scan mass spectrum. For the purpose of this paper, a peptide from Src homology 2 domain-containing-transforming protein C1 (SHC) will be used to demonstrate the manual validation process (*see* **Table 3**). In this case, the precursor ion of interest is 988.51 (+2 charge state) and the area from −1 to 2.5 m/z units around the precursor is free of contamination and the noise around the peak is minimal (below 10% of the precursor peak).

4. The MS/MS spectrum may contain contaminating peptide fragments or species because any peptides or other species within the isolation window (typically −1 to +2.5 *m/z* of the precursor ion (*see* **Note 4**)) will be passed through the first quadrupole and fragmented in the collision cell. To assess the level of possible contamination in the MS/MS spectrum, zoom in on the precursor *m/z* within the MS scan and look for other ions near the precursor ion. A general threshold for contamination is 10% of the parent ion height excluding the peaks in the isotope envelope from the precursor ion, *see* **Fig. 3**.
5. If the full scan mass spectrum indicates potential contamination, return to the MASCOT results and find an alternate MS/MS spectrum for that peptide. Proceeding with a contaminated MS/MS spectrum may be frustrating and inconclusive because fragment peaks generated from the contaminating peak will not be assignable.
6. If the isolation window in the full scan mass spectrum is free of contaminating peaks, select the appropriate MS/MS spectrum and print a full page copy for annotation.
7. At the top of the MS/MS spectrum, write the peptide sequence in the single letter amino acid code assigned by MASCOT in the standard N- to C-terminal direction.
8. Next, label the *m/z* ratio and the charge state for all peaks above the background. The charge state can be found by zooming in on the peak in Analyst and evaluating the spacing between the isotope peaks. (Ex. A spacing of 1 *m/z* unit is a +1, 0.5 *m/z* unit is a +2, 0.33 *m/z* is a +3, etc.). Also, verify that the labeled *m/z* value printed by Analyst on the spectra corresponds to the monoisotopic peak (typically the first peak in the isotope envelope), **Fig. 4**.

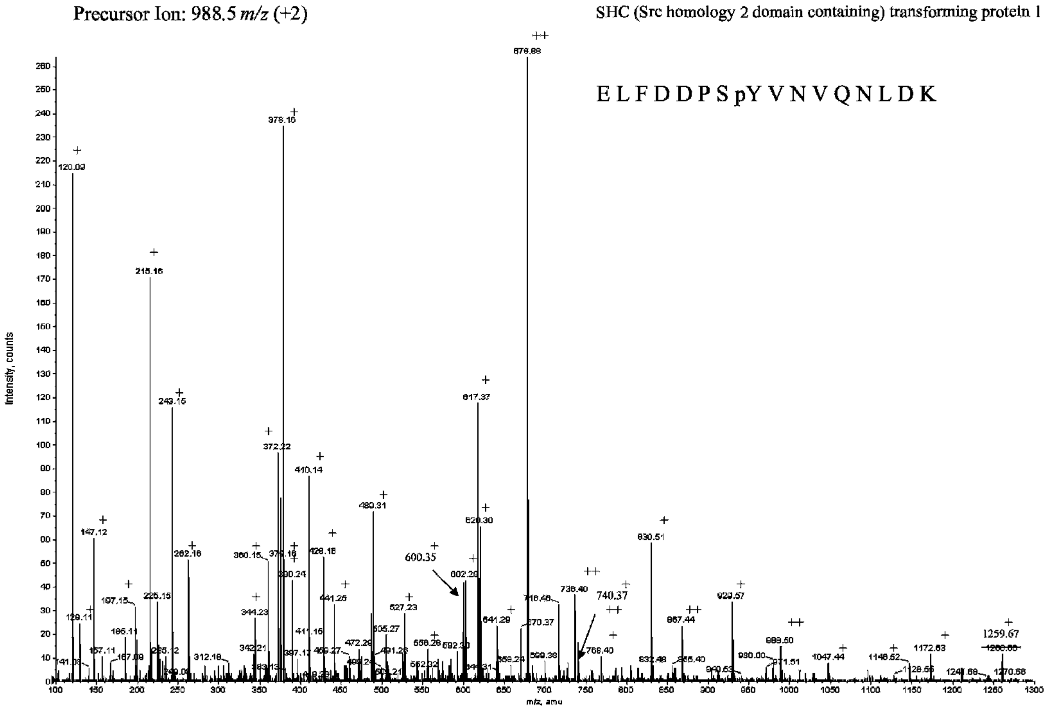
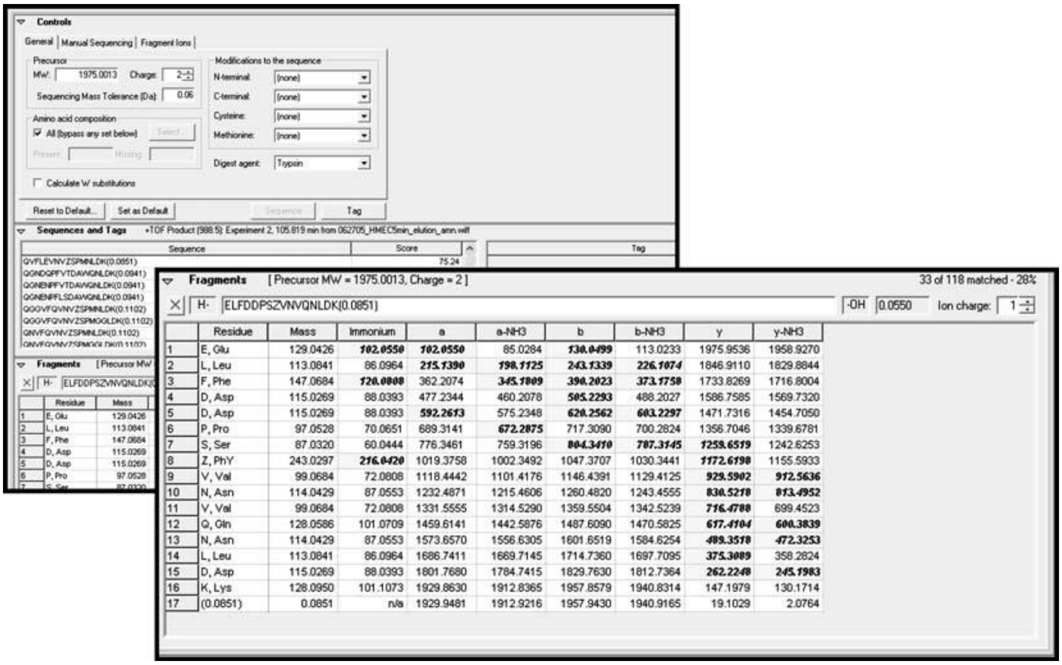


Fig. 4. Annotating charge state and *m/z* for fragment ions in the MS/MS spectrum. The basic annotation for an MS/MS scan includes the charge state of each peak, the protein it is assigned to by MASCOT, the peptide sequence matched by MASCOT and the parent ion *m/z* and charge state. The charge state for each peak in the MS/MS scan can be found by zooming in on the peak in Analyst and evaluating the spacing in the isotope envelope (+1 = 1 *m/z*, +2 = 0.5 *m/z*, +3 = 0.33 *m/z;* +4 = 0.25 *m/z*, etc.).

3.5. Peptide Sequencing

1. De novo peptide sequencing can be performed on each spectrum, but it can be time consuming even with experience. To facilitate fragment ion identification, select the MS/MS spectrum of interest in Analyst 1.0 and select sequence peptide from the BioExplore menu. A new window will appear with a variety of settings that can be specified for the experimental conditions, **Fig. 4**.
2. Once you have selected the settings, click on the sequence button. Analyst will propose a set of sequences with scores reflecting how well each sequence matches the spectrum (*see* **Note 5**).
3. If the sequence proposed by MASCOT is not listed in the Analyst window, type the sequence into the fragment window, **Fig. 5**. Make sure to follow the nomenclature in Analyst for modified residues, such as Z for phosphotyrosine (*see* **Note 6**).
4. Once the sequence is entered, the fragment ions identified by Analyst are highlighted in the table. First, use the table to identify the +1 charge state b and y ions, generated by


Controls
General | Manual Sequencing | Fragment Ions
Precursor
Modifications to the sequence
Amino acid composition
Select...
Present: Missing:
Reset to Default... Set as Default Sequence Tag
Sequences and Tags
Sequence Score Tag

MW: 1975.0013 Charge: 2
Sequencing Mass Tolerance (Da): 0.06
N-terminal: (none)
C-terminal: (none)
Cysteine: (none)
Methionine: (none)
Digest agent: Trypsin
☑ All (bypass any set below)
☐ Calculate W substitutions

+TOF Product (988.5): Experiment 2, 105.819 min from 062705_HMEC5min_elution_amn.wiff

Sequence	Score
QVFLEVNVZSPMNLDK(0.0851)	75.24
QGNDGPFVTDAWQNLDK(0.0941)	
QGNENPFVTDAWQNLDK(0.0941)	
QGNENPFLSDAWQNLDK(0.0941)	
QGGVFQVNVZSPMNLDK(0.1102)	
QGGVFQVNVZSPMGGLDK(0.1102)	
QNVFQVNVZSPMNLDK(0.1102)	

Fragments [Precursor MW = 1975.0013, Charge = 2] 33 of 118 matched - 28%

H- ELFDDPSZVNVQNLDK(0.0851) -OH 0.0550 Ion charge: 1

	Residue	Mass	Immonium	a	a-NH3	b	b-NH3	y	y-NH3
1	E, Glu	129.0426	**102.0550**	**102.0550**	85.0284	**130.0499**	113.0233	1975.9536	1958.9270
2	L, Leu	113.0841	86.0964	**215.1390**	**198.1125**	**243.1339**	**226.1074**	1846.9110	1829.8844
3	F, Phe	147.0684	**120.0808**	362.2074	**345.1809**	**390.2023**	**373.1758**	1733.8269	1716.8004
4	D, Asp	115.0269	88.0393	477.2344	460.2078	**505.2293**	488.2027	1586.7585	1569.7320
5	D, Asp	115.0269	88.0393	**592.2613**	575.2348	**620.2562**	**603.2297**	1471.7316	1454.7050
6	P, Pro	97.0528	70.0651	689.3141	**672.2875**	717.3090	700.2824	1356.7046	1339.6781
7	S, Ser	87.0320	60.0444	776.3461	759.3196	**804.3410**	**787.3145**	**1259.6519**	1242.6253
8	Z, PhY	243.0297	**216.0420**	1019.3758	1002.3492	1047.3707	1030.3441	**1172.6198**	1155.5933
9	V, Val	99.0684	72.0808	1118.4442	1101.4176	1146.4391	1129.4125	**929.5902**	**912.5636**
10	N, Asn	114.0429	87.0553	1232.4871	1215.4606	1260.4820	1243.4555	**830.5218**	**813.4952**
11	V, Val	99.0684	72.0808	1331.5555	1314.5290	1359.5504	1342.5239	**716.4788**	699.4523
12	Q, Gln	128.0586	101.0709	1459.6141	1442.5876	1487.6090	1470.5825	**617.4104**	**600.3839**
13	N, Asn	114.0429	87.0553	1573.6570	1556.6305	1601.6519	1584.6254	**489.3518**	**472.3253**
14	L, Leu	113.0841	86.0964	1686.7411	1669.7145	1714.7360	1697.7095	**375.3089**	358.2824
15	D, Asp	115.0269	88.0393	1801.7680	1784.7415	1829.7630	1812.7364	**262.2248**	**245.1983**
16	K, Lys	128.0950	101.1073	1929.8630	1912.8365	1957.8579	1940.8314	147.1979	130.1714
17	(0.0851)	0.0851	n/a	1929.9481	1912.9216	1957.9430	1940.9165	19.1029	2.0764

Fig. 5. Sequencing peptides with Analyst BioExplore. Under the BioExplore menu in Analyst 1.0, there is a function called sequence peptide which will bring up the window shown in the figure. The function will sequence the MS/MS spectrum selected from the *.wiff file and propose a set of possible peptide sequences. The sequences proposed by Analyst may not be the same as those proposed by MASCOT. The differential can be attributed to the databases and algorithms used by the two programs to match the MS/MS scans with the peptides. In the sequence peptide function, the fragment window is malleable, enabling other potential sequences to be entered, including the assignment provided by MASCOT. Fragment ions detected in the MS/MS scan will be highlighted and shown in *bold* in the fragment window. This function in Analyst is particularly helpful for calculating b ions, y ions and internal fragments. The charge state can also be changed in the upper right hand corner of the fragment window to calculate the *m/z* value for fragment ions with higher charge states.

fragmentation at the peptide bond and charge retention on the N or C terminus, respectively. If Analyst is not available, the theoretical *m/z* value for each b ion can be calculated by totaling the monoisotopic masses of each amino acid in a given N terminal fragment, adding the mass of the proper number of protons and dividing by the charge. The theoretical *m/z* value for each y ion can also be calculated by totaling the monoisotopic masses of the appropriate residues and adding the atomic mass of water, the proper number of protons and dividing by the number of charges. Water is added to the C terminus of the peptide during the peptide bond cleavage by trypsin or some other protease of choice.

5. For each b ion, draw a line over the fragment's C-terminal amino acid to denote the ion's presence. Also, label the fragment ion directly on the spectra with the notation $b_{\#}$, where the # is the number of residues in the fragment ion, Fig. 6a.

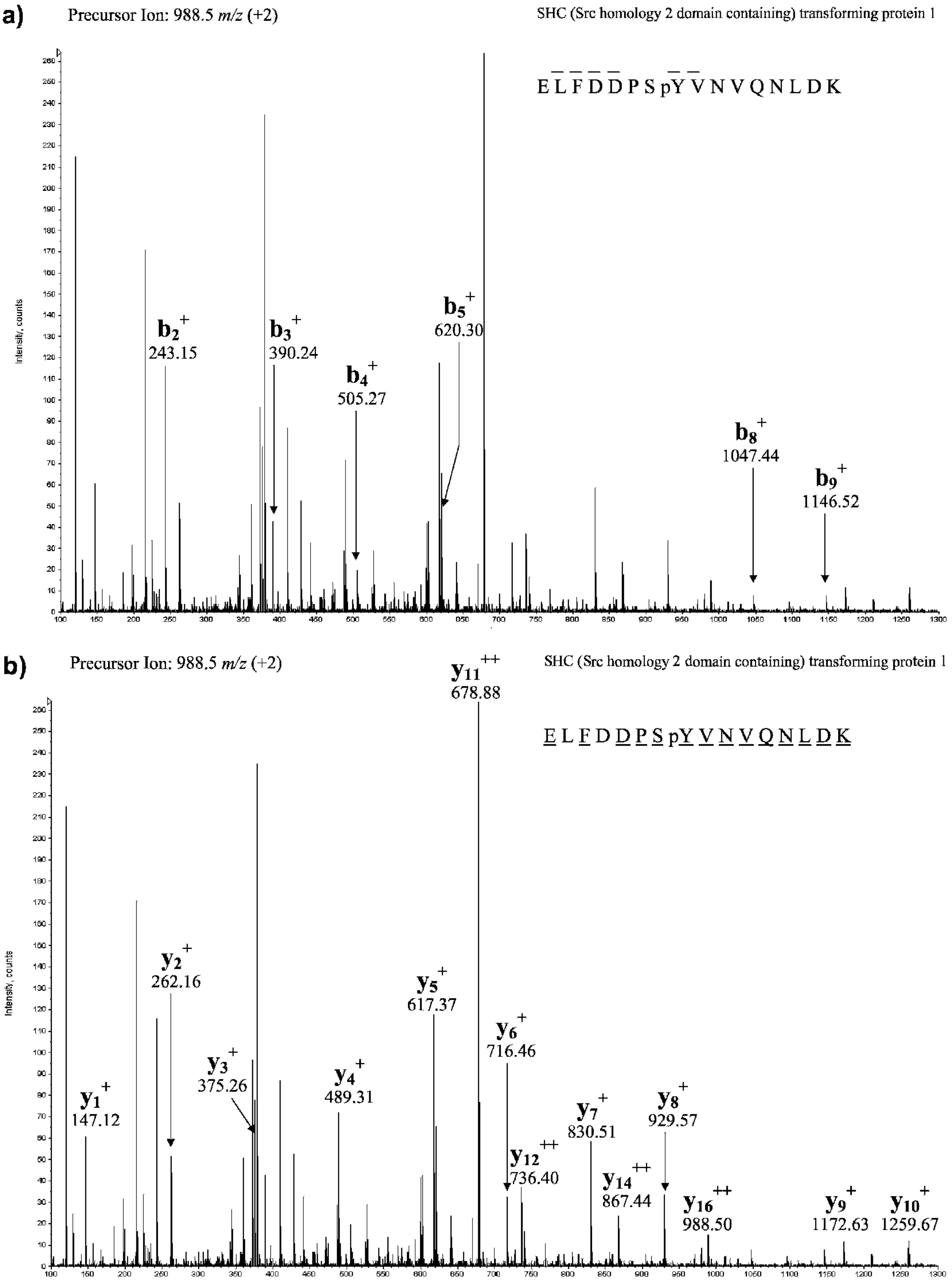


Fig. 6. Assignment of b- and y-type fragment ions in the MS/MS spectrum. Fragmentation of the amide bond results in b ions (**a**) and y ions (**b**), resulting from charge retention at the N or C terminus, respectively. Each b ion identified in the spectra is labeled and a *line* is drawn over the top of the amino acid C-terminal to the peptide fragmentation event. The same procedure is followed for the y ion annotation, but the corresponding amino acid N-terminal to the fragmentation event is underlined on the peptide sequence. The numbering system for b and y ions refers to the total number of residues in the fragment. Initial b and y ion identification is done using the table of values provided by MASCOT or Analyst.

6. For each y ion, draw a line under the fragment's N-terminal amino acid to denote the ion's presence. Also, label the fragment ion directly on the spectra with the notation $y_{\#}$, where the # is the number of residues in the fragment ion, **Fig. 6b**.

7. If the peptide is long enough, all the b and y ions may not appear in the spectrum. To sequence the residues not covered by the +1 charge state b and y ions, identify any +2 b and y ions. Analyst will display the *m/z* for the +2 y and b ions if the ion charge in the upper right hand corner of the fragment ion window is adjusted to +2. If the peptide is of sufficient length, +3 b and y ions may also be found in the spectra.

8. Once all the b and y ions are identified and labeled, the remaining peaks must also be assigned, because unidentified peaks will significantly diminish confidence in the peptide assignment. With most MS/MS spectra, additional fragment ions will help to verify the identity of the peptide. These ions can result from neutral loss of water (H_2O ~ 18 Da), ammonia (NH_3 ~ 17 Da), carbonyl (CO ~ 28 Da (note that b ions displaying loss of CO are commonly referred to as 'a ions')), or phosphate (H_3PO_4 ~ 98 Da or HPO_3 ~ 80 Da) groups, **Fig. 7a, b**. These fragment ions will be separated from the appropriate b or y ion by the indicated number of atomic mass units divided by the charge state of the respective b or y ion.

9. Additional fragment ions may result from internal fragmentation, which is especially prevalent if there is a proline in the sequence. Proline residues normally fragment N-terminal to the residue, producing fragments in either direction, **Fig. 7c**.

10. Identify all the major peaks in the spectra above 300 *m/z* units. Peaks found below 300 *m/z* tend to be more difficult to identify and may contain a large number of immonium ions and internal fragments. However, if one of the top five most abundant peaks in the spectrum is below 300 *m/z*, it should be identified. Due to conservation of mass, all peaks in a given spectrum should be identifiable unless there is contamination in the MS scan. The more peaks that can be identified and assigned in a given spectra, the higher the confidence will be in the peptide identification.

11. If there still remain abundant unassigned peaks in the spectrum after assigning b and y ions, neutral losses and internal fragments, then it is likely that the peptide sequence assignment from Mascot is wrong or there is undefined contamination. To include the spectrum in the data set, attempt to find the actual sequence of the peptide which matches the spectrum, otherwise this peptide should be removed from the data set.

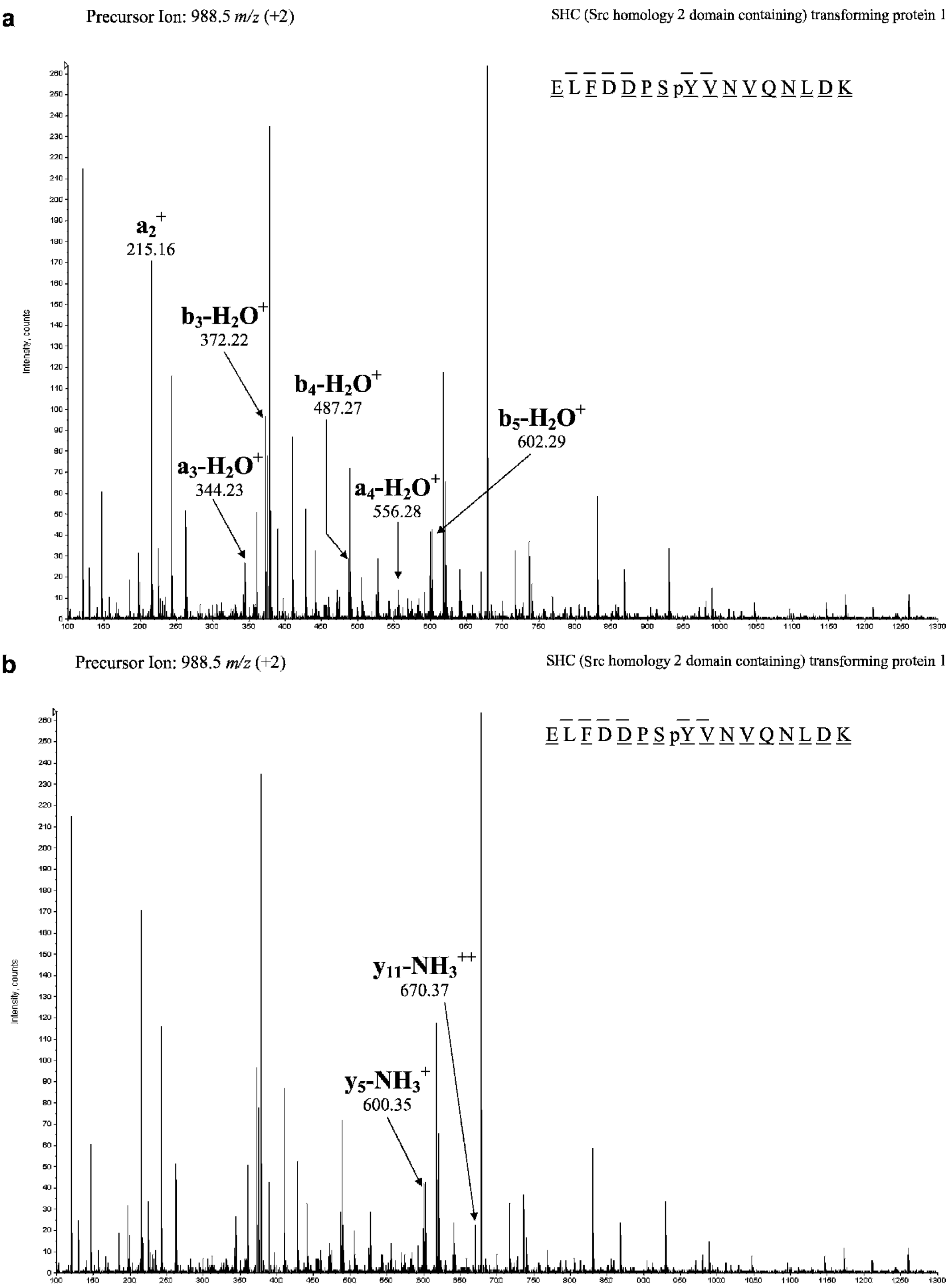


Fig. 7. Assignment of other fragment ions in the MS/MS spectrum. Additional fragment ions in the MS/MS spectrum are the result of neutral loss from b (**a**) or y (**b**) ions, or from internal fragments (**c**). Note that internal fragments can be much harder to identify, but proline cleavage is particularly facile. Internal fragments in either the N- or C-terminal direction often result from an initial cleavage N-terminal to proline and a second cleavage event elsewhere in the peptide sequence.

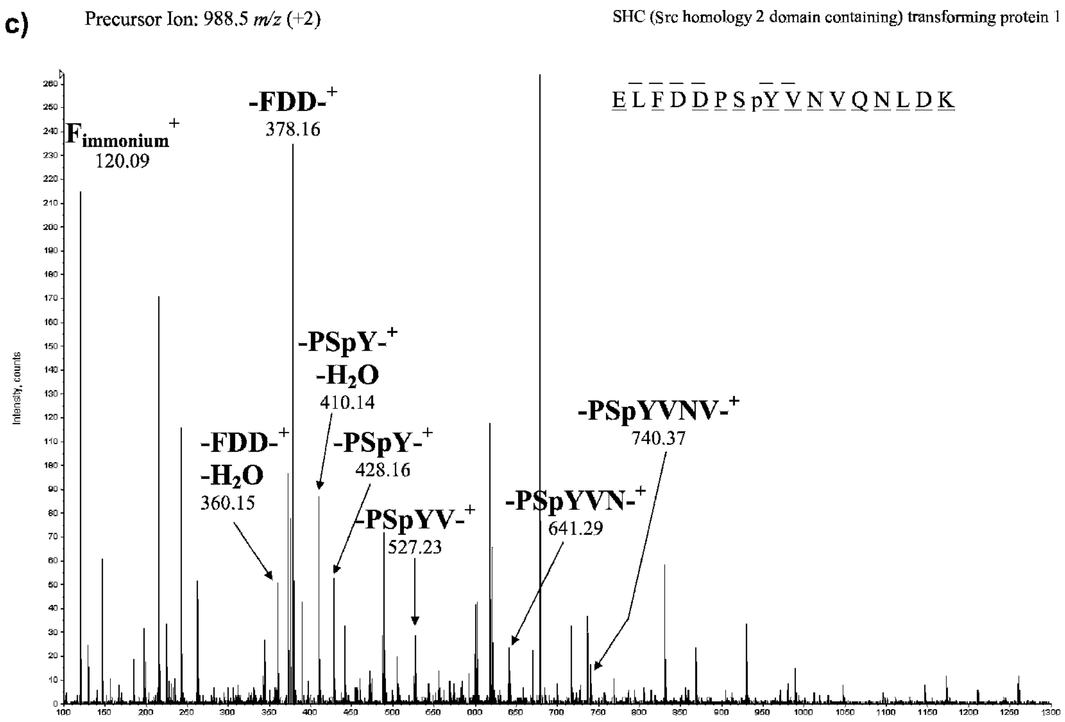


Fig. 7. (continued)

3.6. Phosphorylation Site Identification

1. Tyrosine phosphorylation (pTyr) can be identified by two pieces of evidence: the presence of the phosphotyrosine immonium ion at $m/z = 216.0426$ (+1 charge state) *(8)* and the increased mass to charge of the tyrosine residue in the sequencing from 163.0633 to 243.0297 m/z (+1 charge state) (*see* Note 7).
2. Phosphorylation of serine and threonine residues also increases the residue mass by 79.97 Da. However, serine and threonine phosphorylation is often associated with neutral loss of phosphoric acid from both the precursor ion and from the fragment ions (b, y and internal fragments) containing the phosphorylated residue (*see* Note 8).
3. To identify sites of serine and threonine phosphorylation it is often necessary to look for fragment ions corresponding to the addition of phosphate followed by neutral loss of phosphoric acid. In these instances, final nominal residue masses of 69 Da for phosphorylated serine and 83 Da for phosphorylated threonine are detected. Unfortunately, these masses can also result from neutral loss of water from nonphosphorylated serine and threonine, which can confound phosphorylation site assignment. Fortunately, neutral loss of phosphoric acid from phosphorylated serine and threonine

residues is typically much more abundant that seen for neutral loss of water from non-phosphorylated serine and threonine residues.

4. Pinpointing the exact phosphorylated residue can be difficult if the sequence of b and y ions does not cover the potential site of phosphorylation and multiple tyrosine, serine, or threonine residues are in the unsequenced portion of the peptide.
5. Phosphorylation sites should be reported on a specific residue only if phosphorylation of that residue is explicitly validated from the spectra. Otherwise, it should be explicitly stated that phosphorylation could be on any of several residues in the unsequenced portion of the peptide (Table 3).

Table 3
MS/MS peak list for fragmentation of parent ion 988.51 (+2). All peaks for which charge state can be ascertained are listed and annotated with amino acid sequence and appropriate nomenclature. Unassigned peaks are listed as such and their corresponding intensity is also listed. Other than *m/z* = 441.25, the unassigned peaks either fall below 300 *m/z* or are less abundant than 10% of the maximum spectra intensity of 260. Ideally it should be possible to also assign *m/z* = 441.25; however in practice it is often difficult to account for every ion in the MS/MS spectrum. In the case of a single unidentified ion (at ~15% of the most abundant ion in the spectrum), the confidence in the peptide and phosphorylation site assignment can be very high. The *final two columns* in the table show the theoretical mass of the assigned fragment and the error from the experimental value

Parental ion *m/z:* 988.50 (+2)

Experimental *m/z*	Charge state	Fragment	Peak height	Theoretical *m/z*	Delta
120.09	1	Phenylalanine immonium ion		120.08	0.01
129.11	1	Unassigned			
147.12	1	K (y_1)		147.2	–0.08
185.11	1	Unassigned	20		
197.15	1	Unassigned	35		
215.16	1	EL (a_2)		215.14	0.02
225.15	1	Unassigned	34		
243.15	1	EL (b_2)		243.13	0.02
262.16	1	DK (y_2)		262.22	–0.06
342.21	1	Unassigned	12		

(continued)

Table 3 (continued)

Experimental *m/z*	Charge state	Fragment	Peak height	Theoretical *m/z*	Delta
344.23	1	ELF (a_3-H_2O)		344.2	0.03
360.15	1	–FDD–internal fragment, minus NH_3		360.2	–0.05
372.22	1	ELF (b_3-H_2O)		372.19	0.03
375.26	1	LDK (y_3)		375.31	–0.05
378.16	1	–FDD–internal fragment		378.21	–0.05
390.24	1	ELF (b_3)		390.2	0.04
410.14	1	–PSpY–internal fragment, minus H_2O		410.11	0.03
428.16	1	–PSpY–internal fragment		428.12	0.04
441.25	1	Unassigned	34		
487.27	1	ELFD (b_4-H_2O)		487.22	0.05
489.31	1	NLDK (y_4)		489.35	–0.04
505.27	1	ELFD (b_4)		505.23	0.04
527.23	1	–PSpYV–internal fragment		527.19	0.04
556.28	1	ELFDD (a5–2(H_2O))		556.25	0.03
600.35	1	QNLDK (y_5-NH_3)		600.39	–0.04
602.29	1	ELFDD (b_5-H_2O)		602.25	0.04
617.37	1	QNLDK (y_5)		617.41	–0.04
620.3	1	ELFDD (b_5)		620.26	0.04
641.29	1	–PSpYVN–internal fragment		641.23	0.06
670.37	2	PSpYVNVQNLDK (y_{11}-NH_3)		670.34	0.03
678.88	2	PSpYVNVQNLDK (y_{11})		678.86	0.02
716.46	1	VQNLDK (y_6)		716.48	–0.02
736.4	1	DPSpYVNVQNLDK (y_{12})		736.37	0.03
740.36	1	–PSpYVNV–internal fragment		740.3	0.06
768.4	1	Unassigned	11		
830.51	1	NVQNLDK (y_7)		830.52	–0.01
867.44	2	FDDPSpYVNVQNLDK (y_{14})		867.42	0.02
929.57	1	VNVQNLDK (y_8)		929.59	–0.02
988.5	2	ELFDDPSpYVNVQNLDK (y_{16}), parent ion		988.48	0.02

(continued)

Table 3 (continued)

Experimental *m/z*	Charge state	Fragment	Peak height	Theoretical *m/z*	Delta
1,047.44	1	ELFDDPSpY (b_8)		1,047.37	0.07
1,146.52	1	ELFDDPSpYV (b_9)		1,146.44	0.08
1,172.63	1	pYVNVQNLDK (y_9)		1,172.62	0.01
1,210.61	1	Unassigned	6		
1,259.67	1	SpYVNVQNLDK (y_{10})		1,259.65	0.02

4. Notes

1. Alternative searching programs exist and can be employed for matching MS/MS spectra and peptide sequences. Each program will have its own caveats and biases. Other programs include but are not limited to SEQUEST, PROQUANT, ProteinPilot, SpectrumMill, and others.
2. Our protocol *(7)* specifically isolates phosphotyrosine containing peptides. This method will rarely identify peptides containing phosphoserine and phosphothreonine. To analyze the latter, a general phosphorylation enrichment strategy should be employed, such as IMAC, strong cation exchange *(9)*, or titanium dioxide (TiO_2) *(10, 11)*.
3. The MASCOT results from a search of your data set will produce peptide sequence assignments and associated protein assignments. It is important to realize that multiple proteins may have the same peptide sequence. For a comprehensive list of proteins containing a given peptide sequence, perform a BLAST search through NCBI (http://www.ncbi.nlm.nih.gov/BLAST).
4. The isolation window taken around a precursor ion for fragmentation will depend on the resolution and settings of the mass spectrometer being employed.
5. The sequence assigned by Analyst 1.0 may not be the same as the sequence assigned by MASCOT. Differential sequence assignment between Analyst and MASCOT reflects the different biases in how the databases and algorithms are set up. Disagreement between the two programs exemplifies the need for manual spectral validation prior to presenting mass spectrometry data.

6. Nomenclature for altered residues can be found and modified in the Data Dictionary in the BioTools menu of Analyst 1.0.
7. The intensity and presence of phosphotyrosine immonium ion (*m/z* = 216.0462, +1 charge state) is dependent on the sequence of the peptide and the fragmentation energy of the instrument. The absence of the immonium ion peaks is not sufficient to exclude phosphotyrosine residues from a peptide sequence.
8. As with tyrosine phosphorylation, the peptide sequence plays a significant role in determining the extent of neutral loss of phosphoric acid from serine or threonine phosphorylated precursor or fragment ions. Multiply phosphorylated peptides and peptides with basic residues (R, K, H) tend to exhibit intense neutral loss peaks.

Acknowledgements

The authors thank Paul Huang, Alejandro Wolf-Yadlin, and other members of the White lab, for helpful discussions. This work was supported in part by a National Science Foundation Graduate Research fellowship, NIAID grant R21-AI065354 and by NIH grant P50-GM68762.

References

1. Zhang, Y, Wolf-Yadlin, A, Pappin, DJ, Rush, J, Lauffenburger, DA, White, FM. Time resolved mass spectrometry of tyrosine phosphorylation sites in the EGF receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics* 4, 1240–1250 (2005).
2. Schmelzle, K, Kane, S, Gridley, S, Lienhard, GE, White, FM. Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes* 55(8), 2171–2179 (2006).
3. Schmelzle, K, White, FM. Phosphoproteomic approaches to elucidate cellular signaling networks. *Curr. Opin. Biotechnol.* 17(4), 406–414 (2006).
4. Kim, JE, White, FM. Quantitative analysis of phosphotyrosine signaling networks triggered by CD3 and CD28 costimulation in Jurkat cells. *J. Immunol.* 176(5), 2833–2843 (2006).
5. Moser, K, White, FM. Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. *J. Proteome Res.* 5(1), 98–104 (2006).
6. Kim, JE, Tannenbaum, SR, White, FM. Global phosphoproteome of HT-29 human colon adenocarcinoma cells. *J. Proteome Res.* 4, 1339–1326 (2005).
7. Zhang, Y, Wolf-Yadlin, A, White, FM. Quantitative proteomics analysis of phosphotyrosine-mediated cellular signaling networks. *Methods Mol. Biol.* 359, 203–212 (2007).
8. Steen, H, Küster, B, Fernandez, M, Pandey, A, Mann, M. Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. *Anal. Chem.* 73(7), 1440–1448 (2001).
9. Beausoleil, SA, Jedrychomski, M, Schwartz, D, Elias, JE, Villén, J, Li, J, Cohn, MA, Cantley, LC, Gygi, SP. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl Acad. Sci. U. S. A.* 101(33), 12130–12135 (2004).

10. Pinkse, MWH, Uitto, PM, Hilhorst, MJ, Ooms, B, Heck, AJR. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal. Chem.* 76(14), 3935–3943 (2004).
11. Larsen, MR, Thingholm, TE, Jensen, ON, Roepstorff, P, Jørgensen, TJD. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* 4(7), 873–886 (2005).

Chapter 9

Assigning Glycosylation Sites and Microheterogeneities in Glycoproteins by Liquid Chromatography/Tandem Mass Spectrometry

Yehia Mechref, Milan Madera, and Milos V. Novotny

Summary

Glycosylation of proteins is one of the most common posttranslational modifications which has its bearing on function and biological activity. Assigning the glycosylation sites and their inherent microheterogeneities are key structural issues addressing various glycoprotein functions. This chapter describes three different approaches all based on liquid chromatography/tandem mass spectrometry (LC/MS-MS), which are commonly employed for the assignment of protein glycosylation sites and their microheterogeneities. Comparing the LC/MS-MS analysis of a native glycoprotein tryptic digest to that of a deglycosylated tryptic digest can be accomplished through a routine LC/MS instrument. The use of a scanning mass spectrometer capable of switching between high-voltage and low-voltage scans, combined with monitoring carbohydrate-characteristic oxonium ions, is yet another analytical approach utilized for characterization of the glycosylation sites of glycoproteins. These two approaches do not address the problem originating from the ion suppression associated with coeluting peptides. The use of on-line glycopeptide enrichment in conjunction with LC/MS-MS is a third approach, which reduces ion suppression, thus offering a more sensitive approach to the characterization of protein glycosylation sites.

Key words: Glycoproteins, Glycosylation sites, Microheterogeneity of glycosylation site, LC/MSMS, Deglycosylation of glycopeptides

1. Introduction

It is most common that glycosylation of a particular protein is investigated through the chemical or enzymatic release of glycans and their subsequent characterization, such as sequencing and linkage analysis. While this information is highly important, there are additional structural aspects pertaining to glycosylation that must be addressed. For each site of glycosylation, there are

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_9

possible structural variations (extent of substitution, site-specific microheterogeneities, a site-specific accessibility for particular glycosyltransferases, etc.), which can all have important biochemical consequences *(1–6)*. Investigating protein glycosylation at the level of glycopeptide is at least as important as the investigation of released glycan structures *(7)*.

While techniques based on mass spectrometry (MS) now play an essential role in analyzing the glycan moiety of a glycoprotein, they are also being uniquely useful for identifying the sites of glycosylation *(7–10)*. When using standard proteomic procedures, it is quite difficult to directly identify glycopeptides in a complex protein digest by MS. This is partly due to the low sensitivity of the detection of glycopeptides caused by site heterogeneity and/or the ion-adduct formation and signal suppression in the presence of other peptides, especially if the glycans are terminated with the negatively charged sialic acid moieties *(7)*. Moreover, the glycan heterogeneity and a frequent multiple adduct formation result in the overall glycopeptide signal being distributed into several peaks, resulting in weak signals detected by MS. Although MS using matrix-assisted laser desorption-ionization (MALDI) and electrospray ionization (ESI) on unseparated proteolytic digests have been employed in the assignment of the glycosylation sites, due to the high complexity of heterogeneous peptide mixtures *(11–18)* it appears necessary to use multiple analytical procedures, either in parallel, or sequentially. Interfacing MS with nanoscale liquid chromatography (nano-LC) has been particularly attractive due to small sample handling, different selectivity modes of separation and high separation efficiency.

Currently determination of the glycosylation sites in glycoproteins has been most commonly achieved using ESI mass spectrometers equipped with either a triple-quadrupole analyzer, or more recently, a hyphenated quadrupole/time-of-flight (Q-TOF) instrument. The use of high-cone voltages, to promote “in-source” fragmentation of ions, in conjunction with selected-ion monitoring (SIM) experiments, allows specific detection of glycopeptides from proteolytic digests which are either infused without separation *(19, 20)* or separated by on-line LC *(19, 21–23)*. In the case of Q-TOF MS analysis, determination of the glycosylation sites is based on the precursor ion scanning for the terminal saccharide product ions originating from high-energy CID achieved in the collision cell *(23–25)*. In both cases, a precursor ion scan is performed for one or more of the glycopeptide diagnostic oxonium ions, such as *N*-acetylhexosamine oxonium ion (*m/z* 204), hexose-*N*-acetylhexosamine oxonium ion (*m/z* 366), *N*-acetylneuraminic acid oxonium ion (*m/z* 292), *N*-acetylneuraminic acid, minus water, oxonium ion (*m/z* 274) and *N*-acetylneuraminic acid-hexose-*N*-acetylhexosamine oxonium ion (*m/z* 657). The uses of other mass spectrometers for the assessment of the glycosylation sites of glycoproteins have also been reported *(26–28)*. Recently, protein glycosylation was analyzed by hydrophilic interaction (HILIC) nano-LC/MS of glycopeptides *(29)*. The approach involved a non-specific proteolytic digestion with pronase, resulting in glycopeptides with mostly 2–8 amino acids which are easily resolved by HILIC/MS using the amide-bounded silica beads. In this case, separation of glycopeptides was dependent on the size of a glycan chain, thus prompting the separation of peptide moieties to which different *N*-glycan structures were attached *(29)*.

An alternative approach is based on fractionation of peptides and glycopeptides using different media, including porous graphitized carbon *(30)*, cellulose or Sepharose *(31, 32)*, zwitterionic silica-based materials *(33)* or lectins *(34–36)*, followed by an off-line characterization of the glycopeptide-containing fractions through MS. More recently, an on-line lectin-based glycopeptide enrichment for the enhanced determination of the glycosylation sites in glycoproteins has been introduced *(37)*. This approach is based on the use of silica-based lectin media, capable of sustaining high back pressures commonly associated with the on-line nano-LC systems. The use of lectin microcolumns allowed a substantial enhancement in the sensitivity of the site-of-glycosylation measurements, since it eliminated all coeluting peptides which otherwise suppress ionization of glycopeptides.

Three different approaches commonly employed in the assignment of glycosylation sites and their microheterogeneity will be described below. Each approach offers certain advantages and disadvantages. Employing glycan endoglycosidase, releasing *N*-glycans from the peptide backbone, followed by comparing LC/MS-MS analyses of native and deglycosylated tryptic digest samples of glycoproteins is the first approach. The second approach involves the use of a scanning mass spectrometer, while the third approach involves an on-line enrichment of glycopeptides, utilizing silica-based lectin material. The use of a scanning MS allows the monitoring of carbohydrate-characteristic oxonium ions generated during high-voltage scanning. In contrast, the use of silica-based lectin materials, capable of withstanding high-back pressures, permits the on-line enrichment of glycopeptides prior to their LC/MS-MS analyses.

2. Materials

2.1. Tryptic Digestion

1. Bovine fetuin and monoclonal antibody (mAb) (Sigma-Aldrich, St. Louis, MO).
2. Proteomic-sequencing grade trypsin (Sigma-Aldrich, St. Louis, MO).

3. Dithiothreitol (DTT) and iodoacetamide (IAA) (BioRad, Hercules, CA).
4. 50 mM ammonium bicarbonate buffer/pH 7.5–8.0; 1 mM aqueous hydrochloric acid solution.

2.2. Deglycosylation of a Glycoprotein Tryptic Digest

1. Ammonium bicarbonate buffer (50 mM solution, pH 7.5–8.0).
2. *N*-glycanase F (PNGAse F) from *Chryseobacterium meningosepticum* (EC 3.5.1.52).

2.3. LC/MS

1. *Solvent A*: 3% aqueous acetonitrile containing 0.1% formic acid;
2. *Solvent B*: 0.1% formic acid in acetonitrile.
3. Fetuin and murine monoclonal antibody tryptic digest (mAb) (IgG1K).
4. Nano LC-MS/MS system. Waters CapLC pump with an injector, 10-port valve, Q-TOF Ultima Global mass spectrometer fitted with nanoESI source has been utilized in the methods described below (*see* **Note 1**).
5. C_{18} trapping cartridge (Dionex, Sunnyvale, CA); C_{18} capillary column (75 µm i.d., 15 cm length).

2.4. Activation of Macroporous Silica

1. Macroporous silica (10 µm, 1,000 Å) from Macherey Nagel (Easton, PA) or EMD Chemicals (Gibbstown, NJ).
2. Dry toluene; acetone; ethanol; ether, 3-glycidoxypropyltrimethoxysilane; triethylamine; glacial acetic acid; 10 mM sulfuric acid; 6 M hydrochloric acid; sodium periodate (Sigma-Aldrich, St. Louis, MO).

2.5. Immobilization of SNA Lectin to Aldehyde-Silica

1. Lectin from *Sambucus nigra* (SNA) (Sigma-Aldrich, St. Louis, MO).
2. Coupling solution: 0.1 M sodium bicarbonate buffer containing 0.5 M sodium chloride and 10 mg/mL sodium cyanoborohydride.
3. Sodium borohydride, which is used to reduce aldehyde groups.
4. *Lectin-binding buffer*: 10 mM Tris–HCl (pH 7.4), 150 mM NaCl, 1 mM $CaCl_2$, 0.02% sodium azide.

2.6. Packing of Lectin Microcolumns

1. PEEK tubing, unions and 0.5 µm stainless steel frits (Upchurch Scientific, USA).
2. Acetone.
3. Isocratic HPLC pump.

3. Methods

3.1. Tryptic Digestion of Bovine Fetuin

1. Incubate 100 μL of glycoprotein solution (0.1 μg/μL in 50 mM ammonium bicarbonate) at 95°C for 15 min.
2. Cool down to room temperature prior to the addition of a 2-μL aliquot of 45 mM DTT and incubate at 50°C for 15 min to reduce disulfide bonds to thiols.
3. Cool down to room temperature and add a 2-μL aliquot of 100 mM IAA
4. Allow alkylation of thiols functional groups to proceed for 30 min in the dark at room temperature.
5. Prepare fresh trypsin solution by resuspending trypsin in 1 mM hydrochloric acid to give a 1 μg/μL trypsin solution.
6. Add a 0.1-μL aliquot of the 1 μg/μL trypsin solution to the reduced and alkylated glycoproteins and allow digestion to proceed for 18 h at 37°C.
7. Stop the enzyme activity by boiling the sample for 5 min (*see* **Note 2**).
8. Allow the reaction mixture to cool down to room temperature and use immediately or store at –20°C.

3.2. Deglycosylation of a Glycoprotein Tryptic Digest (See Note 3)

1. Add a 0.5-μL aliquot of PNGase F (5 mU/mL) to the tryptic digest solution prepared as described in **Subheading 3.1**, and incubate at 37°C overnight (*see* **Note 4**).
2. Stop the reaction by boiling for 5 min to denature PNGase F and thus terminate its enzymatic activity (*see* **Note 2**).

3.3. Assigning Glycosylation Sites Through Comparing the LC/MSMS Analyses of Glycoprotein Tryptic Digests and Deglycosylated Glycoprotein Tryptic Digest

The glycosylation sites of a glycoprotein could be determined by comparing the LC/MS-MS analysis of a glycoprotein tryptic digest to the digest which has been subjected to enzymatic release of glycans prior to such an analysis. Peptide-N glycosidases, such as PNGase F, hydrolyze the bond between asparagine and the reducing end *N*-acetylglucosamine (GlcNAc) residue. This reaction converts asparagine to aspartate, with a mass increase of 0.98 Da. Accordingly, the site of glycosylation could be assigned by comparing the corresponding LC/MS-MS analyses of the two samples. Moreover, a mass spectrometer with adequate mass resolution and mass accuracy should be able to distinguish such a difference in *m/z* as a result of deglycosylation and subsequently, should allow the identification of the site of glycosylation. Although this approach has been regularly employed, the data interpretation and processing can be tedious and time consuming.

1. Connect the sample injector to a 10-port valve configured with the trapping column, C_{18} capillary column, which is directly attached to the nano-ESI source.

2. Set the flow rates of the isocratic and system pump to the following values, system pump: 3% of solvent A at 250 nL/min; and isocratic pump: solvent A at 5 μL/min.
3. Inject the native tryptic digest sample with the 10-port valve in a loading position, thus trapping peptides and glycopeptides on the C_{18} trapping cartridge.
4. Wash the trapping cartridge for 10 min with solvent A for a complete sample desalting.
5. Change position of the 10-port valve to injection, thus connecting the trapping cartridge to the C_{18} capillary column interfaced to the mass spectrometer. In this format, the trapped peptides and glycopeptides are eluted from the trapping cartridge to the analytical C_{18} capillary column.
6. Separate the desalted peptides using a 55-min linear gradient (3–55% B).
7. Analyze the separated peptides by mass spectrometry, employing a data-dependent acquisition, which involves tandem MS analysis of the separted peptides.
8. Repeat **steps 3–7** for the *N*-deglycosylated tryptic digest.
9. Process the data collected from *N*-deglycosylated tryptic digest and identify the peptides present in the sample by using the peptide mass fingerprinting strategy or database searching with asparagine-to-aspartate modification.
10. Process the data collected from native tryptic digest using the same strategy utilized for the analysis of deglycosylated tryptic digest and check the masses of corresponding glycopeptides, based on their similarity in retention time. Base-peak intensity chromatograms of native and deglycosylated tryptic digest of fetuin are exemplified in **Fig. 1**. The mass shift originating from PNGase F (and endured by the peptides) provides the information pertaining to the masses of the attached *N*-glycan moieties (*see* **Note 5**). The molecular weights of native and deglycosylated peptides, originating from fetuin, as well as their observed *m/z* values are listed in **Table 1**, while the mass spectra of the deglycosylated peptides are depicted in **Fig. 2**.
11. Run GlycoMod software (http://www.expasy.org/tools/glycomod/) on the peptides which are expected to be glycosylated, in order to predict their *N*-glycan structures.

3.4. Assignment of N-Glycosylation Sites Using Product-Ion Discovery Experiment

The glycosylation sites in glycoproteins can also be determined using a mass spectrometer capable of performing a parent-ion discovery experiment. In such an experiment, the voltage on the gas collision cell is routinely switched between low (8 V) and high (35 V) during every second. Under these conditions, a standard MS spectrum and a high-energy MS spectrum consisting of all

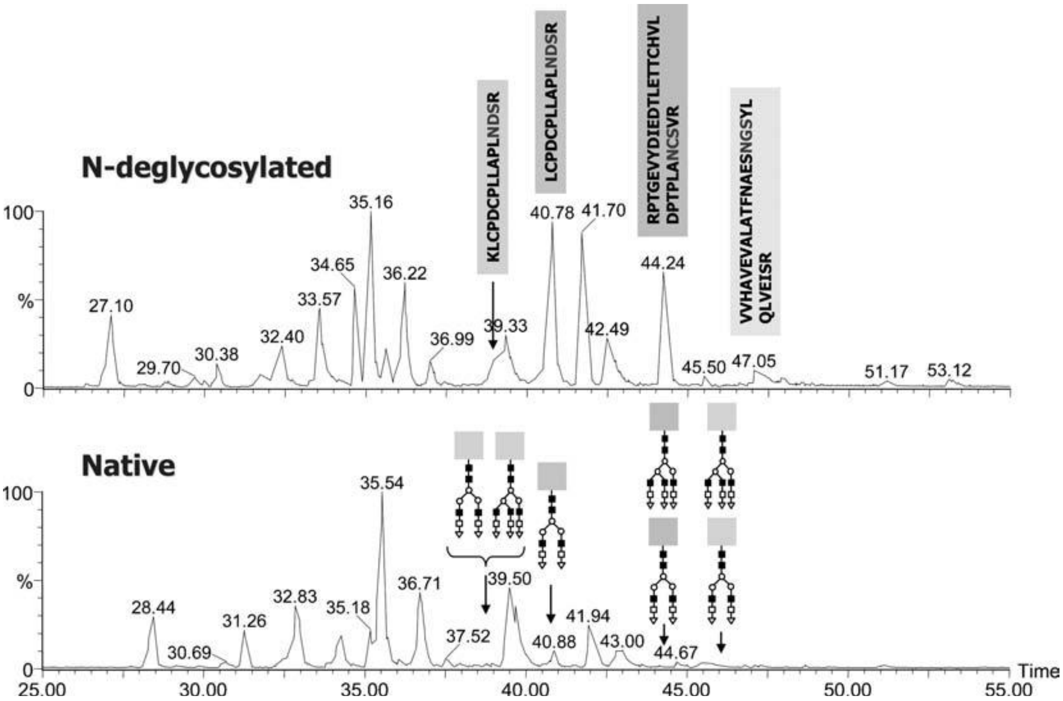


Fig. 1. Base-peak intensity chromatograms of fetuin tryptic digests of deglycosylated and native samples.

product ions generated from the precursor ions are commonly observed in the normal scan (low-voltage scan). In the presence of glycopeptides, the carbohydrate-characteristic oxonium ions (i.e., *m/z* 204, 366, 290, and 308) will be observed in the high-voltage scan. The instrument could be configured to acquire tandem MS of the most intense ions only if the carbohydrate-characteristic oxonium ions are detected in the high-voltage scan. This approach is more effective than the approach described in Subheading 3.3. Interpretation of the generated data here is simpler, since data analysis is only needed where the oxonium ions are detected during the LC/MS-MS experiment.

1. Connect the sample injector to the 10-port valve configured with the trapping column, C_{18} capillary column, which is directly attached to the nano-ESI source.
2. Set the flow rates of the isocratic and the system pumps to the following values: system pump: 3% of solvent A at 250 nL/min; isocratic pump: solvent A at 5 μL/min.
3. Have the 10-port valve in a loading position, inject the sample and allow the peptide and glycopeptide mixture to be trapped and desalted on the cartridge.
4. Switch the 10-port valve to the injection position and separate the desalted components using a 60-min gradient (3–40% B).

Table 1
Molecular weights of fetuin glycopeptides and their respective deglycosylated peptides

Native glycopeptides		*N*-deglycosylated glycopeptides			
				Observed	
Sequence	Calc Mr	Sequence	Exp Mr	m/z	Delta
LCPDCPLLAPL**NDS**R	1739.8334	LCPDCPLLAPL**DDS**R	1740.69	871.35	0.8566
KLCPDCPLLAPL**NDS**R	1867.9284	KLCPDCPLLAPL**DDS**R	1868.77	623.93	0.8416
RPTGEVYDIEIDTLETTCHVLDPT-PLA**NCS**VR	3670.7606	RPTGEVYDIEIDTLETTCHVLDPTPLA**DCS**VR	3671.44	918.87	0.6794
VVHAVEVALATFNAES**NGS**YLQLVEISR	3015.5665	VVHAVEVALATFNAES**DGS**YLQLVEISR	3016.34	1006.45	0.7735

Deglycosylation results in converting asparagine to aspartate with an inherent increase in molecular weight

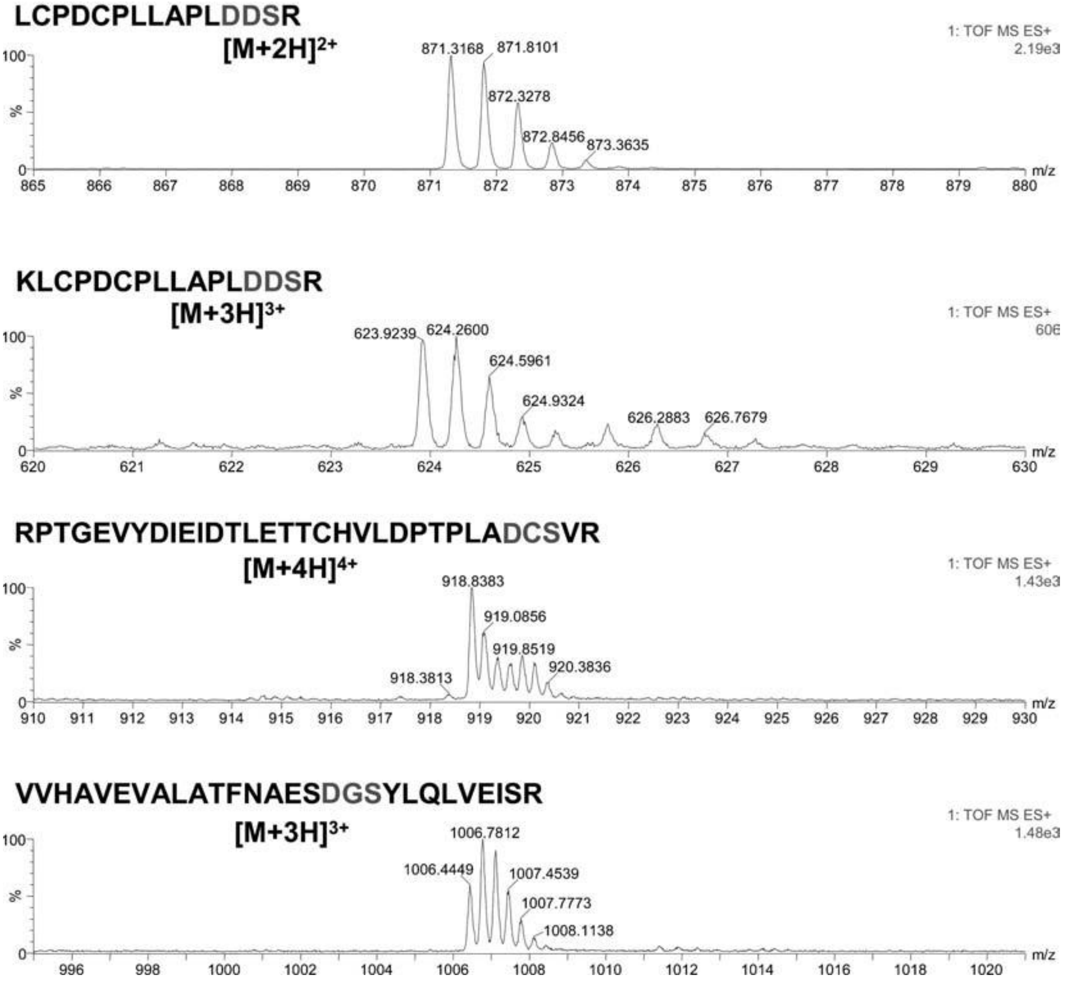


Fig. 2. Mass spectrum of the deglycosylated peptides observed in Fig.1.

5. Set the MS method in such a way that the voltage on collision gas is switched between “high” (35 V) and “low” (8 V) every second. This provides both a standard low-energy MS spectrum and a high-energy MS spectrum of all product ions generated from the precursor ions seen in the normal low-energy scan. After detecting the carbohydrate-characteristic oxonium ions (*see* **Table 2**), the Q-TOF instrument switches to the MS/MS mode and selects the most intense, triply- or multiply- charged-state ions for fragmentation. The MS/MS analysis is performed for 6 s at 1-s scan rate. This process is repeated until the eight most intense precursor ions during a single scan become selected for MS/MS experiments. During the MS/MS experiment, a collision energy ramp from 20 to 40 V is applied to provide as much structural information as possible. The MS/MS spectra collected this way provide

Table 2
Carbohydrate-characteristic oxonium ions commonly observed in PID experiment

Carbohydrate-characteristic oxonium ion ([M+H]$^+$)	Accurate *m/z*	Average *m/z*
Hex	163.0606	163.2
HexNAc	204.0872	204.2
Hex-HexNAc	366.1400	366.3
NeuAc-Hex HexNAc	657.2354	657.6
NeuAc	292.1032	292.3
NeuAc (with water loss)	274.0927	274.3
NeuGc	308.0981	308.3
NeuGc (with water loss)	290.0876	290.3

information pertaining to both the site of glycosylation and the glycan structures attached.

6. The analysis of the glycosylation sites of a monoclonal antibody (IgG) from murine cells, employing this approach, is illustrated in Fig. 3. This analysis has revealed the presence of two sites of glycosylation associated with the glycoprotein. One site was associated with the conserved domain of monoclonal antibodies, while the other was on the variable domain. Only monosialylated structures have been commonly associated with the glycan structures of the conserved domain. This fact was also true for the mAb under investigation. Asialylated and monosialylated glycans were only associated with this site of glycosylation. Sialylation of the glycan structures present on the Fc region glycosylation site is evident from Fig. 4. Three glycopeptides are depicted in Fig. 4 for which the charge states and tandem MS experiments are illustrated in the insets. The depicted glycopeptides are triply charged and their tandem mass spectra reveal the presence of ions at *m/z* values of 204, 290, 308 and 366, corresponding to GlcNAc, NeuGc minus water, NeuGc and Gal-GlcNAc oxonium ions, respectively. In addition, the peptide backbone with only GlcNAc and GlcNAc-Fuc is observed in the spectrum of each glycopeptide.
7. All the other sialylated structures were solely associated with the glycosylation site of the variable domain. The inset of Fig. 3 depicts an MS/MS spectrum of a glycopeptide possessing a sialylated structure. It should be noted that the ions at *m/z* values of 204, 290, 308 and 366 correspond to GlcNAc,

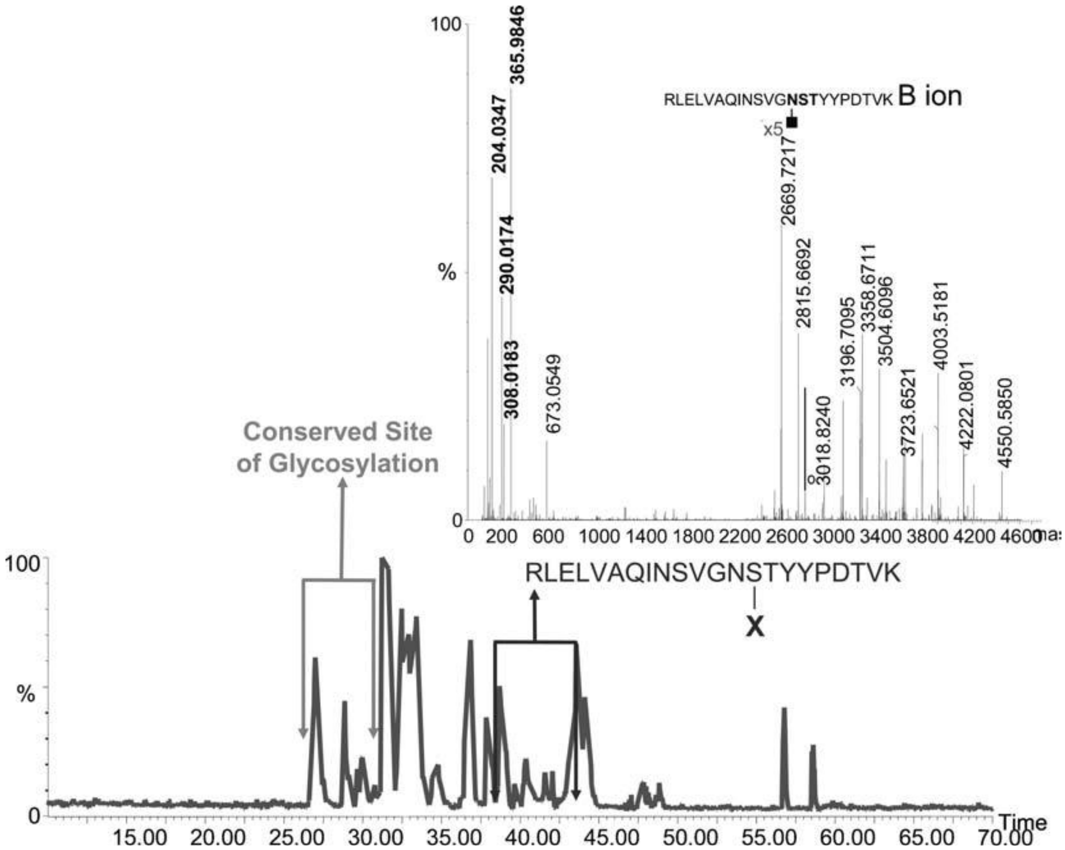


Fig. 3. Base peak intensity chromatogram of tryptically digested mAb. Inset, MS/MS spectrum of a glycopeptide in which the glycan-diagnostic product ions with *m/z* values of 204, 290, 308 and 366 are depicted with blue lines. The signal labeled in the inset represents the B ion of this glycopeptide resulting from the loss of all monosaccharides residues, except the GlcNAc residue attached to the peptide backbone. This ion is commonly the most intense ion observed in the tandem MS spectrum of a glycopeptide. (Reproduced from *(24)* with permission).

NeuGc minus water, NeuGc and Gal-GlcNAc oxonium ions, respectively. The most intense ion observed at high *m/z* values corresponds to the peptide backbone with one GlcNAc residue attached. This ion is known as the B ion. All ions preceding the B ions result from the loss of one or more sugar residues (i.e., the ion at 2815.6692 is that of the peptide backbone with GlcNAc and Fuc attached to it, while another at 3018.8240 is corresponding to the peptide backbone with GlcNAc, Fuc and GlcNAc attached). The detection of NeuGc is expected from the cellular source of this mAb. The selective, ion-extracted chromatograms of the different glycopeptides are depicted in **Fig. 5**. Seven sialylated structures are associated with the variable site of glycosylation, as suggested by the data presented in **Fig. 5**. Moreover, the microheterogeneities of both glycosylation sites of this monoclonal antibody are summarized in **Table 3**.

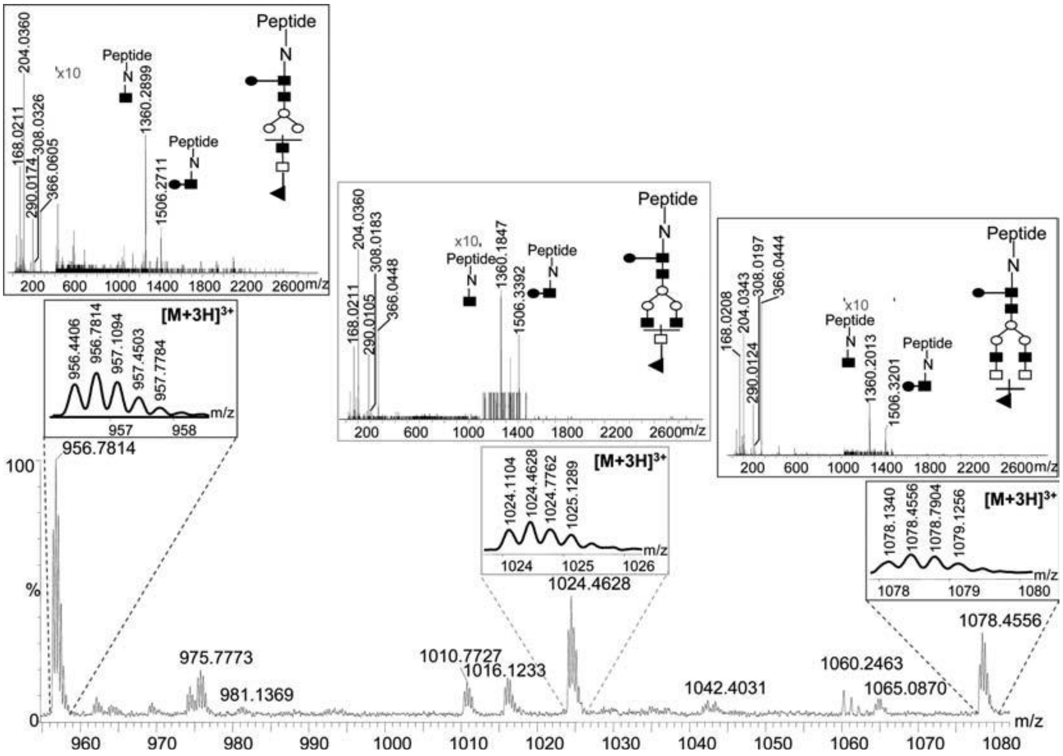


Fig. 4. MS spectrum of sialylated glycopeptides associated with the Fc region of the antibody. The lower inset represents the charge state of different glycopeptides, while the upper inset represents the tandem MS spectra of each of the glycopeptides. Structures of the glycan attached to the peptide backbones are represented by the following symbols: GlcNAc, (*filled square*); Fucose, (*filled circle*); Mannose, (*open circle*); Galactose, (*open square*); *N*-acetylglycolylneuraminic acid, (*filled rectangle*). (Reproduced from *(24)* with permission).

3.5. Online Enrichment of Fetuin Glycopeptides Using Microcolumn Lectin Affinity Chromatography

Although the two approaches described above have been employed effectively for the assignment of the glycosylation sites of glycoproteins, they fail to overcome the issues associated with ion suppression originating from coeluting peptides. An enrichment of glycopeptides using chromatographic media prior to MS analysis, is an approach allowing a complete removal or reduction of coeluting peptides, thus eliminating/reducing ion suppression. One of the chromatographic media that has been effectively employed for the isolation and enrichment of glycoproteins are lectins, which specifically enrich just glycopeptides. Recently, a silica-based format of lectin has been utilized for the online enrichment of glycopeptides of glycoprotein tryptic digests *(37)*.

3.5.1. Derivatization of Macroporous Silica

1. Wash the porous silica sequentially with water, 6 M hydrochloric acid and water again, until the resulting suspension has neutral pH.
2. Filter the activated silica particles on a sintered glass frit and dry at 150°C overnight.

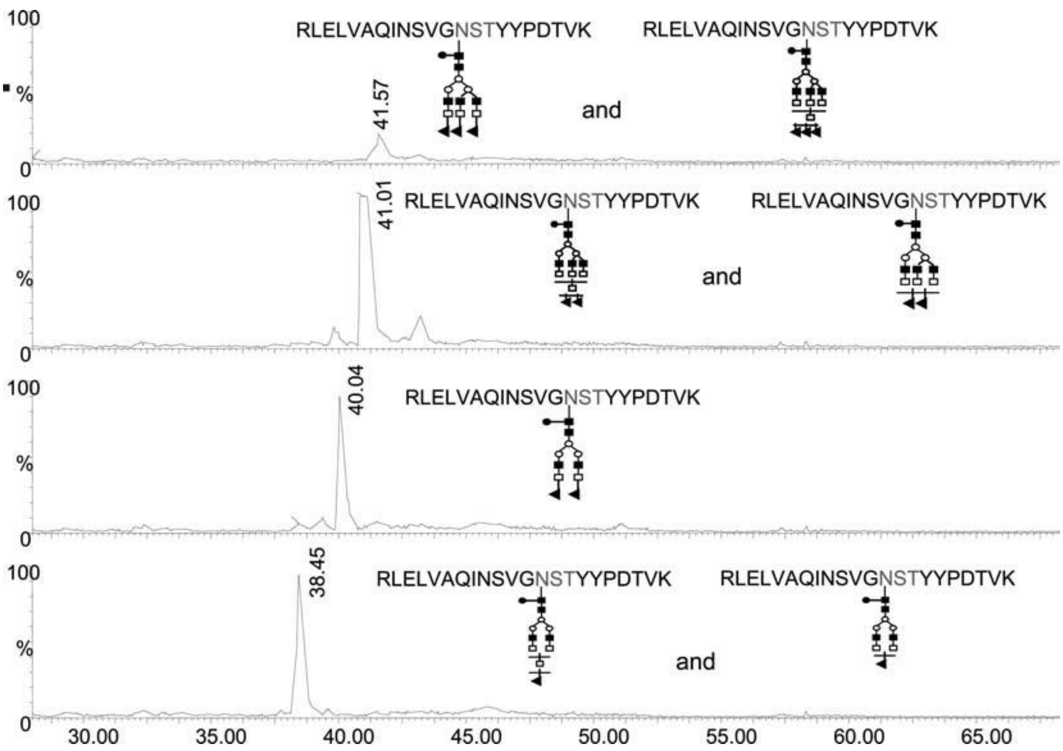
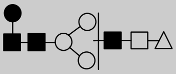
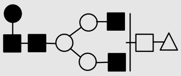
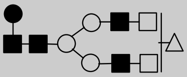
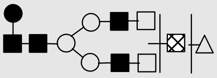
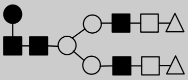
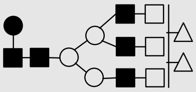
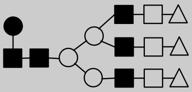
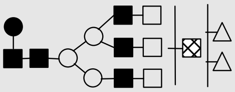
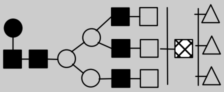
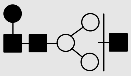
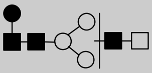
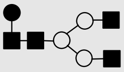
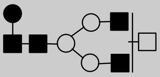
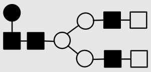
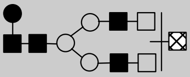


Fig. 5. Selective ion extraction chromatograms of the glycopeptides associated with the variable region. The structures of different glycans attached to this glycosylation site are represented by the same symbols as in **Fig. 3**. (Reproduced from *(24)* with permission).

3. Resuspend a 1-g aliquot of dry silica in 15 mL of dried toluene (molecular sieve).
4. Add a 200-μL aliquot of 3-glycidoxypropyltrimethoxysilane, followed by a 5-μL aliquot of triethylamine which acts as a catalyst.
5. Stir the mixture at 105°C for 16 h under reflux.
6. Filter the resulting epoxy-silica, wash consecutively with toluene, acetone, ether and dry under reduced pressure.
7. Resuspend about 1 g of epoxy-silica in 10 mM sulfuric acid and stir at 90°C for 3 h to hydrolyze the attached epoxy groups to diols.
8. Wash the resulting diol-silica sequentially with water, ethanol, ether and let it dry under reduced pressure.
9. Mix a 1-g aliquot of the freshly prepared diol-silica with 20 mL acetic acid/water (90/10).
10. Add a 1-g aliquot of sodium periodate to the reaction mixture and stir at room temperature for 2 h to oxidize diol groups to aldehydes.

Table 3
Microheterogeneity of the glycosylation sites determined from murine IgG

Structure	Variable region	Preserved region
		✓
		✓
	✓	✓
	✓	
	✓	
	✓	
	✓	
	✓	
	✓	
		✓
		✓
		✓
		✓
		✓
		

11. Wash the aldehyde-silica with water, ethanol and ether, dry it at reduced pressure and store at 4°C prior use.

3.5.2. Immobilization of SNA Lectin to Aldehyde-Silica

1. Solubilize a 2-mg aliquot of SNA lectin in 1 mL of 0.1 M sodium bicarbonate containing 0.5 M sodium chloride and 10 mg sodium cyanoborohydride and stir gently at room temperature for 3 h (*see* **Note 6**).
2. Add 3 mg of sodium borohydride in steps over 30 min to reduce excessive aldehyde groups to diol groups (*see* **Note 7**).
3. Stir the reaction mixture at room temperature for an hour to ensure effective conversion of all free aldehyde groups to diol groups.
4. Wash SNA-silica sequentially with 0.1 M sodium bicarbonate containing 0.5 M sodium chloride, lectin binding buffer and store at 4°C prior to use.

3.5.3. Packing Lectin Microcolumns

1. Cut 250 µm PEEK tubing to 5 cm length, wash thoroughly with acetone to remove possible contaminants and dry under a stream of nitrogen.
2. Place a PEEK union fitted with a 0.5 µm stainless steel frit at the bottom end of the PEEK tubing and connect the upper end to the Capillary Perfusion Kit packing apparatus.
3. Place the apparatus in a vertical position and introduce about 2 mL of a SNA-silica slurry into the reservoir. After filling the reservoir volume with the lectin-binding buffer, close the apparatus and connect it to the isocratic HPLC pump.
4. Start pumping the lectin-binding buffer into the reservoir at 300 µL/min and wait for the backpressure to reach 2,500 psi. Then, continue pumping at the same flow rate for an additional 10 min.
5. Stop the pump and allow the system to depressurize for at least 30 min. Carefully disconnect the apparatus from the pump and disconnect the packed PEEK tubing from the reservoir.
6. Check the packing and if the whole column volume is not filled with the stationary phase, repeat the procedure with a partially packed tubing. Otherwise, place the PEEK union fitted with a 0.5 µm stainless-steel frit to the other end of the tubing, wash the resulting microcolumn with lectin-binding buffer at 2 µL/min for 20 min and store at 4°C prior to use.

3.5.4. Glycopeptide Enrichment

1. Wash the lectin microcolumn with 10mM ammonium bicarbonate at 2 µL/min for 20 min and set the system, as depicted in **Fig. 6**.

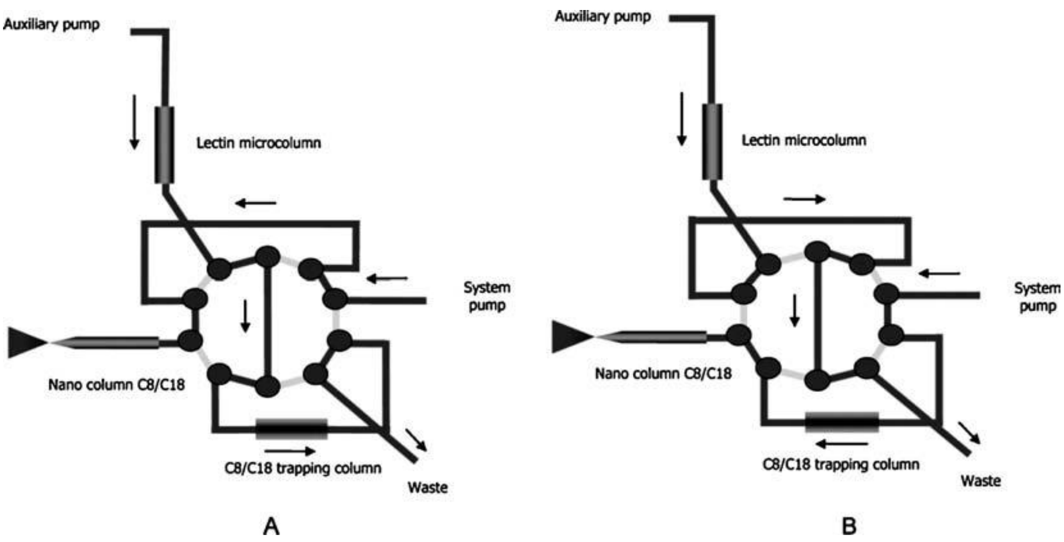

Auxiliary pump
Lectin microcolumn
System pump
Nano column C8/C18
C8/C18 trapping column
Waste
A
Auxiliary pump
Lectin microcolumn
System pump
Nano column C8/C18
C8/C18 trapping column
Waste
B

Fig. 6. High-performance affinity chromatography setup coupled online to electrospray mass spectrometry, including a system for loading of proteins/glycoproteins (**a**); and analysis of proteins/glycoproteins (**b**).

2. Set the flow rates of the isocratic and the system pumps to the following values: system pump: 3% of solvent A at 250 nL/min; isocratic pump: 10 mM ammonium bicarbonate at 2 μL/min.
3. Set up the data-dependent MS acquisition (DDA) method in such a way that the instrument will automatically isolate the seven most intense precursor ions, which will be subjected to tandem MS individually.
4. Make sure 10-port valve is in a loading position, where the injector and the lectin microcolumn are connected to the trapping cartridge. Inject 1 μL sample and wash the lectin microcolumn and the trapping cartridge for 15 min with 10 mM ammonium bicarbonate.
5. Switch the valve to the inject position (trapping cartridge is now connected to the system pump and capillary C_{18} column). If the analysis of the lectin-unbound fraction is to be performed, start the set-up MS method and separate the trapped peptides using a 50-min linear gradient from 3 to 55% acetonitrile.
6. After the 10-port valve is switched back to the loading position, displace fetuin glycopeptides bound to the lectin medium by injecting a 20-μL aliquot of the lectin elution buffer and washing the lectin microcolumn with the trap for 20 min using 10 mM ammonium bicarbonate (*see* Note 8).

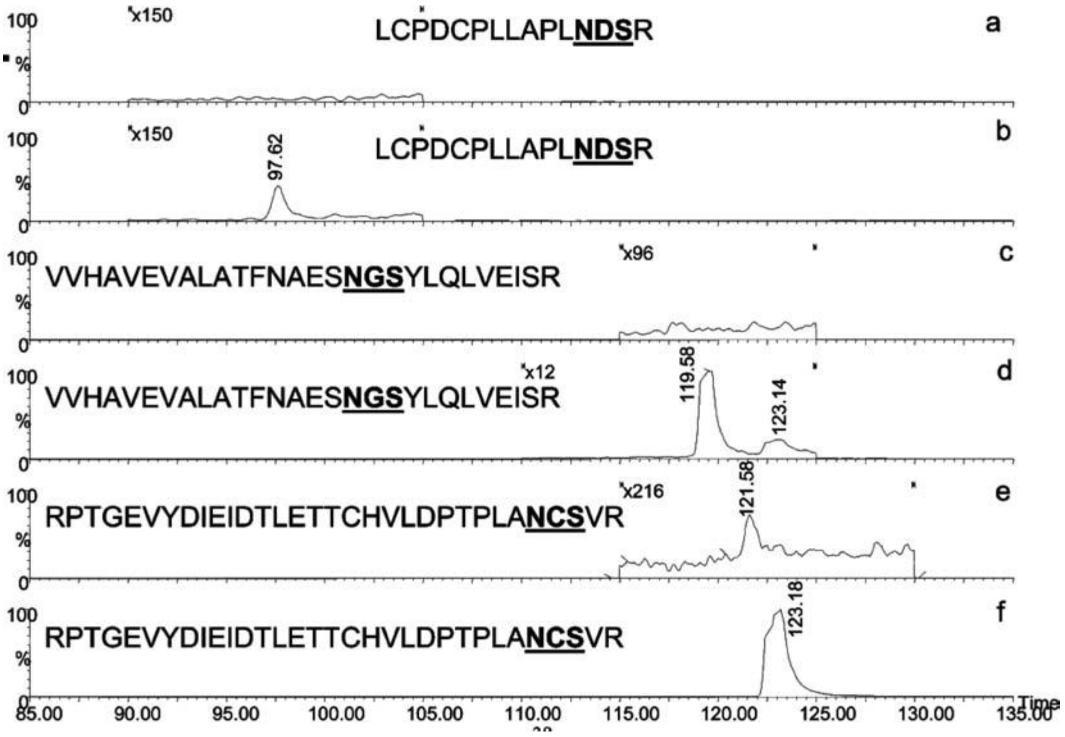


Fig. 7. Extracted ion chromatograms of fetuin glycopeptides analyzed by online SNA-lectin trapping and LC/MS. Lectin-unbound fractions (**a**, **c** and **e**) and lectin-bound fractions (**b**, **d** and **f**).

7. Separate the enriched glycopeptides using a linear gradient from 3 to 55% acetonitrile for over 50 min, while performing tandem MS experiment on the separated glycopeptides.

3.5.5. Data Processing

1. Open the file with the data collected from lectin-bound fraction. Check the MS/MS spectra and look for the characteristic glycan oxonium ions listed in **Table 2**. The enrichment of fetuin glycopeptides using silica-based SNA lectin material is depicted in **Fig. 7**.
2. Run GlycoMod software (http://www.expasy.org/tools/glycomod/) which allows the identification of a particular *N*-glycopeptide backbone sequence along with prediction of the structure of an attached glycan.

4. Notes

1. The LC/MS system described here consists of a Waters CapLC pump with an injector, 10-port valve, Q-TOF Ultima Global mass spectrometer fitted with a nanoESI source. However,

any LC/MS system could be used with the first described approach, including an LC/MS system with scanning capabilities such as QTOF or a triple-quadrupole.

2. Quenching of trypsin digestion is commonly achieved either by boiling the reaction mixture and denaturing trypsin or by acidifying the reaction mixture through the addition of formic acid, which renders the enzyme inactive.

3. PNGase F activity is inhibited in the presence of SDS; therefore, it is essential to have NP-40 (a non-ionic detergent which is known to counteract the action of SDS) present in the reaction mixture prior to the addition of PNGase F, if SDS is known to be present. Deglycosylation of a native glycoprotein requires a longer incubation time as well as more enzyme quantity. *N*-glycans with fucose α 1,3-linked to the asparagine-bound *N*-acetylglucosamine are resistant to the action of PNGase F.

4. In general, 10 mU of enzyme is sufficient to deglycosylate up to 100 μg denatured glycoprotein or 20 μg native glycoprotein in 18 h at pH 7.5 and 37°C. However, PNGase F is also effective in the release of glycans from glycopeptides using the same incubation buffer, as that employed for tryptic digestion (50 mM ammonium bicarbonate buffer, pH 8.0). This is due to the fact that the optimum pH range of PNGase is 7.5–8.6.

5. Hydrophobicity of eluting peptides and glycopeptides is mainly attributed by the peptide backbone. Therefore, the glycopeptides and unmodified peptides, as a result of *N*-deglycosylation of the same glycopeptide, will elute with very comparable retention times (*see* Fig. 1).

6. Phosphate or carbonate buffers with pH 8 are also suitable for the coupling reaction; however, amine containing buffers should be avoided since they may interfere with the coupling reaction.

7. The addition of sodium borohydride in steps is very crucial to ensure a complete conversion of all free aldehyde groups. Sodium borohydride degrades in the presence of water, so that its addition in steps will ensure that sufficient amount is reacting with aldehyde groups prior to losing its activity through its reaction with water.

8. Washing the lectin microcolumns with 10 mM ammonium bicarbonate during the analysis of peptides and enriched peptides has been determined not to adversely influence the binding capacity of the silica-based lectin materials. Moreover, it has been demonstrated recently that the binding capacity of the silica-based lectin does not change even when the microcolumns are washed with water for more than 2 days *(37)*.

Acknowledgments

This work was supported by grants No. GM24349 from the National Institutes of Health (NIH) and No. RR018942 from NCRR/NIH as a contribution from the National Center for Glycomics and Glycoproteomics at Indiana University.

References

1. Mackiewicz, A., Dewey, M. J., Berger, F. G., and Baumann, H. (1991) Acute phase mediated change in glycosylation of rat alpha-1-acid glycoprotein in transgenic mice. *Glycobiology* 1, 265–269
2. Maguire, T., and Breen, K. (1995) A decrease in neural sialyltransferase activity in Alzheimers disease. *Dementia* 6, 185–190
3. Maguire, T., Thakore, J., Dinan, T. G., Hopwood, S., and Breen, K. C. (1997) Plasma sialyltransferase levels in psychiatric disorders as a possible indicator of HPA axis function. *Biol. Psychiatry* 41, 1131–1136
4. Parekh, R., Isenberg, D., Rook, G., and Roitt, A. (1989) Comparative analysis of disease-associated changes in the galactosylation of serum IgG. *J. Autoimmun.* 2, 101–114
5. Parekh, R. B., Dwek, R. A., Sutton, B. J., Fernandes, D. L., Leung, A., Stanworth, D., and Rademacher, T. W. (1985) Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* 316, 452–457
6. Turner, G. (1995) Haptoglobin-A potential reporter molecule for glycosylation changes in disease. *Adv. Exp. Med. Biol.* 376, 231–238
7. Mechref, Y., and Novotny, M. V. (2002) Structural investigations of glycoconjugates at high sensitivity. *Chem. Rev.* 102, 321–370
8. Harvey, D. J. (2003) Identification of sites of glycosylation. *Methods in Molecular Biology* 211 (Protein Sequencing Protocols (2nd Edition), 371–383
9. Orlando, R., and Yang, Y. (1998) *in* "Mass Spectrometry of Biological Materials" (Larsen, B., and McEwen, C. N., Eds.), pp. 215–245, Dekker, New York, NY
10. Packer, N. H., and Harrison, M. J. (1998) Glycobiology and proteomics. Is mass spectrometry the holy grail?. *Electrophoresis* 19, 1872–1882
11. Alving, K., Korner, R., Paulsen, H., and Peter-Katalinic, J. (1998) Nanospray-ESI low-energy CID and MALDI post-source decay for determination of O-glycosylation sites in MUC4 peptides. *J. Mass Spectrom.* 33, 1124–1133
12. Iwase, H., Tanaka, A., Hiki, Y., Kokubo, T., Ishii-Karakasa, I., Hisatani, K., Kobayashi, Y., and Hotta, K. (1998) Application of matrix-assisted laser desorption ionization time-of-flight mass spectrometry to the analysis of glycopeptide-containing multiple O-linked oligosaccharides. *J. Chromatogr. B* 709, 145–149
13. Kuster, B., and Mann, M. (1999) ^{18}O-Labeling of *N*-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. *Anal. Chem.* 71, 1431–1440
14. Nemeth, J. F., Hochensang, G. P., Marnett, L. J., and Caprioli, R. M. (2001) Characterization of the glycosylation sites in cyclooxygenase-2 using mass spectrometry. *Biochemistry* 40, 3109–3116
15. Stults, N. L., and Cummings, R. D. (1993) O-linked fucose in glycoproteins from Chinese hamster ovary cells. *Glycobiology* 3, 589–596
16. Wolf, S. M., Ferrari, R. P., Traversa, S., and Biemann, K. (2000) Determination of the carbohydrate composition and the disulfide bond linkages of bovine lactoperoxidase by mass spectrometry. *J. Mass Spectrom.* 35, 210–217
17. Yang, Y., and Orlando, R. (1996) Identifying the glycosylation sites and site-specific carbohydrate heterogeneity of glycoproteins by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* 10, 932–936
18. Zhu, X. G., Borchers, C., Bienstock, R. J., and Tomer, K. B. (2000) *Biochemistry* 39, 11194–11204
19. Huddleston, M. J., Bean, M. F., and Carr, S. A. (1993) Collisional fragmentation of glycopeptides by electrospray ionization LC MS and LC MSMS- methods for selec-

tive detection of glycopeptides in protein digests. *Anal. Chem.* 65, 877–884

20. Wilm, M., Neubauer, G., and Mann, M. (1996) Parent ion scans of unseparated peptide mixtures. *Anal. Chem.* 68, 527–533
21. Medzihradszky, K. F., Maltby, D. A., Hall, S. C., Settineri, C. A., and Burlingame, A. L. (1994) Characterization of protein *N*-glycosylation by reversed-phase microbore liquid chromatography/electrospray mass spectrometry, complementary mobile phases, and sequential exoglycosidase digestion. *J. Am. Soc. Mass Spectrom.* 5, 350–358
22. Itoh, S., Kawasaki, N., Ohta, M., and Kayakawa, T. (2002) Structural analysis of a glycoprotein by liquid chromatography-mass spectrometry and liquid chromatography with tandem mass spectrometry application to recombinant human thrombomodulin. *J. Chromatogr. A* 978, 141–152
23. Ritchie, M. A., Gill, A. C., Deery, M. J., and Lilley, K. (2002) Precursor ion scanning for detection and structural characterization of heterogenous glycopeptide mixtures. *J. Am. Soc. Mass Spectrom.* 13, 1065–1077
24. Mechref, Y., Muzikar, J., and Novotny, M. V. (2005) Comprehensive assessment of *N*-glycans derived from a murine monoclonal antibody: a case for multimethodological approach. *Electrophoresis* 26, 2034–2046
25. Harazono, A., Kawasaki, N., Kawanishi, T., and Hayakawa, T. (2005) Site-specific glycosylation analysis of human apolipoprotein B100 using LC/ESI MS/MS. *Glycobiology* 15, 447–462
26. Dage, J. L., Ackermann, B., and Halsall, H. B. (1998) Site localization of sialyl LewisX antigen on 1-acid glycoprotein by high-performance liquid chromatography-electrospray mass spectrometry. *Glycobiology* 8, 755–760
27. Wang, F., Nakouzi, A., Angeletti, R. H., and Casadevall, A. (2003) Site-specific characterization of the *N*-linked oligosaccharides of a murine immunoglobulin M by high-performance liquid chromatography/electrospray mass spectrometry. *Anal. Biochem.* 314, 266–280
28. Wuhrer, M., Balog, C. I. A., Koeleman, C. A. M., Deelder, M. A., and Hokke, C. H. (2005) New features of site-specific horseradish peroxidase (HPR) glycosylation uncovered by nano-LC-MS with repeated ion-isolation/fragmentation cycles. *Biochim. Biophys Acta* 1723, 229–239
29. Wuhrer, M., Koeleman, C. A. M., Hokke, C. H., and Deelder, A. M. (2005) Protein glycosylation analyzed by normal-phase nano-liquid chromatography-mass spectrometry of glycopeptides. *Anal. Chem.* 77, 886–894
30. Larsen, M. R., Hojrup, P., and Roepstorff, P. (2005) Characterization of gel-separated glycoproteins using two-step proteolytic digestion combined with sequential microcolumns and mass spectrometry. *Mol. Cell. Proteomics* 4, 107–119
31. Tajiri, M., Yoshida, S., and Wada, Y. (2005) Differential analysis of site-specific glycans on plasma and cellular fibronectins: application of a hydrophilic affinity method for glycopeptide enrichment. *Glycobiology* 15, 1332–1340
32. Wada, Y., Tajiri, M., and Yoshida, S. (2004) Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* 76, 6560–6565
33. Hagglund, P., Bunkenborg, J., Elortza, F., Jensen Ole, N., and Roepstorff, P. (2004) A new strategy for identification of *N*-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J. Proteome Res.* 3, 556–566
34. Bunkenborg, J., Pilch, B. J., Podteleinikov, A. V., and Wisniewski, J. R. (2004) Screening for *N*-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* 4, 454–465
35. Fu, D., and Van Halbeek, H. (1992) *N*-Glycosylation site mapping of human serotransferrin by serial lectin affinity chromatography, fast atom bombardment-mass spectrometry, and proton nuclear magnetic resonance spectroscopy. *Anal. Biochem.* 206, 53–63
36. Garcia, R., Rodriguez, R., Montesino, R., Besada, V., Gonzalez, J., and Cremata, J. A. (1995) Concanavalin A- and wheat germ agglutinin-conjugated lectins as a tool for the identification of multiple *N*-glycosylation sites in heterologous protein expressed in yeast. *Anal. Biochem.* 231, 342–348
37. Madera, M., Mechref, Y., and Novotny, M.V. (2005) Combining lectin microcolumns with high-resolution separation techniques for enrichment of glycoproteins and glycopeptides. *Anal. Chem.* 77, 4081–4090

Chapter 10

Structure Analysis of *N*-Glycoproteins

Stefanie Henning, Jasna Peter-Katalinić, and Gottfried Pohlentz

Summary

General mass spectrometry-based strategies for analysis of *N*-glycosylated peptides are described. The well-established method utilizes Peptide-*N*-glycosidase F (*PNGase* F) for in-gel or in-solution release of *N*-linked glycans from the polypeptide chains (along with the conversion of the formerly *N*-glycosylated Asn to Asp), thus allowing separate analysis of glycan moieties and deglycosylated peptides. However, no assignment of individual glycans to a glycosylation site can be realized. Intact glycopeptides (i.e., proteolytic mixtures in which the glycan chains stay attached at their original glycosylation sites) can be analyzed either by a direct infusion or with HPLC separation prior to MALDI or ESI mass spectrometric analysis to provide both information on the glycan structure and glycosylation site in the same experiment. Several different strategies for efficient in-solution digestion of glycoproteins are described, such as proteolytic digestion in the electrospray capillary and simultaneous analysis of the resulting (glyco)peptides.

Key words: PNGase F glycan release, In-gel, In-solution, In-capillary, Proteolysis, MALDI MS, ESI MS, HPLC MS, Glycopeptides.

1. Introduction

Until a few decades ago, carbohydrates were thought to be rather boring components that more or less fulfill storage or scaffolding functions. Moreover, protein chemists considered them as a nuisance since they frequently hampered protein purification. Research over the last 50 years or so, however, demonstrated that glycosylation is one of the most common and diverse posttranslational modification of eukaryotic proteins. Presumably more than 50% of all proteins are glycosylated while their carbohydrate content varies from less than 1% to more than 90%, by weight *(1)*. A variety of biological functions have been proposed for the carbohydrate part of glycoproteins. The blood group antigens

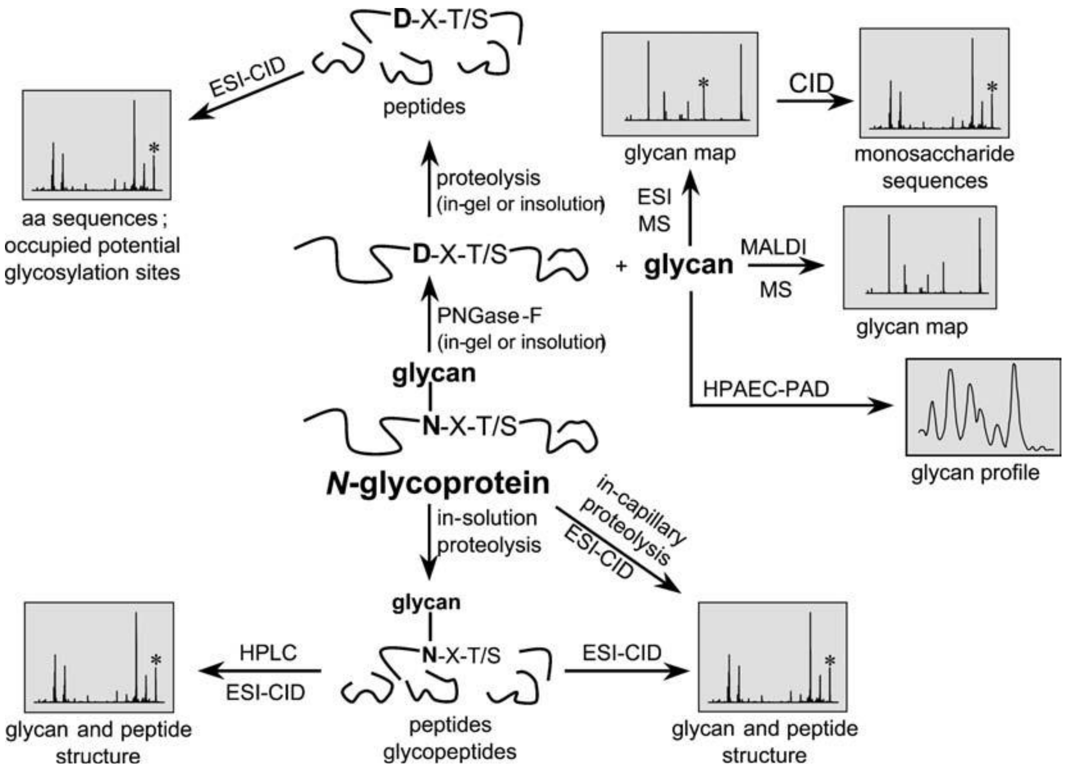
M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_10

are composed of oligosaccharides bound to proteins (or lipids) *(2)*; glycans are involved in the proper folding and the quality control of newly synthesized glycoproteins in the endoplasmic reticulum *(3, 4)* and control the serum lifetime of glycoproteins *(5, 6)*; *N*- and *O*-glycans play a role in the polarized biosynthetic sorting of proteins *(7)*; carbohydrates participate in the sperm–egg binding *(8)* just to mention a few.

Glycosylation is a highly complex and dynamic process, which frequently results in the biosynthesis of multiple glycoforms of the same protein by either incomplete glycosylation or processing of the oligosaccarides.

There are two main types of protein glycosylation. For the *N*-glycosylation, the polypeptide sequence has to contain the consensus sequence asparagine–X–serine/threonine (Asn–X–Ser/Thr, where X may be any amino acid except proline, Pro) in which the oligosaccharide is attached to the Asn residue via an *N*-glycosidic bond. *O*-glycosylation occurs at Ser or Thr residues. No general consensus sequence is known for *O*-glycosylation.

Structure analysis of glycoproteins to reveal the glycosylation sites and the corresponding glycoforms is still a challenge. The most common and some recently developed strategies are summarized in **Scheme 1**.


D-X-T/S
peptides
ESI-CID
*
aa sequences ;
occupied potential
glycosylation sites
proteolysis
(in-gel or insolution)
*
glycan map
CID
*
monosaccharide
sequences
ESI
MS
D-X-T/S
+ glycan
MALDI
MS
glycan map
PNGase-F
(in-gel or insolution)
glycan
N-X-T/S
HPAEC-PAD
glycan profile
N-glycoprotein
in-capillary
proteolysis
ESI-CID
in-solution
proteolysis
glycan
N-X-T/S
HPLC
ESI-CID
ESI-CID
*
glycan and peptide
structure
peptides
glycopeptides
*
glycan and peptide
structure

Scheme 1. Strategies for structure elucidation of *N*-glycosylated proteins.

Two procedures for structural analysis of glycoproteins by means of mass spectrometry are widely used. One involves the enzymatic or chemical release of the intact oligosaccharide prior to either in-gel or in-solution proteolytic digest. *N*-linked oligosaccharide moieties can be released as glycosylamines from the polypeptide chain using Peptide-*N*-glycosidase F (*PNGase* F) *(9–11)*. Simultaneously, the formerly *N*-glycosylated Asn residue will be converted to aspartic acid (Asp) by action of the enzyme which leads to a mass shift in the remaining peptide of roughly 1 Da.

Separate analyses of the carbohydrate and the protein portion by mass spectrometry enables the characterization of the overall glycan pattern and determination of the previously occupied glycosylation sites by virtue of the mass shift. However, employing this strategy, no assignment of an individual oligosaccharide to a certain glycosylation site can be made.

Another method of choice for analysis of glycoproteins is the separation of intact proteolytic glycopeptides by high performance liquid chromatography (HPLC) prior to mass spectrometric analysis *(12, 13)*. This technique provides information on the glycan structure as well as on the glycosylation site and is inevitable in the case when glycopeptides are not directly detectable in a proteolytic glycopeptide/peptide mixture (see below). Both methods are, however, usually laborious and time-consuming.

For structure elucidation of glycoproteins, the analysis of proteolytic glycopeptides is highly advantageous because determination of the glycan structure and the glycosylation site are accomplished in the same experiment. However, significant losses of proteolytic (glyco)-peptides during purification procedures following in-gel or in-solution digestion of (glyco)proteins frequently occur. In order to overcome this drawback and enhance structural information, a simpler and faster method for structure elucidation of glycoproteins based on the analysis of proteolytic glycopeptides, has been recently developed. This strategy relies either on the identification and/or sequencing of previously purified proteins or protein mixtures by proteolytic digest in the electrospray (ESI) capillary and simultaneous analysis of the resulting (glyco-)peptides or on the direct analysis of (glyco-)peptides derived from an in-solution proteolytic digest by nanoESI MS and MS/MS *(14, 15)*.

Since there is no enzyme available that can be universally applied to de-*O*-glycosylate proteins, β-elimination is still the only method for analysis of *O*-glycosides *(16, 17)*. Under reducing conditions, the reaction yields the oligosaccharides as sugar alcohols and results in the formation of unique amino acid residues derived from the formerly *O*-glycosylated Ser and Thr residues. Again, independent analyses of the released glycans and the deglycosylated protein and/or proteolytic peptides can be performed, but the entire glycan-protein linkage information will be lost. A detailed review on *O*-glycosylation of proteins and its analysis has been recently to published *(18)*. Therefore, the present article will focus on strategies for the MS-based structure elucidation of *N*-glycosylated proteins.

2. Materials

1. Glycoprotein standards (R*N*Ase B, transferrin, human IgG); Sigma (Deisenhofen, Germany).
2. Microcon® centrifugal filter devices and ZipTip® pipette tips; Millipore (Bedford, USA).
3. Micro Bio-Spin Chromatography Columns packed with Bio-Gel P-6; Bio-Rad Laboratories (München, Germany).
4. MALDI matrices: 2,5-dihydroxybenzoic acid (DHB) or α-cyano-4-hydroxycinnamic acid (CHCA).

2.1. In-Gel Procedures for N-Glycoproteins

1. *Solutions for preparation of protein bands*: 100 mM NH_4HCO_3; acetonitrile (ACN); 10 mM dithioerytritol (DTT) in 100 mM NH_4HCO_3; 50 mM iodoacetamide in 100 mM NH_4HCO_3; 100 mM $NaHCO_3$ pH 7.8.
2. *Solutions for release of glycans by peptide-N4-(N-acetyl-*β*-glucosaminyl)-asparagine amidase (PNGase F)*: PNGase F solution (200 U/mL) in 20 mM $NaHCO_3$ pH 7.8; 20 mM $NaHCO_3$ pH 7.8; acetonitrile (ACN); 50% ACN.
3. *Desalting solutions*: 50% ACN/0.1% v/v trifluoroacetic acid (TFA); 80% ACN/0.1% TFA; 0.1% TFA.
4. *Solutions for in-gel proteolytic digestion*: 100 mM NH_4HCO_3; 50 mM NH_4HCO_3; 20 mM NH_4HCO_3; protease solution: trypsin, chymotrypsin and/or Glu-C; 25 ng/μL in 50 mM NH_4HCO_3; 50% ACN/5% formic acid (FA); ACN.
5. *Solutions for purification of the extracted proteolytic peptides*:

Solution	Composition	Preparation
	2.5% trifluoroacetic acid (TFA)	25 μL TFA + 975 μL deionized water
Wetting solution	50% ACN	500 μL ACN + 500 μL deionized water
Equilibration solution	0.1% TFA	40 μL 2.5% TFA + 960 μL deionized water
Wash solution	0.1% TFA	40 μL 2.5% TFA + 960 μL deionized water
Elution solution 1	50% ACN/0.1% TFA	40 μL 2.5% TFA + 500 μL ACN + 460 μL deionized water
Elution solution 2	80% ACN/0.1% TFA	40 μL 2.5% TFA + 800 μL ACN + 160 μL deionized water
Elution solution 3	pure ACN	

2.2. In-Solution Procedures for N-Glycoproteins

1. *Solutions used to prepare proteins for digestion*: glycoprotein solution (20–100 pmol/μL); 100 mM NH_4HCO_3; 10 mM DTT; 50 mM iodoacetamide; 50 mM $NaHCO_3$ pH 7.8.
2. *PNGase F solution*: PNGase F (200 U/mL) in 20 mM $NaHCO_3$ pH 7.8.
3. 20 mM $NaHCO_3$ pH 7.8.

2.3. In-Capillary Procedures for N-Glycopeptides

1. Micro Bio-Spin Chromatography Columns packed with Bio-Gel P-6 (Bio-Rad Laboratories, München, Germany).
2. *Solutions for SEC*: deionized water and sample protein solution (20–75 μL).
3. Microcon Centrifugal Filter Devices with molecular weight cut-offs (MWCOs) varying from 30 to 100 kDa.
4. *Solutions for ultrafiltration*: deionized water and sample protein solution (50–500 μL).
5. *Solutions for in-capillary proteolytic digestion*: sample protein solution (desalted), 20–100 pmol/μL; ammoniumhydrogen carbonate (NH_4HCO_3), 100 mM; dithiothreitol (DTT), 20 mM; methanol; protease (trypsin, chymotrypsin, Glu-C), 0.1 μg/μL; deionized water.

2.4. In-Solution Procedures for N-Glycopeptides

1. Sample protein solution (desalted), 20–100 pmol/μL.
2. Ammoniumhydrogen carbonate (NH_4HCO_3), 100 mM.
3. Dithiothreitol (DTT), 20 mM.
4. Deionized water.
5. Protease (trypsin, chymotrypsin, Glu-C), 0.1 μg/μL.
6. Methanol.
7. Formic acid (FA; 99.9%, p.a.)

2.5. HPLC-MS Analysis of N-Glycopeptides

1. *NanoHPLC*: UltiMate (LC Packings, Dionex).
2. *Column*: NAN75-15-03-C18PN (LC Packings, Dionex).
3. *Gradient*: 99% A to 50% A; 2h; solvent A: 5% acetonitrile/0.05% formic acid; solvent B: 80% acetonitrile/0.04% formic acid.
4. *Mass spectrometer*: quadrupole time-of-flight (Q-TOF) nanoESI mass spectrometer (Waters/Micromass, Manchester, UK).

3. Methods

3.1. Mass Spectrometry

Nanoelectrospray ionization mass spectrometry (nanoESI MS) experiments were carried out using a quadrupole time-of-flight (Q-TOF) mass spectrometer (Waters/Micromass, Manchester, UK) in the positive ion mode (ESI(+)). A Z-spray atmospheric

3.1.1. Nanoelectrospray Ionization Mass Spectrometry

pressure ionisation (API) source was used with the source temperature set to 80°C and a desolvation gas (N_2) flow rate of 75 L/h. Home made nanospray capillaries were used, the capillary tip was set to a potential of 1.1 kV and the cone voltage was 40 V.

For low energy collision-induced dissociation (CID) experiments peptide precursor ions were selected in the quadrupole analyzer and fragmented in the collision cell using a collision gas (Ar) pressure of 3.0×10^{-5} mbar and collision energies of 20–40 eV (E_{lab}).

3.1.2. Matrix Assisted Laser Desorption/Ionization Mass Spectrometry

Matrix assisted laser desorption/ionization mass spectrometry (MALDI MS) was performed using a REFLEX III MALDI mass spectrometer (Bruker Daltonik, Bremen, Germany) in the reflectron mode. An acceleration voltage of 25 kV was applied and the reflectron voltage was set to 28.5 kV.

3.2. Structure Elucidation of N-Glycoproteins by Separate Analysis of N-Glycans and the Deglycosylated Protein

The first part of the methodological section of the present article deals with "classical" strategies of *N*-glycoprotein analysis. In-gel or in-solution release of *N*-glycans, in-gel or in-solution proteolytic digest of the deglycosylated protein, and subsequent separate analysis of glycans and peptides by means of mass spectrometry have been successfully employed over the last two decades.

3.2.1. N-Glycoprotein Analysis via In-Gel Procedures

The prerequisite for these investigations is the separation of (glyco-)proteins using one-dimensional or two-dimensional sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE). The protein bands or spots have been stained either with Coomassie® brilliant blue or silver.

Preparation of Protein Bands

1. Excise protein bands or spots and transfer to a reaction vial (1.5 mL, Eppendorf).
2. Add 250 μL of 100 mM NH_4HCO_3; shake the mixture for 30 min at ambient temperature.
3. Discard the supernatant and add 500 μL of acetonitrile; shake the mixture for 30 min at ambient temperature.
4. Discard the supernatant and add 250 μL of 100 mM NH_4HCO_3; shake the mixture for 30 min at ambient temperature.
5. Discard the supernatant; for the reduction step, the gel pieces are covered with 10 mM DTT in 100 mM NH_4HCO_3. The mixture is shaken for 1 h at 56°C, cooled to ambient temperature and the supernatant is discarded.
6. For alkylation, the gel pieces are then incubated with 50 mM iodoacetamide in 100 mM NH_4HCO_3 for 1 h at ambient temperature under exclusion of light.
7. Finally, in order to remove the excess of alkylation reagent and rebuffer, the gel pieces are shaken consecutively with (1) 250 μL of 100 mM $NaHCO_3$ pH 7.8; (2) acetonitrile; (3) 100 mM $NaHCO_3$; and (4) acetonitrile for 30 min at ambient temperature with in between disposal of the respective supernatants.

8. The gel pieces are dried in vacuo (generally a 10 min run in a "speed-vac" should be sufficient).

Release of Glycans by PNgase F

1. The in-gel digestion procedure follows the method described by Kuster et al. *(19)*.
2. Add 10–50 µL of a *PNGase* F solution (200 U/mL) in 20 mM $NaHCO_3$ pH 7.8 to the prepared gel pieces and incubate for 60 min at 37°C.
3. Cover the gel pieces with an additional 20 mM $NaHCO_3$, and incubate overnight at 37°C.
4. Glycans were extracted from the gel pieces by removing the incubation buffer followed by two extractions with 200 µL pure water, two extractions with 200 µL 50% ACN, and a final extraction with 200 µL pure ACN. Perform each extraction by sonication for 30 min; combine all extracts and dry in vacuo.

Desalting of Released Glycans

The desalting of released glycans is performed according to the method described by Packer et al. *(20)* using graphitised carbon cartridges.

1. Prepare micro cartridges by dismantling a normal-sized GlycoClean H cartridge (Oxford GlycoSciences) and resuspending the packing in water.
2. Using this suspension, microcolumns with approximately 2–5 µL material are made by filling an appropriate amount of the slurry into an Eppendorf pipette tip (0.5–10 µL) that has been previously plugged with a small piece of glass microfiber filter (Whatman, Brentford, UK) and by pressing out the water using disposable 5 mL syringe.
3. Prior to the use, wash the columns with three changes of 30 µL of water, two changes of 30 µL of 80% ACN/0.1% v/v TFA followed by 3 × 50 µL of water.
4. Dissolve the sample in 20–50 µL water and apply to the column. Salts are washed off with 50 µL of water followed by 3 × 50 µL of 0.1% TFA, and the glycans are eluted by the subsequent application of 3 × 50 µL of 50% ACN/0.1% TFA and 1 × 50 µL of 80% ACN/0.1% TFA. The eluates are combined and dried in vacuo.

In-Gel Proteolytic Digestion of Deglycosylated Proteins

1. For structure analysis of the remaining deglycosylated proteins, the gel pieces are dried in vacuo and then prepared for the proteolytic digest.
2. Rebuffering is achieved by subsequent addition of 250 µL of 100 mM NH_4HCO_3, 250 µL of ACN, 250 µL of 100 mM NH_4HCO_3, and 250 µL of ACN with in between shaking for 30 min at ambient temperature and disposal of the respective supernatants. Then, the gel pieces are dried in vacuo.

3. Add 30 µL of the protease solution and incubate for 30 min at 37°C.
4. Apply additional 50 mM NH_4HCO_3 to cover the gel pieces completely.
5. After incubation for 16 h at 37°C, the incubation buffer is transferred to a collecting vial. The proteolytic peptides are extracted by subsequent addition of 250 µL of 20 mM NH_4HCO_3, 3 × 250 µL of 50% ACN/5% formic acid (FA), 3 × 250 µL of 90% ACN/5% FA, and 250 µL of pure ACN. In between the sample is shaken for 30 min at ambient temperature and the respective supernatant is transferred to the collecting vial. The combined supernatants are dried in vacuo.

Purification of the Extracted Proteolytic Peptides

Peptide purification is preferentially performed by reversed-phase separation using C18 ZipTip® pipette tips (Millipore, Bedford, USA). The procedure follows the manufacturer's instructions, as is briefly described below.

1. Dissolve the sample in 10 µL of 0.5% TFA and shake vigorously.
2. Mount the ZipTip® pipette tip onto a 10 µL Eppendorf pipette. For equilibration 3 × 10 µL of the wetting solution and subsequently 3 × 10 µL of equilibration solution are aspirated and dispensed to waste.
3. Load the peptides by aspirating and dispensing the sample solution ten times.
4. Salts are removed by aspirating 3 × 10 µL of wash solution and dispensing to waste.
5. Finally, peptides are eluted by subsequently aspirating 5 × 10 µL of elution solution 1, 5 × 10 µL of elution solution 2, and 3 × 10 µL of elution solution 3 and dispensing the respective elution solutions to a collection vial. The combined eluates are dried in vacuo.

As an example, the MALDI mass spectrum obtained from the in-gel *PNGase* F-released oligosaccharides derived from the α3 chain of laminin 5 *(11)* is shown in **Fig. 1**. All detectable peaks represent singly charged sodiated molecular ions ($[M+Na]^+$) and are tentatively assigned according to their *m/z* values.

By subsequent in-gel tryptic digestion of the deglycosylated α3 chain, a proteolytic mixture containing the formerly glycosylated peptides were characterized. Due to the *PNGase* F reaction, these peptides contained aspartic acid (D) instead of asparagine (N) at the *N*-glycosylation sites. The fragment ion spectrum derived from a CID experiment utilizing the doubly charged precursor ion ($[M+2H]^{2+}$) with *m/z* 638.43 is depicted in **Fig. 2**. The increment $m/z_{y6}-m/z_{y5} = 115.04$ clearly proves the presence of a D instead of the N in position 245, thus indicating the former *N*-glycosylation of N_{245}.

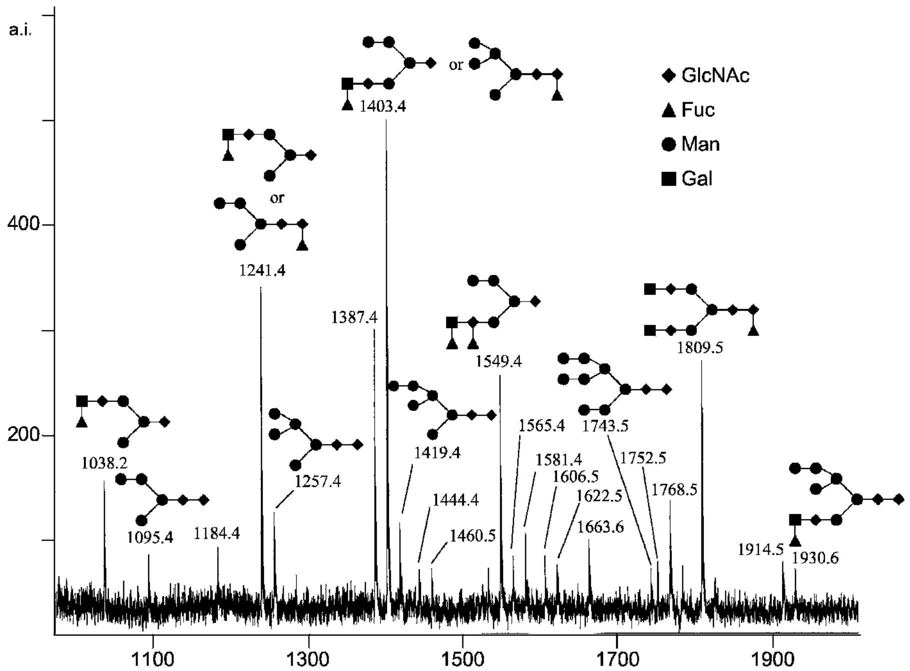


Fig. 1. MALDI mass spectrum of PNGase F released *N*-glycans derived from the α3 chain of laminin 5.

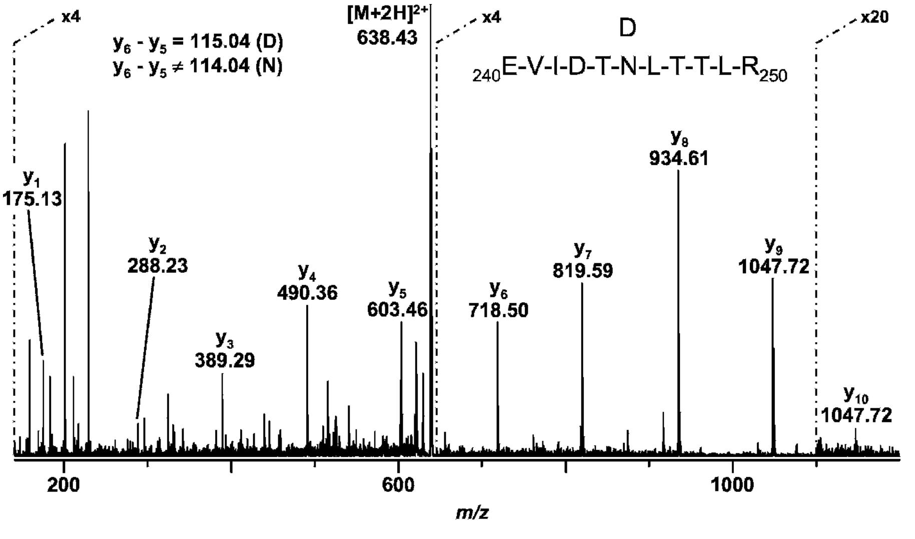


Fig. 2. NanoESI CID spectrum of the doubly charged precursor ions at *m/z* 638.43 derived from a tryptic digest of the deglycosylated α3 chain of laminin 5.

3.2.2. N-Glycoprotein Analysis via In-Solution Procedures

This strategy is preferably applied to purified proteins (e.g., recombinant proteins) or in cases when a general *N*-glycan pattern derived from a protein mixture (e.g., plasma proteins) has to be determined.

1. Add 2 µL of 100 mM NH_4HCO_3 and 10 µL of 10 mM DTT to 2–5 µL of the glycoprotein solution (20–100 pmol/µL) and adjust to 20 µL with deionized water.
2. The sample is mixed and incubated at 56°C for 1 h.
3. When the sample is cooled to ambient temperature, add 10 µL of 50 mM iodoacetamide and shake the mixture under exclusion of light for 1 h.
4. To remove the excess of reduction/alkylation reagent and to rebuffer, the sample is treated as described in **Subheading "Size Exclusion Chromatography"** (see below) except for the spin column being equilibrated in 20 mM $NaHCO_3$ pH 7.8. The spin column eluate contains the reduced and alkylated glycoprotein ready for PNGase F treatment.

N-Glycan Release by PNgase F

1. Mix 5–10 µL of the glycoprotein solution (depending on the starting concentration of the protein) with 5 µL of PNGase F solution and adjust to 20 µL with 20 mM $NaHCO_3$ pH 7.8. After mixing, the sample is incubated at 37°C for 16 h.
2. For the isolation of the liberated *N*-glycans, one of the following procedures may be performed:
   (a) *Precipitation of the deglycosylated protein*: The deglycosylated protein is precipitated by addition of four volumes of ethanol at 0°C, incubated for 2 h at this temperature, and separated by centrifugation (30 min in an Eppendorf centrifuge at full speed). Evaporate the supernatant containing the *N*-glycans in vacuo. Desalt the *N*-glycan sample on graphitized carbon as described in the **Subheading "Desalting of Released Glycans"** (*See* **Note 1**).
   (b) *Separation of the released N-glycans by ultrafiltration*: The procedure described under **Subheading "Ultrafiltration"** is used with the following exceptions (1) the first two filtrates are combined (they should contain >90% of the released *N*-glycans), evaporated in vacuo, and desalted on graphitized carbon as described above and (2) in the last ultrafiltration run the retenate containing the deglycosylated protein is concentrated to 10 µL, according to the manufacturer's instructions.

3.3. Structure Elucidation of N-Glycoproteins by Analysis of Proteolytic N-Glycopeptides

3.3.1. In-Capillary Proteolytic Digest

By far the fastest method for analyzing (glyco-)proteins is the proteolytic digestion in the ESI capillary and simultaneous analysis of the resulting (glyco-)peptides since the first results may have been already obtained after 30 min *(14)*. However, the proteins under investigation have to be essentially salt-free. Dialysis is a very effective desalting method, but it is extremely time consuming (for sufficient desalting a dialysis time of at least 48 h is necessary) and typically accompanied with significant sample loss (due to the adhesion of (glyco-)proteins to the dialysis membrane). The alternative methods for desalting of (glyco-)protein preparations, size exclusion chromatography (SEC) on Bio-Gel P-6 spin columns and ultrafiltration by use of Microcon Centrifugal Filter Devices, are both faster and show better sample recovery (*see* **Note 2**).

Size Exclusion Chromatography

According to the manufacturer's instructions for buffer exchange, desalting is performed as follows:

1. The settled gel is resuspended by sharply inverting the column several times.
2. After removal of the remaining air bubbles by tapping the column, the tip is snapped off and the column is placed in a 2 mL microcentrifuge tube. Then the cap is removed and the excess of packing buffer is drained by gravity (contingently, the cap has to be pushed back on the column and removed again to start the flow).
3. The drained buffer is discarded, the column is placed back into the tube and centrifuged in a microcentrifuge (e.g., Centrifuge 5415 C, Eppendorf, Hamburg, Germany) for 2 min at 1,000 × *g*. The flow-through is discarded, deionized water is applied in four aliquots of 500 μL each. After each application the water is drained either by gravity or by centrifuging for 1 min at 1,000 × *g* and the eluted water is discarded.
4. The column is placed in a new 1.5 mL reaction tube and the sample protein solution (20–75 μL) is carefully applied onto the top center of the gel bed. The assembly is centrifuged for 4 min at 1,000 × *g*. The eluate should now contain the desalted protein.

Ultrafiltration

Desalting by ultrafiltration was carried out using Microcon Centrifugal Filter Devices with molecular weight cut-offs (MWCOs) varying from 30 to 100 kDa and depending on the size of the (glyco-)protein under investigation. The manufacturer's instructions describe the concentration of a protein solution from 500 μL down to 10 μL. However, in our hands a 50-fold concentration leads to an unacceptable loss of protein (probably by adhesion to the filter membrane). Therefore, in the following protocol a repeated concentration from 500 to 100 μL, is used to reduce the sample loss.

1. The (glyco-)protein solution is transferred to the filter device and, if necessary, adjusted to a total volume of 500 μL by addition of deionized water.
2. After mixing by gently inverting, the device is placed into a microcentrifuge and spun for 7 min at 12,000 × *g* at ambient temperature.

3. The filtrate is discarded and the following steps are repeated four times (1) replenish the retenate to 500 μL with deionized water; (2) mix by gentle inversion of the assembly; (3) centrifuge for 6 min at 12,000 × *g* at ambient temperature; and (4) discard the filtrate. After the last run, the retenate is isolated by placing the filter device upside down into a new 1.5 mL reaction tube and centrifuging the assembly for 30 s at 1,000 × *g*.

In-Capillary Proteolytic Digestion

(a) *Digestion without reduction* (*see* **Note 3**): Mix 2–5 μL of the sample protein solution (the final concentration in a 20 μL digest mixture should be 5–10 pmol/μL) with 2 μL of 100 mM NH_4HCO_3, 2 μL of methanol, 1–4 μL protease solution and adjusted to 20 μL with deionized water. After mixing the solution and centrifugation for 1 min in an Eppendorf centrifuge at maximum power, an aliquot (5–10 μL) is immediately transferred to the ESI capillary and MS data acquisition is started. The (glyco-)peptide molecular ions were monitored over a time period of 45 min to 2 h.

(b) *Digestion with reduction* (*see* **Note 4**): Add 2 μL of 100 mM NH_4HCO_3, 2 μL of 20 mM DTT and the appropriate volume of deionized water to 2–5 μL of the sample protein solution (for a total volume of 20 μL). Subsequently, the mixture was shaken either for 2 h at ambient temperature or for 5 min at 95°C. After addition of the protease solution, mixing, and centrifugation as described above, an aliquot (5–10 μL) is immediately transferred to the ESI capillary and MS data acquisition is started.

Taking into account the fact that glycans are frequently located on the surface of glycoproteins and that protease cleavage sites around these glycosylation sites are accessible to the enzymes, it may be useful to run in-capillary digestion using virtually native proteins i.e., without reductive cleavage of the disulfide bonds (procedure a). Since only accessible sites can be utilized by the proteases and a large portion of the protein molecules remain uncleaved this strategy may yield a relatively high abundance of glycopeptides compared to the "naked" peptides. As an example the nanoESI mass spectrum obtained during an in-capillary tryptic digestion of ribonuclease B is shown in **Fig. 3**. A number of glycopeptide ions were detected. Besides proteolytic (glyco-)peptide ions in the early phase of the in-capillary digestion, molecular ions derived from the intact protein ($[M+7H]^{7+}$ and $[M+8H]^{8+}$), which also reflect the glycosylation pattern, were found. In the course of the analysis the relative intensity of these intact protein ion species decreased with corresponding increase in the intensity of (glyco-)peptide ions. This phenomenon is frequently observed during in-capillary digestion of smaller proteins *(14)*. The sequence coverage provided by the detected tryptic peptides

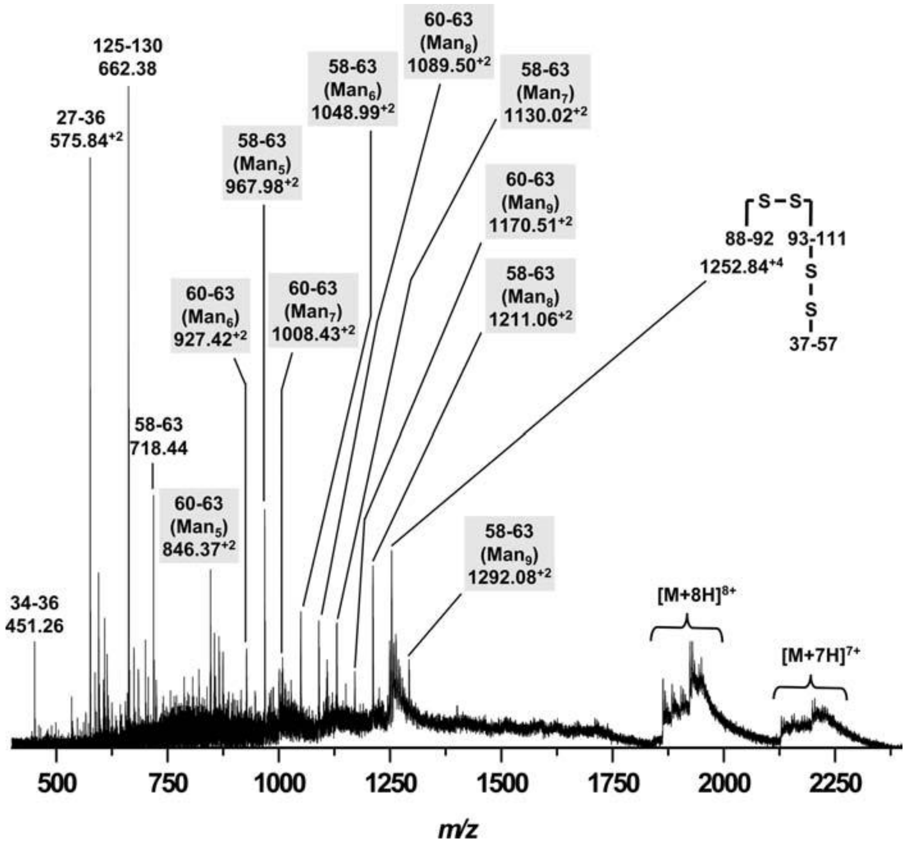


Fig. 3. NanoESI Q-TOF mass spectrum obtained from an in-capillary tryptic digest of ribonuclease B (RNAse B/trypsin 25:1; adapted from (*14*)).

was as low as 53%. Incubation with dithiothreitol prior to in-capillary digestion (procedure b) led to a much faster degradation of the protein and the relative abundances of glycopeptide ions decreased in favor of non-glycosylated peptide ions. Moreover, the sequence coverage under these conditions was 100% (data not shown) (*see* **Note 5**).

3.3.2. In-Solution Proteolytic Digest

Glycopeptides containing acidic residues, e.g., NeuAc, fail to ionize under in-capillary digestion conditions due to the relatively high pH (7.4–7.8). For these cases, an in-solution proteolytic digest with subsequent addition of methanol and formic acid and direct nanoESI MS analysis is the method of choice.

1. Mix 2–5 μL of the sample protein solution (the final concentration in a 20 μL digest mixture should be 5–10 pmol/μL) with 2 μL of 100 mM NH_4HCO_3, 2 μL of 20 mM DTT and the appropriate volume of deionized water (total volume: 20 μL including protease solution).

2. Shake the mixture either for 2 h at ambient temperature or for 5 min at 95°C.
3. After addition of 1– 4 µL of the protease solution and shaking, the mixture is incubated for 2–16 h at 37°C.
4. An aliquot of the mixture is adjusted to 25% methanol and 10% FA (e.g., 6.5 µL reaction mixture plus 2.5 µL methanol plus 1 µL FA). After mixing the solution and centrifugation for 1 min in an Eppendorf centrifuge at maximum power, an aliquot (5–10 µL) is transferred to the ESI capillary and data acquisition is started.

In **Fig. 4** the nanoESI mass spectrum obtained by in-solution digestion of human transferrin (TRFE) is shown. In order to obtain smaller (glyco-)peptides, a 1:1 mixture of trypsin and chymotrypsin was used. Again, the main advantage of a direct analysis of the (glyco-)protein digest, i.e., high structure information content, is impressively demonstrated. Besides single amino acids (K, *m/z* 147.11; F, *m/z* 166.10; R, *m/z* 175.13) and dipeptides (e.g., $_{679}$RP$_{698}$, *m/z* 272.18), a large number of non-glycosylated peptides of diverse chain lengths (from 3 to >20 amino acids) are detectable as protonated molecular ions in various charge states

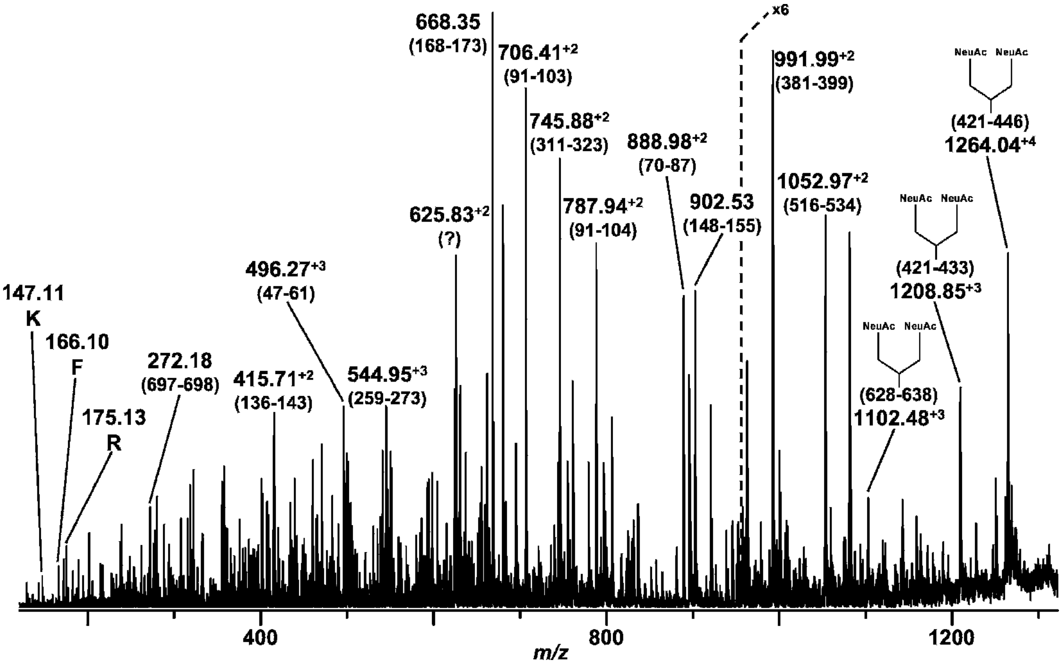


Fig. 4. NanoESI Q-TOF mass spectrum obtained from an in-solution tryptic/chymotryptic digest of human transferrin.

($[M+nH]^{n+}$), covering more than 82% of the TRFE sequence. Moreover, sizable glycopeptide ions derived from peptides containing one of both *N*-glycosylation sites (N_{432}, N_{630}) can be observed (e.g., $[M+3H]^{3+}$, *m/z* 1,102.48; ($[M+3H]^{3+}$, *m/z* 1,208.85; ($[M+4H]^{4+}$, *m/z* 1,264.04).

The glycopeptide structures can be characterized using CID experiments. The fragment ion spectrum derived from the quadruply charged precursor ions with *m/z* 1,264.04 and the corresponding fragmentation scheme are shown in **Fig. 5**. The almost complete series of Y- and B-type ions (nomenclature of the glycan fragments is according to Domon and Costello *(20)*, cf. **Fig. 6**), in addition to fragment ions derived from double cleavages, allowed the unambiguous elucidation of the glycan structure (i.e., a complex type disialylated biantennary oligosaccharide). A few low, abundant, but clearly detectable, b and y ions originating from the peptide backbone enabled the identification of the peptide moiety as $aa_{421-446}$ containing the *N*-glycosylation site N_{432}.

3.4. HPLC-MS Analysis of N-Glycopeptides

Glycopeptide derived ions may not be detected by direct mass spectrometric analysis of glycoprotein digests either due to high glycan heterogeneity at a given *N*-glycosylation site or because of the dynamic range limitations (e.g., low relative abundance of the proteolytic glycopeptides compared to non-glycosylated peptides are observed for human IgG preparations). In these cases, HPLC-MS analysis of the proteolytic *N*-glycopeptides is the method of choice.

1. The HPLC system is coupled to the ESI mass spectrometer via commercially available or home-made interfaces.
2. An aliquot of an in-solution digest (see above) is diluted two to fivefold (depending on the initial protein concentration) with solvent A and 1 μL of the mixture is injected into the HPLC system.
3. HPLC gradient and mass spectrometric data acquisition are started simultaneously.

A total ion chromatogram obtained from an HPLC-MS experiment using a tryptic/chymotryptic digest derived from a human IgG preparation is shown in **Fig. 7a**. Extracted ion chromatograms scanning for characteristic in-source fragments of glycans such as Gal-GlcNAc$^+$ at *m/z* 366.14, provide preliminary information on glycopeptide-containing peaks (not shown). Combining the respective scans, mass spectra showing, amongst others, glycopeptide derived ions (e.g., **Fig. 7b**) were obtained. The glycopeptide ion peaks can be tentatively assigned according to their *m/z* values (cf. **Table 1**) or their structures may be corroborated by subsequent HPLC-MS/MS fragmentation experiments.

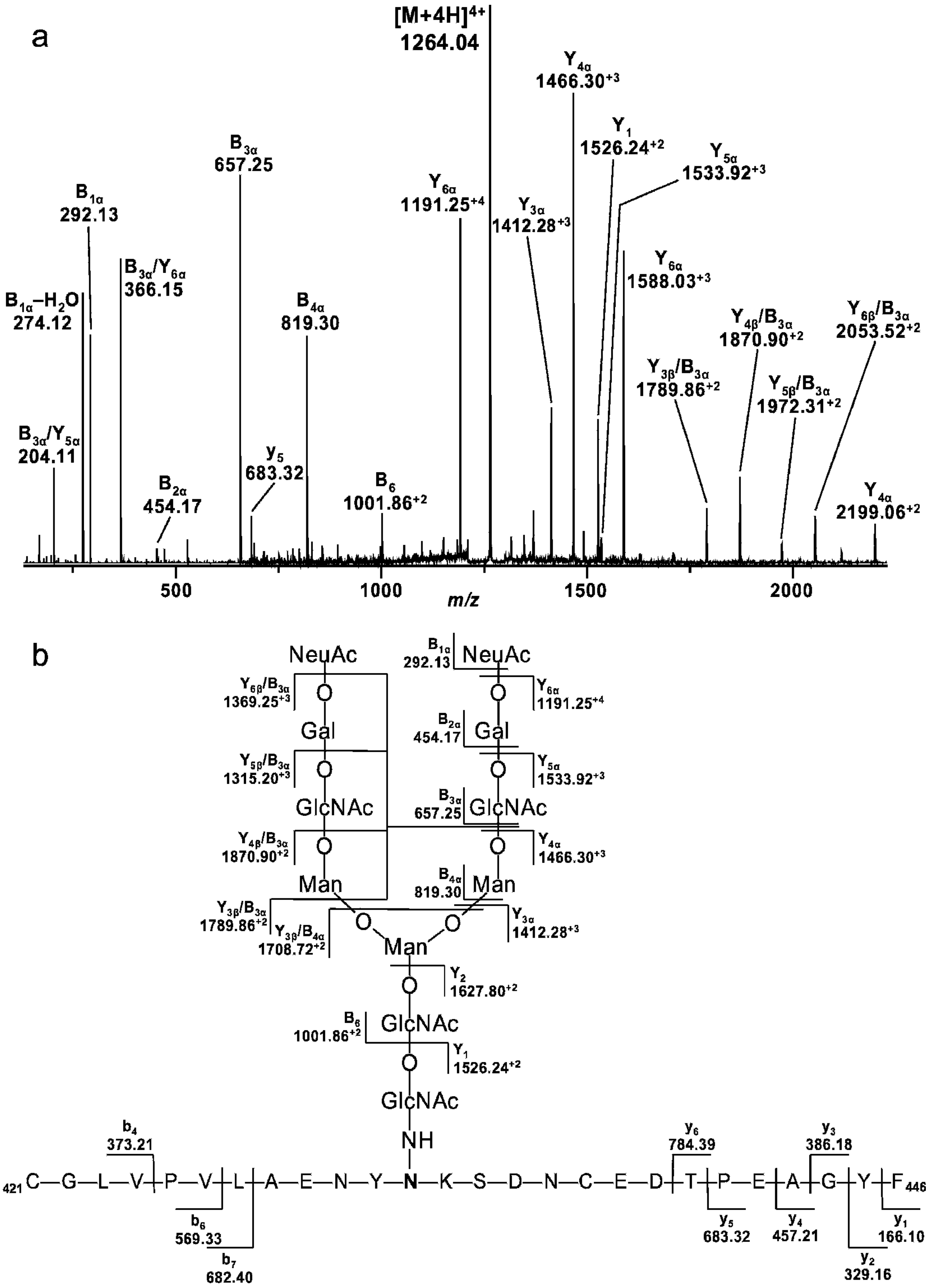


Fig. 5. **a** NanoESI Q-TOF fragment ion spectrum obtained from a CID experiment on the quadruply charged precursor ions at *m/z* 1,264.04 derived from an in-solution tryptic/chymotryptic digest of human transferrin. **b** Fragmentation scheme deduced from **a**).

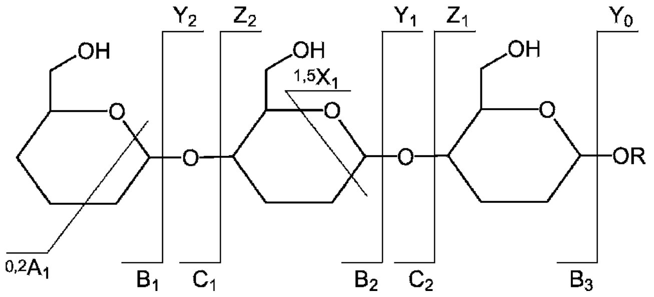


Fig. 6. Nomenclature of fragment ions derived from glycan moieties according to Domon and Costello (*21*).

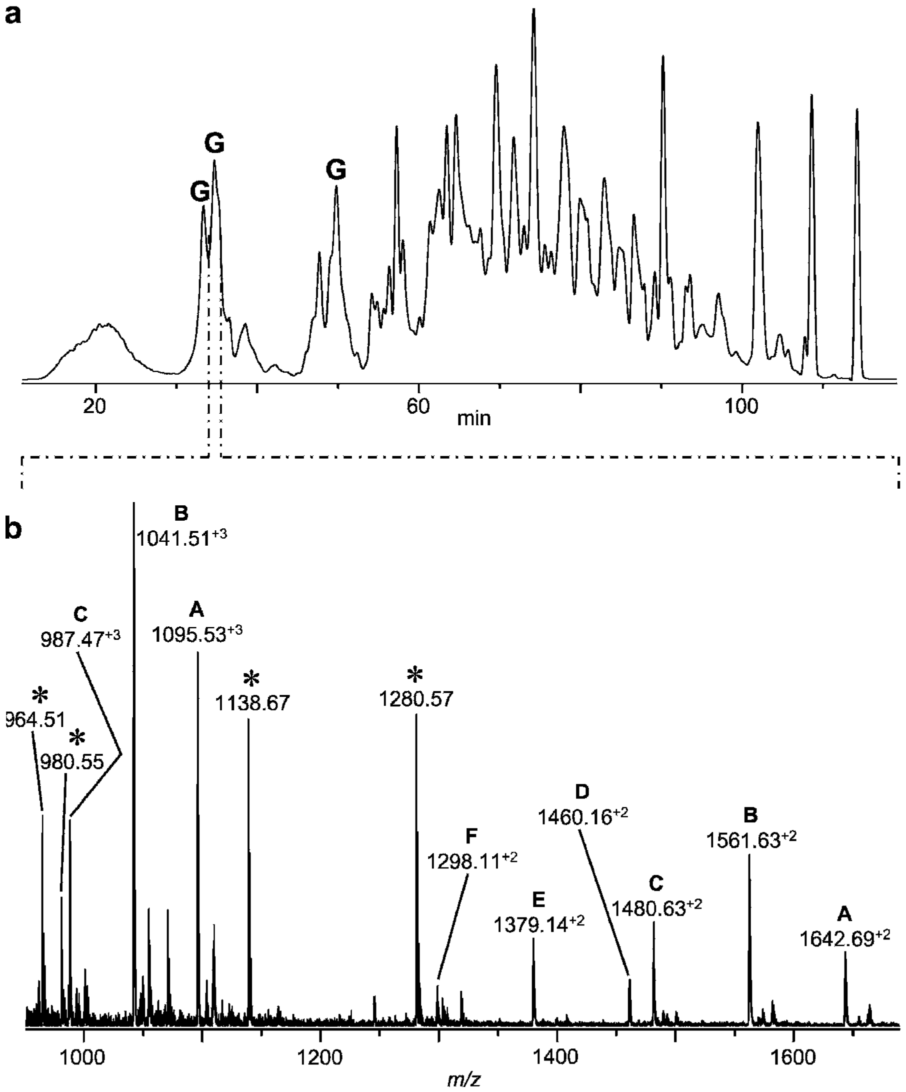
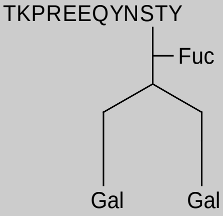
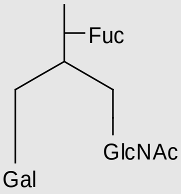
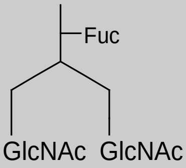
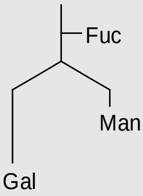
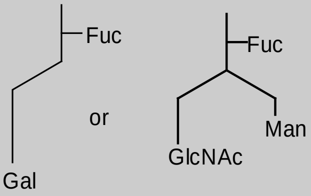
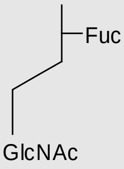


Fig. 7. **a** Total ion chromatogram obtained from an HPLC-MS of a tryptic/chymotryptic digest of a human IgG preparation. Peaks labeled with **G** contain glycopeptide ions. **b** NanoESI mass spectrum obtained from the peak at min 34.7. Peaks marked with an asterisk represent singly charged ions derived from non-glycosylated peptides. The structures of the glycopeptides A to F are listed in **Table1**.

Table 1
Glycopeptide structures deduced from the HPLC mass spectrum of the chromatogram peak at min 34.7 shown in Fig.7

Oligosaccharide	*M/z*	Glycan structure
A	1,642.69^{+2} ≡ 3,284.37^{+1} (calc.: 3,284.35) (1,095.53^{+3})	 TKPREEQYNSTY Fuc Gal Gal
B	1,561.63^{+2} ≡ 3,122.25^{+1} (calc.: 3,122.30) (1,041.51^{+3})	 Fuc GlcNAc Gal
C	1,480.63^{+2} ≡ 2,960.25^{+1} (calc.: 2,960.24) (987.47^{+3})	 Fuc GlcNAc GlcNAc
D	1,460.16^{+2} ≡ 2,919.31^{+1} (calc.: 2,919.22)	
E	1,379.14^{+2} ≡ 2,757.27^{+1} (calc.: 2,757.16)	 Fuc Gal or Fuc Man GlcNAc
F	1,298.11^{+2} ≡ 2,595.17^{+1} (calc.: 2,595.11)	

4. Notes

1. This procedure is applicable only if the deglycosylated protein will not be further analyzed because it is nearly impossible to redissolve the protein pellet from an ethanol precipitate.
2. SEC is restricted to smaller sample volumes (20–75 µL) but provides a high (glyco-)protein recovery while ultrafiltration may be used for larger volumes but involves, especially at low starting protein concentrations, significant sample loss.
3. Suitable for proteins without disulfide bonds or native proteins with disulfide bonds.
4. Applicable for proteins containing disulfide bonds.
5. The (glyco-)protein under investigation may turn out to be resistant to proteolytic digestion without initial protein denaturation. In these cases, procedure b is the only option.

Acknowledgement

We thank the "Deutsche Forschungsgemeinschaft" (SFB 492/Z2) for financial support.

References

1. Rademacher, T. W., Parekh, R. B., Dwek, R. A. (1988) Glycobiology. *Ann. Rev. Biochem.* 57, 785–838
2. Lloyd, K. O. (2000) The chemistry and immunochemistry of blood group A, B, H, and Lewis antigens: Past, present and future. *Glycoconj. J.* 17, 531–541
3. Williams, D. W. (2006) Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. *J. Cell Sci.* 119, 615–623
4. Mitra, N., Sinha, S., Ramya, T. N. C., Surolia, A. (2006) *N*-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* 31, 156–163
5. Paulson, J. C. (1989) Glycoproteins: what are the sugar chains for?. *Trends Biochem. Sci.* 14, 272–276
6. Rasmussen, J. R. (1992) Effect of glycosylation on protein function. *Curr. Op. Struct. Biol.* 2, 682–686
7. Potter, B. A., Hughey, R. P., Weisz, O. A. (2006) Role of *N*- and *O*-glycans in polarized biosynthetic sorting. *Am. J. Physiol. Cell Physiol.* 290, C1–C10
8. Clark, G. F., Dell, A. (2006) Molecular models for murine sperm–egg binding. *J. Biol. Chem.* 281, 13853–13856
9. Sagi, D., Kienz, P., Denecke, J., Marquardt, T., Peter-Katalinić, J. (2005) Glycoproteomics of *N*-glycosylation by in-gel deglycosylation and matrix-assisted laser desorption/ionisation-time of flight mass spectrometry mapping: application to congenital disorders of glycosylation. *Proteomics* 5, 2689–2701
10. Müthing, J., Kemminer, S. E., Conradt, H. S., Sagi, D., Nimtz, M., Karst, U., Peter-Katalinić, J. (2003) Effects of buffering conditions and culture pH on production rates and glycosylation of clinical phase I anti-melanoma mouse IgG3 monoclonal antibody R24. Biotechnol. Bioeng. 83, 321–334
11. Kunneken, K., Pohlentz, G., Schmidt-Hederich, A., Odenthal, U., Smyth, N., Peter-Katalinić, J., Bruckner, P., Eble, J. A. (2004) Recombinant human laminin-5 domains: effects of heterotrimerization, proteolytic processing,

and *N*-glycosylation on α3β1 integrin binding. *J. Biol. Chem.* 279, 5184–5193

12. Imre, T., Schlosser, G., Pocsfalvi, G., Siciliano, R., Molnar-Szollosi, E., Kremmer, T., Malorni, A., Vekey, K. (2005) Glycosylation site analysis of human alpha-1-acid glycoprotein (AGP) by capillary liquid chromatography – electrospray mass spectrometry. *J. Mass Spectrom.* 40, 1472–1483
13. Wuhrer, M., Deelder, A. M., Hokke, C. H. (2005) Protein glycosylation analysis by liquid chromatography–mass spectrometry. *J. Chromatogr B Analyt. Technol. Biomed. Life Sci.* 825, 124–133
14. Pohlentz, G., Kölbl, S., Peter-Katalinić, J. (2005) High sequence coverage by in-capillary proteolysis of native proteins and simultaneous analysis of the resulting peptides by nanoelectrospray ionization-mass spectrometry and tandem mass spectrometry. *Proteomics* 5, 1758–1763
15. Henning, S., Peter-Katalinić, J., Pohlentz, G. (2007) Structure elucidation of glycoproteins by direct nano ESI MS and MS/MS analysis of proteolytic glycopeptides. *J. Mass Spec.* 42, 1415–1421
16. Steiner, K., Pohlentz, G., Dreisewerd, K., Berkenkamp, S., Messner, P., Peter-Katalinić, J., Schaffer, C. (2006) New insights into the glycosylation of the surface layer protein SgsE from *Geobacillus stearothermophilus* NRS 2004/3a. *J. Bacteriol.* 188, 7914–7921
17. Hanisch, F. G., Jovanovic, M., Peter-Katalinić, J. (2001) Glycoprotein Identification and Localization of *O*-glycosylation sites by mass spectrometric analysis of deglycosylated/alkylaminylated peptide fragments. *Anal. Biochem.* 290, 47–59
18. Peter-Kataliniæ, J. (2005) O-glycosylation of proteins. *Methods Enzymol.* 405, 139–171
19. Kuster, B., Wheeler, S. F., Hunter, A. P., Dwek, R. A., Harvey, D. J. (1997) Sequencing of *N*-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography. *Anal. Biochem.* 250, 82–101
20. Packer, N. H., Lawson, M. A., Jardine, D. R., Redmond, J. W. (1998) A general approach to desalting oligosaccharides released from glycoproteins. *Glycoconj. J.* 15, 737–747
21. Domon, B., Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* 5, 397–409

Chapter 11

Capillary Zone Electrophoresis-Mass Spectrometry for the Characterization of Isoforms of Intact Glycoproteins

Christian Neusüß, and Matthias Pelzing

Summary

Electrospray ionization-mass spectrometry (ESI-MS) is a powerful tool for the characterization of intact proteins. However, complex samples require separation in order to obtain clear mass spectra and avoid matrix interaction; capillary electrophoresis (CE) has been shown to be a powerful separation technique for intact proteins. Herein, we present a method based on capillary zone electrophoretic (CZE) separation coupled online with mass spectrometric protein detection. While this approach is suitable for the separation and characterization of various intact proteins, the emphasis is placed on the separation of glycoforms of various and rather complex glycoproteins. The method has been successfully applied to the analysis of glycoproteins, e.g., erythropoietin, fetuin, and alpha-acid glycoprotein.

Key words: Capillary electrophoresis-mass spectrometry, Intact proteins, Glycosylation, Accurate mass, Carbohydrates, Glycoproteins.

1. Introduction

An increasing number of drugs are based on recombinant proteins and their importance for biopharmaceutical developments are expected to increase in the future. However, there is a lack of fast methods that are appropriate for both quality control and support of optimization of these complex molecules. This is primarily due to the often observed microheterogeneity of posttranslationally modified proteins. Especially glycosylation frequently leads to a broad range of isoforms. This complexity challenges direct mass spectrometric analysis of such complex mixtures and necessitates high-quality separation, which is often difficult to realize due to variable properties of proteins.

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_11

Capillary zone electrophoresis has proven to be a powerful separation technique for intact proteins *(1–3)*. First, surface interaction is more limited than in, e.g., liquid chromatography and the remaining surface (i.e., the capillary wall) can be passivated against interaction. Second, the separation efficiency is extremely high especially for large molecules like proteins as longitudinal diffusion – principally the only contributor to peak broadening during separation – has a rather low effect. Third, and perhaps of greatest interest in the context of this chapter, many posttranslational modifications like glycosylation or phosphorylation can be differentiated as they change the size and/or the charge of the protein and thus potentially its electrophoretic mobility.

Coupling of the capillary electrophoresis (CE) to the mass spectrometry creates a powerful analytical tool that is a rather good alternative or complement for the characterization of peptides and especially proteins *(4–6)*. The method described herein enables separation of both different proteins and glycoforms of a single protein. The excellent glycoform separation results in high-quality mass spectra, high dynamic range, and good sensitivity, enabling the characterization of minor glycan modifications *(7, 8)*. CZE-MS is an ideal complement to the established techniques for glycopeptide and glycan analysis, not differentiating branching or linkage isoforms, but leading to an overall composition of the glycoprotein with the possibility to get quantitative information on the isoforms. This method has been successfully applied for the characterization of several proteins; however, due to the extreme variety of proteins (e.g., with respect to the isoelectric point) certain experimental parameters might need adjustment for different applications.

2. Materials

2.1. Chemicals and Supplies

1. All chemicals should be of highest purity. As most of them are available at LC-MS grade nowadays, this is the best choice. However, HPLC/gradient grade will suffice in most cases, especially where sensitivity is not the key issue.
2. Isopropanol (Lichrosolv, HPLC grade), acetic acid (Suprasolve), and Ammonia (25%) were obtained from VWR International (Darmstadt, Germany).
3. Methanol (HPLC gradient grade) was obtained from J.T. Baker (Phillipsburg, NJ, USA).
4. TuneMix was obtained from Agilent Technologies (Waldbronn, Germany) and 1:50 diluted in 95% ACN 5% water

before use. The diluted solution can be stored for some weeks at 6–8°C.

5. Hexadimethrine bromide (Polybrene), bovine fetuin, bovine α_1-acid glycoprotein (bovine AGP), ammonium hydrogen carbonate, and 1 M NaOH solution were supplied by Sigma/Fluka (Taufkirchen, Germany).
6. UltraTol^{TM} Dynamic Pre-Coat LN was provided by Target Discovery (Palo Alto, CA, USA).
7. Standard recombinant human erythropoietin (rHuEPO) was obtained as BRP from Pharmacopoeia (EDQM, European Pharmacopoeia, Council of Europe, Strasbourg, France).
8. Deionised and organic-eliminated water was obtained using a Milli-Q water purification system (Millipore, Schwalbach, Germany).
9. All electrolytes and the sheath liquid need to be degassed by ultrasonication for at least 10 min. The sheath liquid especially needs prior thorough shaking of freshly prepared mixtures in the reservoir bottle.
10. The storage of the electrolytes, coating solutions, and sheath liquid is best in high-quality glass bottles (borosilicate) or Teflon bottles. Use bottles always for the same solutions in order to prevent cross contamination.
11. For ultrafiltration Microcon-10 cartridge (Millipore, Schwalbach, Germany) with a cutoff around 10 kDa has been used.

2.2. Instrumentation

2.2.1. Capillary Electrophoresis

1. Any CE instrument can be used as long as the end of the capillary can be taken out of the instrument. The power supply should be able to produce stable 30 kV (or slightly less) at both polarities at the inlet vial. Here a HP^{3D}CE instrument (Agilent Technologies, Waldbronn, Germany) is used.
2. Fused-silica capillaries of 50 μm ID × 360 μm OD are supplied by Polymicro Technologies (Phoenix, AZ, USA). The length of the capillary is typically in the order of 80 cm. The time needed for the analysis is related quadratically with the length of the capillary. While theoretically the separation is not influenced by the length of the capillary, slight dependence has been observed in practice (*see* **Note 1**).
3. Proper (flat) cutting is important especially for the outlet of the capillary. Either a rotating diamante tool or a ceramic cutter can be used. In the latter case, the best cut is obtained if the scribing pressure is as low as possible and the actual cut is performed by pulling at both ends while slightly beginning to cant. The quality of the cut should be controlled under a microscope. Especially a "nose" of glass prolonging over the cut should be avoided.

4. The insertion of the capillary is performed according to the standard procedure for the actual instrument. The outlet is left outside the sprayer for conditioning and coating procedures (*see* **Note 2**).
5. The instrument is controlled by the Chemstation software.

2.2.2. Coupling

1. For CZE-MS coupling, a coaxial sheath-liquid sprayer is used (Agilent Technologies). Sheath liquid is delivered by a 5-mL gas-tight syringe (Hamilton, Reno, NV, USA) using a Cole-Parmer syringe pump (Vernon Hill, IL, USA) (*see* **Note 3**).
2. Both the CE and the MS are controlled by applying the respective software. A contact closure cable provides the connection between both instruments by transferring start/stop signals. Thus, unattended sequences are possible by proper timing of the CE and MS method, respectively.

2.2.3. Mass Spectrometry

1. Different types of mass spectrometers may be used. These instructions assume the use of a micrOTOF or micrOTOF-Q mass spectrometer from Bruker Daltonik, Bremen, Germany.
2. Acceleration of the ions with high potential in the first vacuum stages is essential for efficient transfer of highly charged intact proteins cleavage of the clusters formed by the attachment of the solvent molecules.
3. To avoid interference of the calibrant with the analyte external mass calibration is recommended.
4. The micrOTOF and micrOTOF-Q is equipped with an analog to digital converter (ADC) for a discrimination free measurement of the isotopic pattern which allows using singly charged ions for the calibration of highly charged molecules like intact proteins.

3. Methods

Different methods can be used for the characterization of proteins by CZE-MS. For instance, a formic acid based system is suitable for the separation of many peptides and proteins *(4, 6)*. However, for the separation of the glycoforms of complex glycoproteins like fetuin or erythropoietin, the use of the acetic acid as a background electrolyte (BGE) yields better results.

3.1. Sample Treatment

While the analysis typically requires a total protein concentration of about 1 μg/μl for the characterization of large and/or complex proteins, less complex or smaller proteins can be analyzed at the low to medium ng/μl level (*see* **Note 4**).

The purity of the sample may influence both the injection and the separation. For proteins highly soluble in water, a precleaning step by ultrafiltration is recommended in the cases where higher amounts of salts, detergents, or other small molecules are present (*see* **Note 5**):

1. Protein solutions should be prepared fresh. Depending on the protein, solutions can be stored at 6–8°C for several days. Longer storage at −20°C may be possible for certain proteins.
2. Ultrafiltration is performed in the following way: The filter membrane is initially washed with deionized water for 10 min at 14,000 × *g*. Then the sample is added to the filter and centrifuged at the same acceleration for 10 min. The residue is washed three times with 10 μl of water. The residue is recovered from the inverted cartridge by centrifugation for 3 min at 1,000 × *g*. This step is repeated a few times with the addition of 2 μl of water to the inverted cartridge in order to increase the recovery.
3. The sample is acidified to a final concentration of 0.2 M acetic acid by the addition of 1.1 μl of 20% acetic acid and adjusted to a final volume of 20 μl. Finally, the sample is degassed by ultrasonication for 5 min (*see* **Note 6**).

3.2. Capillary Zone Electrophoresis

The separation is the most critical step after the setup of the instrumentation. Careful attention has to be given to the subsequent steps:

1. All solutions and BGEs are degassed by ultrasonication before use.
2. New capillaries are conditioned with 1 M NaOH solution by flushing at maximum pressure of the CE instrument for 20 min followed by flushing with water for 20 min (*see* **Note 7**).
3. Coating of the capillary is performed by flushing the capillary with UltraTrol solution for 40 min (*see* **Note 8**).
4. Five minutes of flushing with water is followed by 40 min of the respective BGE containing 2 M acetic acid (*see* **Note 9**).
5. The capillary is now coupled to the mass spectrometer (see below).
6. The overall performance and the stability of the system are tested by applying 30 kV for 10 min. The current should be stable within ±0.5 μA. For a 80-cm-long capillary a current of around 14 μA is expected (*see* **Note 10**).
7. Before starting the run, the volume of the capillary should be exchanged at least twice by flushing the capillary (3 min 1 bar) (*see* **Note 11**).

8. Injection is performed by pressure. An initial injection plug of about 1% of the capillary (750 mbar = 50 mbar for 15 s) should be used. Higher injection (up to 10% of the capillary) can be performed with the sample of low ionic strength (e.g., following the preconcentration protocol described above).
9. After the start of the analysis, the CE current should be observed. Slightly lower and at first decreasing current is caused by low conductivity of the sample zone, thus, is more pronounced when a long injection plug is used (*see* **Note 10**).
10. Postconditioning is performed by flushing the capillary for several minutes with BGE (*see* **Note 12**).

A separation of fetuin is shown in **Fig. 1**. The Base peak electropherogram already shows some separation. Four ion traces are displayed in order to show the actual separation. The charge envelope used to display the different isoforms has been derived from mass spectra (see below). The first two proteins belong to

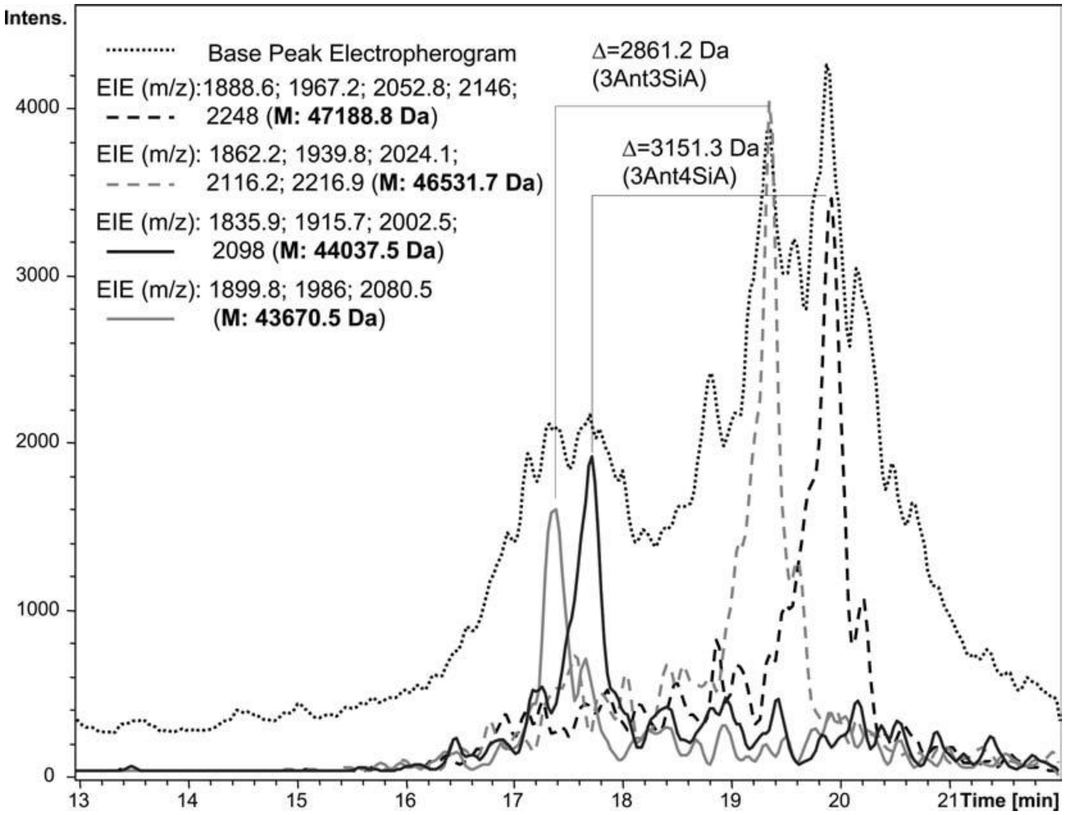


Fig. 1. Base peak electropherogram and selected extracted ion electrophreograms acquired for bovine fetuin glycoforms (capillary 1 m × 50 μm, coated with UltraTol™ Dynamic Pre-Coat LN, BGE: 2 M acetic acid, Sep. Volt. +30 kV) (modified from ref.*8*).

a group of glycoforms with one complete glycan less than the group of proteins in the time window around the third and fourth mass trace. The first pair of traces corresponds to two glycoforms of fetuin differing by one sialic acid. Much better separation for sialic acid isoforms is achieved for EPO *(9)*, where even differences in (noncharged!) hexose-*N*-acetylhexosamine units could be resolved on the basis of different migration.

3.3. Coupling

For the CE-MS coupling, mass spectrometers with the electrospray potential placed on the MS inlet are beneficial since stable spray conditions can be obtained more easily and BGEs resulting in a higher current can be used. Thus, the coupling described in **Subheading 3.4** is only possible by employing the mass spectrometer with the grounded sprayer and ESI potential applied to the inlet of the mass spectrometer (i.e., only commercial suppliers: Agilent Technologies and Bruker Daltonics).

1. The CE capillary should be placed into the sprayer after the capillary has been preconditioned, coated, and washed with BGE as described earlier.
2. The sheath liquid is prepared by mixing 1% acetic acid in a mixture of isopropanol:water 1:1. The sheath liquid requires degassing prior to filling the syringe. A small air bubble at the end of the syringe helps damping potential pulsation. The syringe is coupled via standard fittings and PEEK tubing to the sprayer.
3. The capillary is fed into the sprayer according to the instructions given. The actual positioning of the capillary at the tip of the sprayer is controlled by a magnifying glass. The position is best when only a slice of the capillary is protruding from the metal needle. For more details on the preparation and reproducibility of the CE-MS coupling, *see* **ref.** ***10***.

3.4. Mass Spectrometry

Different type of mass spectrometers can be used. However, time-of-flight (TOF) or quadrupole-time-of-flight (Qq-TOF) type instruments are best suited with respect to resolution and mass accuracy. Quadrupole type instruments are slow in scanning and limited in resolution and mass accuracy and therefore inappropriate for this purpose. Ion trap instruments bear the risk of adduct formation (ion molecule reactions) due to the relatively high pressure in the trap (up to 10^{-2} Torr) and relatively long dwell time of the ions (up to 100 ms). Fourier transform ion cyclotron resonance (FT-ICR) instruments provide a very high resolution on one hand but on the other hand show limitations particularly for large proteins due to the space charge effects in the ICR trap at low magnetic field strength (12T). The overall performance of an ESI-TOF or ESI-Qq-TOF system with respect to resolution, sensitivity, and mass accuracy make this type of mass spectrometer ideally suited for this application.

1. Electrospray ionization is used in positive mode (negative potential on the inlet of the MS), –4.5 kV.
2. A nebulizer gas pressure of 0.2 mbar is applied to assist the spraying process (*see* **Note 13**).
3. The dry gas (nitrogen) temperature is set to 180°C. Higher temperatures can lead to instabilities of the CE-MS interface.
4. The acceleration voltage of the mass spectrometer is set to 9,000 V which allows high repetition rates (improving the signal-to-noise ratio).
5. Instrument operation is conducted by using a method developed by optimizing the transfer conditions for the expected mass range of about 1,000–3,000. For this purpose TuneMix at a dilution of 1:50 is used.
6. The trigger time is set to 90 µs corresponding to a mass range of 50–3,000 m/z.
7. Spectra were acquired by summarizing 20,000 single spectra, leading to a time resolution of 1.8 s.
8. Due to their additional negatively charged groups, glycosylated proteins are shifted in their major charge distribution toward higher m/z values. For fetuin and erythropoietin the charge envelope stretches over the region between 2,000 and 3,000 m/z. Nonglycosylated proteins of this size normally appear in the mass range of around 1,500 m/z.
9. Prior to the start of the acquisition the instrument is calibrated with TuneMix at dilution of 1:50 in the mass range of 600–3,000 m/z.
10. The "maximum entropy" algorithm in DataAnalysis^{TM} version 3.4 (Bruker Daltonics) is used for charge deconvolution of the averaged protein mass spectra (**Fig. 2a**).
11. The major ions from a charge distribution belonging to one mass are added up to one extracted ion electropherogram (EIE) to increase the selectivity of the separated traces (compare **Fig. 1**).

Figure 2 shows a mass spectrum derived from the analysis shown in **Fig. 1**. Even at this small time interval of about 0.3 min, a strong heterogeneity is observed (**Fig. 2a**). After charge deconvolution (**Fig. 2b**), the mass differences can be attributed to differences in glycosylation.

Two dozen glycoforms can be distinguished for fetuin (and more than 50 for erythropoietin). Moreover, even small modifications can be resolved. This is illustrated in **Fig. 3** for one glycoform of erythropoietin. The mass precision is generally better than 1 Da (intraday, different user, different instrument) for these isotopically nonresolved proteins. The measured masses indicate not only acetylation ($\Delta = 42$), but also the ability to resolve the

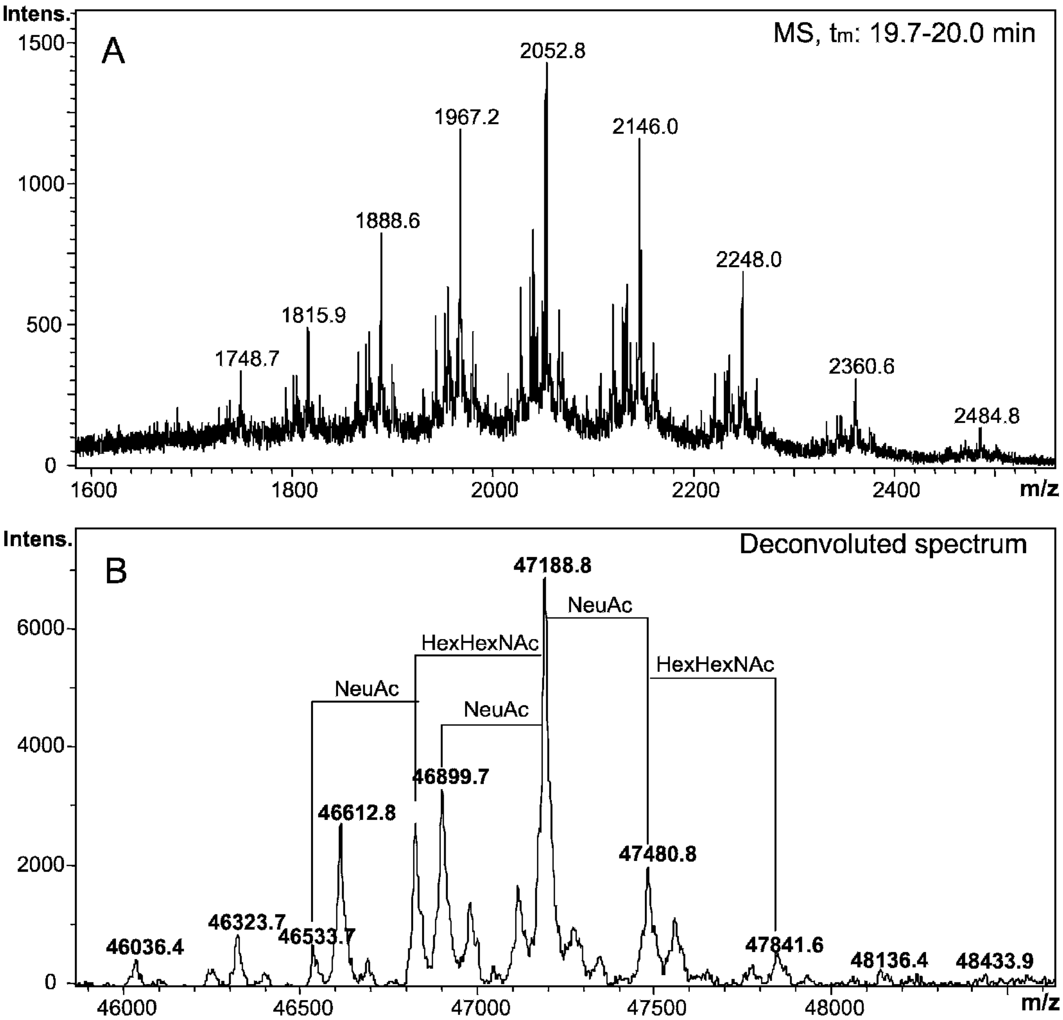


Fig. 2. Mass spectrum (**a**) and corresponding charge-deconvoluted spectrum (**b**) of time window 19.7–20 min from the analysis shown in **Fig.1**.

difference in oxygen content. This has been previously demonstarted for alpha-acid glycoprotein, where a broad distribution of oxygenation (Acetyl-neuraminic acid vs. Glycoyl-neuraminic acid) could be resolved.

4. Notes

1. In practice, a shorter capillary length has only minor negative influence on the separation as long as a minimum length of

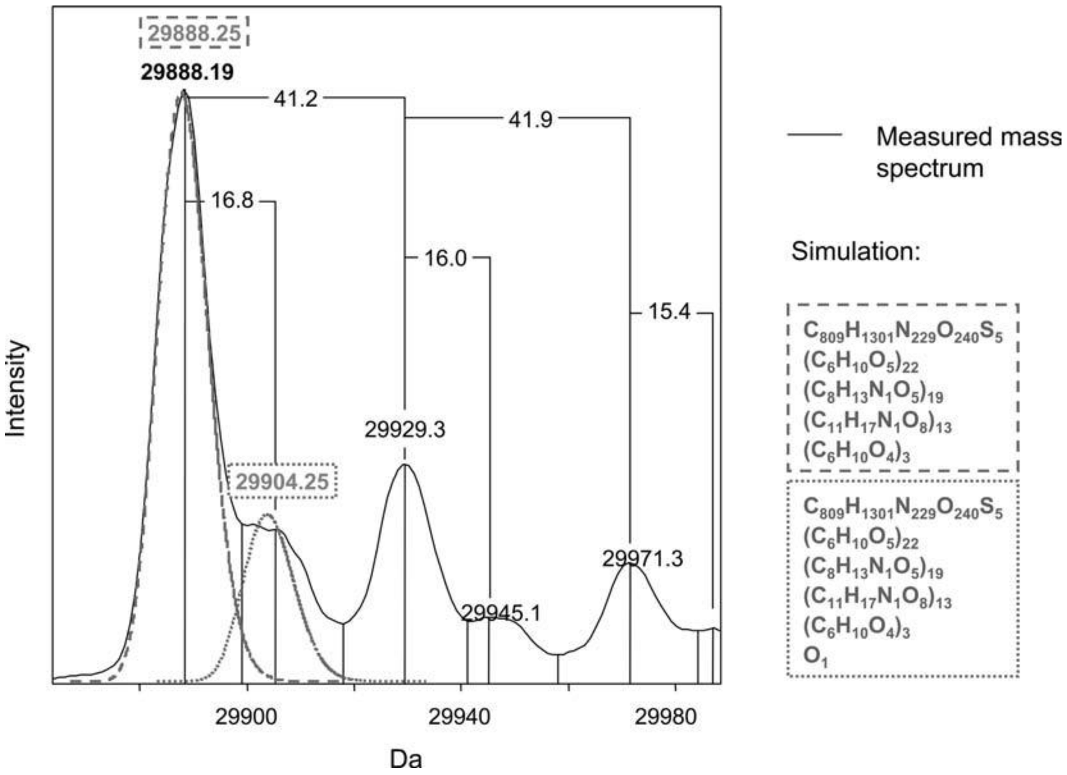


Fig. 3. Detail of a measured charge-deconvoluted mass spectrum of erythropoietin in comparison to results of simulation.

the capillary of about 50–60 cm is used, which is the shortest possible length for commercial instruments anyway.

2. Contamination of the interface and the mass spectrometer with nonvolatile substances such as NaOH (preconditioning and rinsing of the capillary) or coating solution should be strictly avoided. It is best to have the capillary during these procedures not installed in the sprayer. In order to repeat frequent rinsing and conditioning steps in a sequence, at least the electrospray potential should be switched off. In the remote controlled system used here, the easiest way to do so, is by adding a respective short (seconds) time segment at the beginning of the run.
3. The UltraTrol gives an electro-osmotic flow (EOF) close to zero and enables efficient separation of glycoforms of erythropoietin, fetuin, and similar proteins. However, other coating solutions may be more appropriate for other proteins. Depending on the protein or the sample matrix, a coating with a strong reversed EOF might be advantageous. Successful glycoform separation has been achieved by coating the capillary with a polybrene solution (2% hexadimethrine
)

bromide (Polybrene) in watery solution containing 10% isopropanol and 2% polyethylene glycol).

4. The overall concentration depends on the number of isoforms the protein might display. Furthermore, the required concentration depends on the information expected. In case that a higher concentration is used, isoforms of minor concentration can be detected.
5. Ultrafiltration can be used for preconcentration as well. The volume for reconstitution of the protein can be as low as a few tens of microliters, whereas the loadability of the cartridges can be as high as milliliters.
6. The acidification step ensures the compatibility of the sample with the electrolyte. A lower ionic strength than the electrolyte, however, allows for a larger injection plug and results in increased sensitivity. Therefore, a 1:10 diluted BGE is useful as standard method.
7. Cleaning of the electrodes is necessary in case different coating has been in contact with the electrodes previously. To avoid contamination, the level of the vial containing the coating solution should be low enough so that it does not reach the level of the electrode. Alternatively, the coating procedure can be performed outside of the CE instrument using a syringe (syringe pump, standard connection to capillary, though attention has to be paid not to damage the (perfectly cut) end of the capillary).
8. In order to avoid any contamination with other coatings potentially run on the same instrument, *all* vials should be used only for a certain application.
9. Alternatively, a HPLC pump can be used to provide the sheath liquid, if the flow rate can be adjusted to stable flow rates down to 2 μl/min. (A smaller syringe is acceptable, but, frequent refilling of the syringe is then required.)
10. BGE: The addition of organic solvents may increase the separation (e.g., 20% methanol showed better results in the polybrene-coated capillaries). However, it was not tried here, because of the risk of solubilizing the coating.
11. As an alternative to acetic acid, formic acid based BGE can be used. However, acetic acid showed better results for the separation of the glycoforms of erythropoietin or fetuin. The addition of ammonia in order to adjust the pH might be critical for the ESI process and/or wall interaction of the protein. No spectra were obtained for erythropoietin.
12. The actual value depends on the capillary length and may differ slightly (e.g., typically it is higher for aged/re-used electrolyte); however, the value needs to be kept constant.

Instability of the current can be due to different reasons, causing negative flow in the system. Sheath liquid will then enter the capillary and can cause current breakdown or a continuous decrease of the current. In this case any siphoning caused by height differences of the CE and the MS should be eliminated and the level of CE system relative to the MS increased. Alternatively, the nebulizer pressure can be slightly increased or a low pressure can be applied during separation. These two latter parameters are useful in order to find out whether siphoning is the reason for the current breakdown. Generally, errors in the sheath liquid setup (air bubbles, low flow, etc.) may lead to low currents in the CE. A current breakdown during the run (with injection) is often associated with the incompatibility of the sample and the BGE caused by differences in pH (acidification of a basic system causes gas bubbles, i.e., CO_2) or organic solvent content (different solubility of air) could be the reason.

13. In case impurities are observed at initial conditions, it often helps to repeat the application of high voltage (also reversed!) with subsequent change of the vial.
14. In case that the migration time shifts, a recoating should be applied by flushing (at 1 bar) subsequently with water (3 min), 1 M NaOH (3 min), water (3 min), coating solution (≥3 min), water (3 min), and BGE (10 min).
15. The nebulizer pressure should be low in order to prevent a siphoning effect in the capillary. Depending on the actual position of the inner metal capillary inside the sprayer the value may be slightly (!) higher in order to achieve a stable spray.

Acknowledgment

We thank Jonathan Moss for proof reading of the manuscript.

References

1. Hutterer, K., and Dolník, V. (2004) Capillary electrophoresis of proteins 2001–2003, *Electrophoresis*, 22, 3998–4012.
2. Patrick, J. S., and Lagu, A. L. (2001) Review applications of capillary electrophoresis to the analysis of biotechnology-derived therapeutic proteins, *Electrophoresis*, 22, 4179–96.
3. Kamoda, S., and Kakehi, K. (2006) Capillary electrophoresis for the analysis of glycoprotein pharmaceuticals, *Electrophoresis*, 27, 2495–504.
4. Hernández-Borges, J., Neusüß, C., Cifuentes, A., and Pelzing, M. (2004) On-line capillary electrophoresis-mass spectrometry for the analysis of biomolecules, *Electrophoresis*, 25, 2257–81.

5. Simpson, D. C., and Smith, R. D. (2005) Combining capillary electrophoresis with mass spectrometry for applications in proteomics, *Electrophoresis*, 26, 1291–305.
6. Bateman, K. P., Kelly, J. F., Thibault, P., Ramaley, L., and White, R. L. (2003) Glycoprotein analysis by capillary zone electrophoresis-electrospray mass spectrometry, in *Capillary Electrophoresis of Carbohydrates* (ed. P. Thibault and S. Honda), Methods in Molecular Biology, Vol. 213, Humana, Totowa, NJ, pp. 219—39.
7. Neusüß, C., Demelbauer, U., and Pelzing, M. (2005) Glycoform characterization of intact erythropoietin by capillary electrophoresis-electrospray-time of flight-mass spectrometry, *Electrophoresis*, 26, 1442–50.
8. Balaguer, E., and Neusüß, C. (2006) Glycoprotein characterization combining intact protein and glycan analysis by capillary electrophoresis-electrospray ionization-mass spectrometry, *Analytical Chemistry*, 78, 5384–93.
9. Balaguer, E., Demelbauer, U., Pelzing, M., Sanz-Nebot, V., Barbosa, J., and Neusüß, C. (2006) Glycoform characterization of erythropoietin combining glycan and intact protein analysis by capillary electrophoresis – electrospray – time-of-flight mass spectrometry, *Electrophoresis*, 27, 2638–50.
10. Ohnesorge, J., Neusüß, C., and Wätzig, H. (2005) Quantitation in capillary electrophoresis – mass spectrometry, *Electrophoresis*, 26, 3973–87.

Chapter 12

Top-Down Proteomics on a High-field Fourier Transform Ion Cyclotron Resonance Mass Spectrometer

Séverine A. Ouvry-Patat, Matthew P. Torres, Craig A. Gelfand, Hung Hiang Quek, Michael Easterling, J. Paul Speir, and Christoph H. Borchers

Summary

Mass spectrometry is the tool of choice for sequencing peptides and determining the sites of posttranslational modifications; however, this bottom-up approach lacks in providing global information about the modification states of proteins including the number and types of isoforms and their stoichiometry. Recently, various techniques and mass spectrometers, such as high-field Fourier Transform Ion Cyclotron Resonance (FTICR) mass spectrometers, have been developed to study intact proteins (top-down proteomics). While the protein molecular mass and the qualitative and quantitative information about protein isoforms can be revealed by FTICR-MS analysis, their primary structure (including the identification of modifications and their exact locations in the amino acid sequence) can directly be determined using the MS/MS capability offered by the FTICR mass spectrometer. The distinct advantage of top-down methods are that modifications can be determined for a specific protein isoform rather than for peptides belonging to one or several isoforms. In this chapter, we describe different top-down proteomic approaches enabled by high-field (7, 9.4, and 12 T) FTICR mass spectrometers, and their applicability to answer biological and biomedical questions. We also describe the use of the free flow electrophoresis (FFE) to separate proteins prior to top-down mass spectrometric characterization.

Key words: Top-down Proteomics, FTICR-MS, FFE.

1. Introduction

Mass spectrometry is well known for its ability to sequence peptides (bottom-up proteomics) typically obtained by proteolysis and to determine their posttranslational modifications (PTMs) *(1)*. However, bottom-up approaches provide only little (if any)

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_12

information about the physical status of the intact protein including the number and type of isoforms and their relative proportions. Indeed some proteins, such as histones, exhibit very complex PTM profiles, which are crucial to their biological function *(2)*. Thus, methods and instruments capable of determining the global modification patterns of proteins, rather than the “bits and pieces” typically obtained from a bottom-up approach, are highly desired *(3, 4)*. In top-down proteomics, larger peptides and intact proteins are intensively analyzed in the mass spectrometer. In general, this means that no prior enzymatic or chemical protein digestion required as is the case for bottom-up proteomics. With recent development and improvement of instruments and techniques, protein analysis via top-down methods has become feasible *(5–9)*. In particular, high-field Fourier Transform Ion Cyclotron Resonance (FTICR) mass spectrometry used in electrospray ionization (ESI) mode is very suitable to perform top-down proteomics *(7, 10, 11)*. High-field FTICR mass spectrometry has the distinct advantage of combining ultra-high mass accuracy and ultra-high resolution while maintaining high sensitivity, which makes this technique well suited for studying proteins extracted from biological samples. This technology is capable of resolving isotopic clusters of protein ions and determining monoisotopic masses for high molecular mass proteins (up to ca. 60 kDa). Furthermore, FTICR mass spectrometry offers a range of methods to carry out gas phase dissociation of proteins useful for protein sequencing, such as collision induced dissociation (CID), electron capture dissociation (ECD), in-source and thermal fragmentation, infrared multiphoton dissociation (IRMPD) *(12–14)*. While CID is available for the determination of the amino acid sequence, ECD is extremely useful in determining labile PTM sites, such as phosphorylation, acylation, and O-linked glycosylation *(15, 16)*. IRMPD has proven valuable for determining the amino acid composition including modified amino acid residues *(17, 18)*. These dissociation approaches can be used alone or in combination, allowing multiple stages of tandem MS, which is useful for sequencing fragment ions that provide detailed structural analysis of modifications.

To perform top-down experiments, the proteins need to be in a relatively pure form, otherwise ion suppression effects can deteriorate the quality of the data and render the mass spectra too complex for analysis. In addition, proteins need to be present in solvents compatible with ESI, limiting the applicability to soluble (e.g., cytosolic) proteins (i.e., ESI is not very tolerant to salts or detergents). Online top-down proteomic experiments on an FTICR instrument (e.g., LC-MS) are still in their infancy (primarily due to the slow duty cycle of the instrument). Free flow electrophoresis (FFE) is particularly well suited for offline coupling with the FTICR-MS because it offers high-resolution protein separation while sustaining proteins in solution.

Top-down proteomics on an FTICR instrument is a challenging endeavor and requires a lot of savoir faire. Unfortunately there is no “set in stone” protocol as every sample is unique. In this chapter, we describe the basic steps to perform top-down experiments with a Bruker Daltonics 7, 9.4, or 12- FTICR. Some of the protocols are applicable to other instruments and some are specific for the configuration of our system. We also describe the approach that combines offline protein separation via FFE with top-down experiments on the FTICR and demonstrate its high potential in the analysis complex biological systems. As this separation method is still fairly new, detailed protocols have been included.

2. Materials

2.1. Isoelectric Focusing Free Flow Electrophoresis (pH 3–10)

2.1.1. Denaturing IEF-FFE of Intact Proteins

1. BD™ FFE System with 0.4-mm spacer (Becton Dickinson, Franklin Lakes, NJ).
2. Water, 18.2 MΩ (Milli-Q, Millipore, Billerica, MA).
3. Urea (Serva Electrophoresis GmbH, Heidelberg, Germany).
4. Mannitol (98%, Sigma, St. Louis, MO).
5. Glycerol (>99.5%, Sigma).
6. Isopropanol (100%, Fisher Scientific, Fair Lawn, NJ).
7. Sodium Hydroxide (ACS grade, Sigma).
8. Sulfuric Acid (ACS grade, Sigma).
9. BD™ FFE-IEF kit (cat. #441118; Becton Dickinson) pH 3–10 gradient kit, which includes hydroxypropyl methyl cellulose (HPMC) powder; pI mix, and buffers 1, 2, and 3 for the pH 3–10 gradient.
10. Microtest flat bottom 96-well plates #35–1172 (Fisher Scientific).
11. Trifluoroacetic acid, HPLC-grade (Sigma).
12. Methanol, HPLC-grade (Sigma).
13. Acetic acid, glacial (Sigma).

2.1.2. Buffers for Protein Separation via FFE

To ensure accuracy, all reagents (including liquids) should be weighed out as follows and used the same day. Store at 4°C until needed.

1. 1 M sodium hydroxide: 16.0 g sodium hydroxide, Milli-Q water to 400.0 g.
2. 5 M sulfuric acid: 49.0 sulfuric acid, 73.3 g Milli-Q water.
3. 1 M sulfuric acid: 9.81 g Sulfuric acid, 94.67 g Milli-Q water.
4. Anodic stabilization medium: 3.0 g 5 M sulfuric acid, 72.0 g urea, 6.8 g mannitol, 69.7 g Milli-Q water.

5. Separation medium 1: 30.3 g IEF-FFE buffer 1, 72.0 g urea, 6.8 g mannitol, 69.7 g Milli-Q water.
6. Separation medium 2: 100.0 g IEF-FFE buffer 2, 144.0 g urea, 13.6 g mannitol, 100.0 g Milli-Q water.
7. Separation medium 3: 37.0 g IEF-FFE buffer 3, 72.0 g urea, 6.8 g mannitol, 63.0 g Milli-Q water.
8. Cathodic stabilization medium: 30.0 g 1 M sodium hydroxide, 144.0 g urea, 13.6 g mannitol, 170.0 g Milli-Q water.
9. Counterflow medium: 360.0 g urea, 33.7 g mannitol, 500.0 g Milli-Q water.
10. Electrolyte anode buffer: 40.0 g 1 M sulfuric acid, 360.0 g Milli-Q water.
11. Electrolyte cathode buffer: 40.0 g 1 M sodium hydroxide, 360.0 g Milli-Q water.
12. 0.8% HPMC solution (*see* **Note 1**): 8.0 g HPMC powder, 1000.0 g Milli-Q water. Add water to large clean glass vessel with a large stirrer bar. Stir the water on a stirring plate until a strong vortex is created so that the base of the vortex is touching the stirrer bar. Gradually add the HPMC powder over a 10-min period. Allow mixing to occur overnight and then store at 4°C in the dark.
13. 0.1% HPMC solution: 105.0 g 0.8% HPMC solution, 525.0 g Milli-Q water

2.1.3. Microscale Reverse-Phase Purification of IEF-FFE Protein Fractions

1. C_4 Zip-Tips (Millipore).
2. Water, HPLC-grade (Pierce, Rockford, IL).
3. Acetonitrile, HPLC-grade (Pierce).
4. Trifluoroacetic acid, HPLC-grade (Sigma).

2.2. FTICR-MS

1. An Apex-Qe Qq-FTICR mass spectrometer (Bruker Daltonics, Billerica, MA) equipped with a 7-, 9.4-, and 12-T actively shielded magnet and an Apollo II (Bruker Daltonics) ESI source.
2. Mass spectrometry grade Chromasolv solvents: acetonitrile, 0.1% formic acid and water, 0.1% formic acid (Sigma).
3. HPLC grade water, acetonitrile, and methanol (Fisher Scientific).
4. Acetic acid and formic acid (Sigma).
5. Amber glass with Teflon tops to store solvents (Fisher Scientific).
6. Ammonium acetate (Sigma).
7. Protein standards for calibration: ubiquitin, myoglobin, and carbonic anhydrase (Sigma).

3. Methods

3.1. Protein Separation by IEF-FFE off-line Coupled to Top-down Analysis

FFE is distinguished from other protein separation techniques by its resolving power and its ability to separate proteins in solution (**Fig. 1**). The separation can be performed under denaturating (urea) or a variety of nondenaturating conditions, allowing separation of functional proteins and intact protein complexes. Although most FFE protocols require a low conductivity environment for optimal protein separation, a desalting step via reversed phase purification is typically performed prior to ESI-MS top-down experiments to eliminate interferences which may arise from, e.g., urea, sugars, glycerol, ampholytes, and other small molecules present in various FFE buffers. In particular, we find that prompt removal of urea from the FFE fractions can significantly reduce protein carbamylation, which is caused by prolonged exposure to urea as it breaks down to isocyanic acid and leads to a mass increment of 43 Da per modified amino-group.

3.1.1. Protein Separation via FFE

1. Ensure that assembly of the FFE system (with 0.4-mm spacer) is complete, free of air bubbles and in compliance with the manufacturer's specifications.
2. Add each of the FFE inlet and counterflow inlet lines into the 0.1% HPMC solution and set the media pump flow rate to 150 mL/h. Allow this to flow for 5 min (*see* **Note 2**).

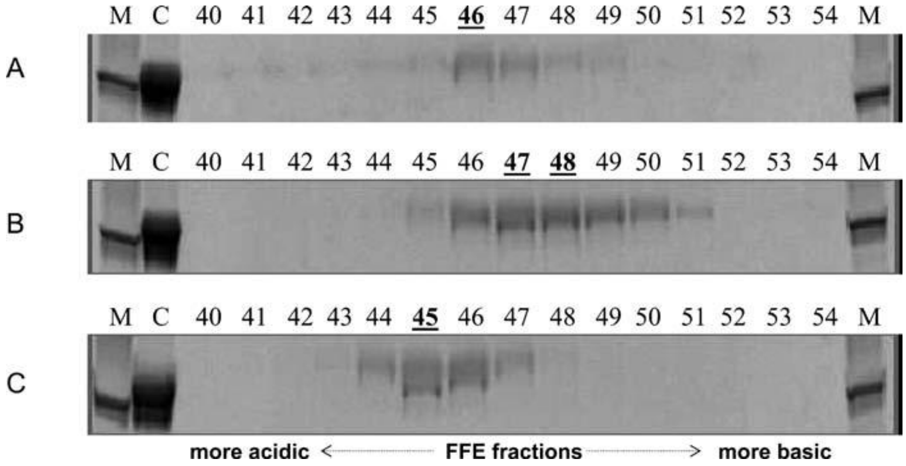

M C 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 M
A
M C 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 M
B
M C 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 M
C
more acidic <— FFE fractions —> more basic

Fig. 1. FFE for separating protein isoforms. SDS-PAGE of fractions resulting from IEF-FFE separations, using a 4–7 pH range under denaturing conditions (8 M urea, 4.4% mannitol, and appropriate proprietary FFE ampholytes). **a** Wild-type CDA, **b** E91A CDA mutant, and **c** H102N CDA mutant. *Numbers* indicate FFE fraction in each lane, with the fractions having the majority of the protein concentration *underlined*. M, molecular mass marker (31 kDa band showing); C, unfractionated CDA protein. In the pH 4–7 protocol, the resulting fractions are separated by a pH increment of approximately 0.04 pH units.

3. Stop the pump, remove the FFE inlet and counterflow inlet lines from the HPMC solution and place in a 1 L beaker of Milli-Q water. Turn the pump on at 150 mL/h and allow this to flow for 5 min.
4. Stop the pump, remove the FFE inlet and counterflow inlet lines from the water and place into the counterflow media. Turn the pump on at 150 mL/h and allow this to flow for 15 min.
5. Stop the pump and place each FFE inlet and counterflow inlet line into the following IEF buffer:
   (a) Inlet 1: anodic stabilization media
   (b) Inlet 2: separation media 1
   (c) Inlet 3–4: separation media 2
   (d) Inlet 5: separation media 3
   (e) Inlet 6–7: cathodic stabilization media
   (f) Counterflow inlets: counterflow media
6. Run the media pump at 72 mL/h for 10 min (keep the flow going from this point forth).
7. Prime the sample inlet by uncapping the inlet and allowing flow from the separation chamber for 2 min. Connect the sample inlet to the sample pump.
8. Place the electrode contact lid onto the separation chamber.
9. Set the voltage, current, and power maxima to the following:
   (a) Voltage (V_{max}) = 650 V
   (b) Current (I_{max}) = 50 mA
   (c) Power (P_{max}) = 60 W
10. Turn the voltage on and allow the current to stabilize for 20 min. Once stabilized, the current should be 21 mA ± 3 mA.

3.1.2. Isoelectric Focusing of Cytidine Deaminase by IEF-FFE

1. Dilute 10 μL of cytidine deaminase (1 nmol or 31.5 μg) into 90 μL of separation buffer 3 (*see* **Note 3**).
2. Inject the sample into the separation chamber using a sample pump flow rate of 1 mL/h. Continue sample injection until the entire 100 μL volume has been loaded into the separation chamber.
3. Stop the sample pump after injection is complete.
4. Begin collecting plate fractions using a microtest flat bottom 96-well plate 17 min after sample injection is stopped (*see* **Note 4**).
5. Collect two plates at 11 min per plate.
6. After plate collection, plates are frozen at −80°C or processed immediately using C_4 reverse-phase chromatography (*see* **Note 5**).

3.1.3. C4 Reversed Phase Purification of IEF-FFE Fractions

1. Collect IEF-FFE fractions of interest into individual microfuge tubes.
2. Bind the sample to a C_4 Zip-tip pre-equilibrated with 0.1% trifluoroacetic acid.
3. Wash with 10 × 10 μL volumes of 0.1% trifluoroacetic acid.
4. Elute with 5 μL 80% methanol, 1% acetic acid.
5. Freeze sample at −80°C or proceed to MS analysis.

3.2. Preparation of Proteins for Direct Infusion into the FTICR

The sample introduction via direct infusion mode (i.e., without any analytical separation coupled online to the mass spectrometer) is advantageous for top-down experiments as the time constraint for the MS analysis is set by the sample amount and the flow rate and therefore can be tailored for a particular application. However, prior to the ESI-MS analysis, the proteins must be desalted and present in relatively simple mixtures.

3.2.1. Sample Preparation

1. Commercially available proteins can be used without further process by diluting them with spraying solvent (*see* **Subheading 3.3.1**).
2. Complex mixtures of proteins resulting from cell lysates, immunoprecipitation, or other extraction means, as well as tagged-proteins need to be separated by reversed phase HPLC, nickel column, gel filtration, FFE, or some other analytical separation method.
3. HPLC fractions can be sprayed directly onto the mass spectrometer after lyophilization and resuspension with spraying solvent (*see* **Subheading 3.3.1**).
4. Fractions resulting from separation methods containing salts and detergent need to be desalted with C4 ziptips (*see* **Subheading 3.1.2**).

3.2.2. Protein Concentration

Dilute proteins to a concentration of 0.5–5 μM with spraying solvent (*see* **Subheading 3.3.1**). In general, consider that the larger the protein, the higher the concentration should be. For instance, 0.5 μM solution will be prepared for ubiquitin (M_r ~8,560 Da) vs. 3 μM solution for carbonic anhydrase (M_r ~29,000 Da).

3.3. Top-down Experiments by High-Field FTICR-MS

In most top-down experiments, the molecular masses of the proteins are initially determined in the ESI MS-mode. The mass spectrum of the stem-loop binding protein (SLBP) depicted in **Fig. 2** demonstrates the ability of FTICR to accurately determine the masses of proteins present in a simple protein mixture and differentiate various protein isoforms even when they are present in low abundance. The minor peak at *m/z* 922.2 has one fewer phosphoryl group than the major peak at *m/z* 927.5 *(14)*. The FTICR analysis of proteins can be performed under denaturating

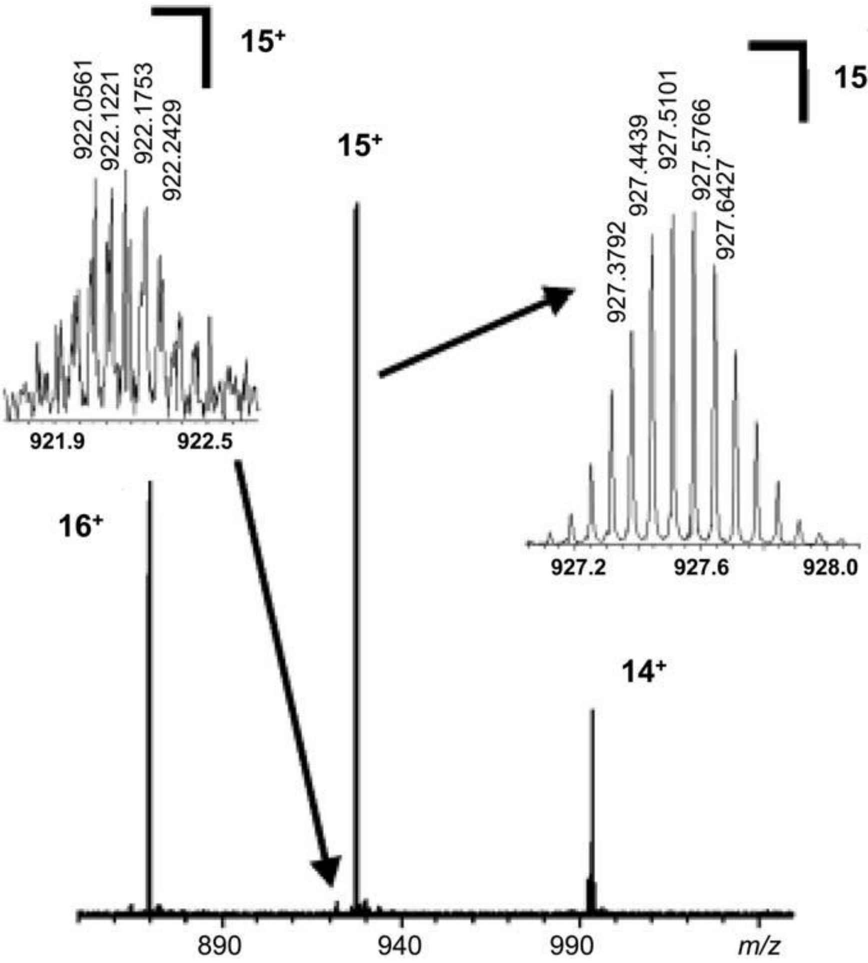


Fig. 2. Molecular mass determination of the *Drosphilia* truncated His-tagged stem-loop binding protein (dSLBP-RPD) using a Bruker Apex Qq-FTICR-MS equipped with a 9.4 T magnet. The minor peak at *m/z* 922.2 has one fewer phosphoryl group than the major peak at *m/z* 927.5. Reproduced from ref. *14* with permission from the National Academy of Sciences.

(e.g., high concentration of organic solvents and low pH) or native conditions, which is useful for characterization of protein complexes including the determination of the components and their stoichiometry within the complex. **Figure 3** shows the ESI-MS spectrum of cystidine deaminase (CDA) enzyme introduced under denatured conditions; **Fig. 4** depicts the ESI-MS spectrum of the same protein introduced into the FTICR under nondenaturing conditions and reveals the formation of an eight component complex (i.e., dimeric CDA, two zinc ions, two water molecules, and two molecules of the hydrated transition state analog) *(19)*.

For direct analysis of the primary structure of proteins via top-down proteomics, numerous and complementary dissociation mechanisms are offered by modern hybrid FTICR-MS instruments. Protein sequence information can be obtained through skimmer-

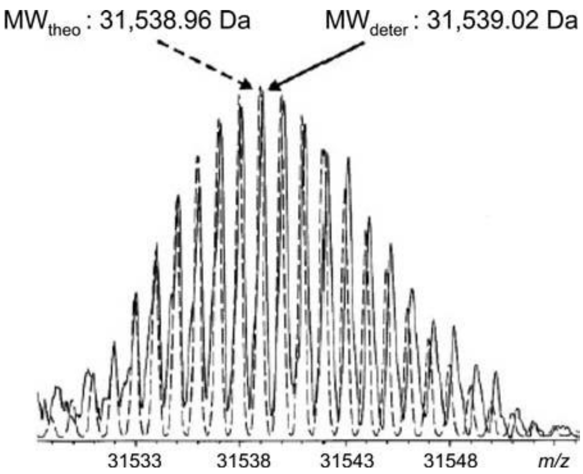


Fig. 3. FTICR-MS analysis of cytidine deaminase (CDA) under denaturing conditions by 9.4 T FTICR-MS. Deconvoluted spectrum (*solid line*) is superimposed on the theoretical isotopic distribution of CDA (dashed line) derived from its elemental composition. The most abundant mass agrees with the theoretical mass within 2 ppm (0.13 Da). theo., theoretical; deter., experimentally determined; MW, molecular mass. Reproduced from ref. *19* with permission from the National Academy of Sciences .

induced fragmentation, which leads mainly to sequence-specific y- and b-ions. Hybrid FTICR mass spectrometers also offer CID capability. **Figure 5** shows tandem mass spectrum obtained by CID of a fairly large protein, CDA (M_r = 31,520) purified by FFE prior to MS analysis.

A major advantage of FTICR is that high mass accuracy is maintained in the tandem MS analysis, facilitating sequencing of unknown proteins as demonstrated by de novo sequencing of SLBP (**Fig. 6a**) *(14)*. Combining of skimmer-induced dissociation and CID is particularly useful for confirmation of the fragment ion assignments and for detailed structural analysis of modified fragment ions. **Figure 6b** illustrates FTICR-MS^3 analysis, where a fragment ion produced by skimmer-induced dissociation of SLBP was selected for CID, revealing that the N-terminal amino acid residue of the protein is removed and that the new N-terminus is acetylated *(14)*. ECD is especially valuable for sequencing modified proteins because it almost exclusively and exhaustively breaks peptide amide bonds, creating mainly c- and z-ions. **Figure 6c** shows the wealth of data generated in the ECD-MS/MS analysis of SLBP, which resulted in direct identification of a phosphorylation site in SLBP *(14)*. IRMPD is well known to mimic typical low energy CID patterns in proteins, but the associated secondary fragmentation processes are very efficient at producing immonium ions. While immonium ions have little specificity for protein identification through database searching, they are useful for de novo sequencing strategies (**Fig. 7**). In FTICR, IRMPD is usually performed by irradiating ions that are trapped

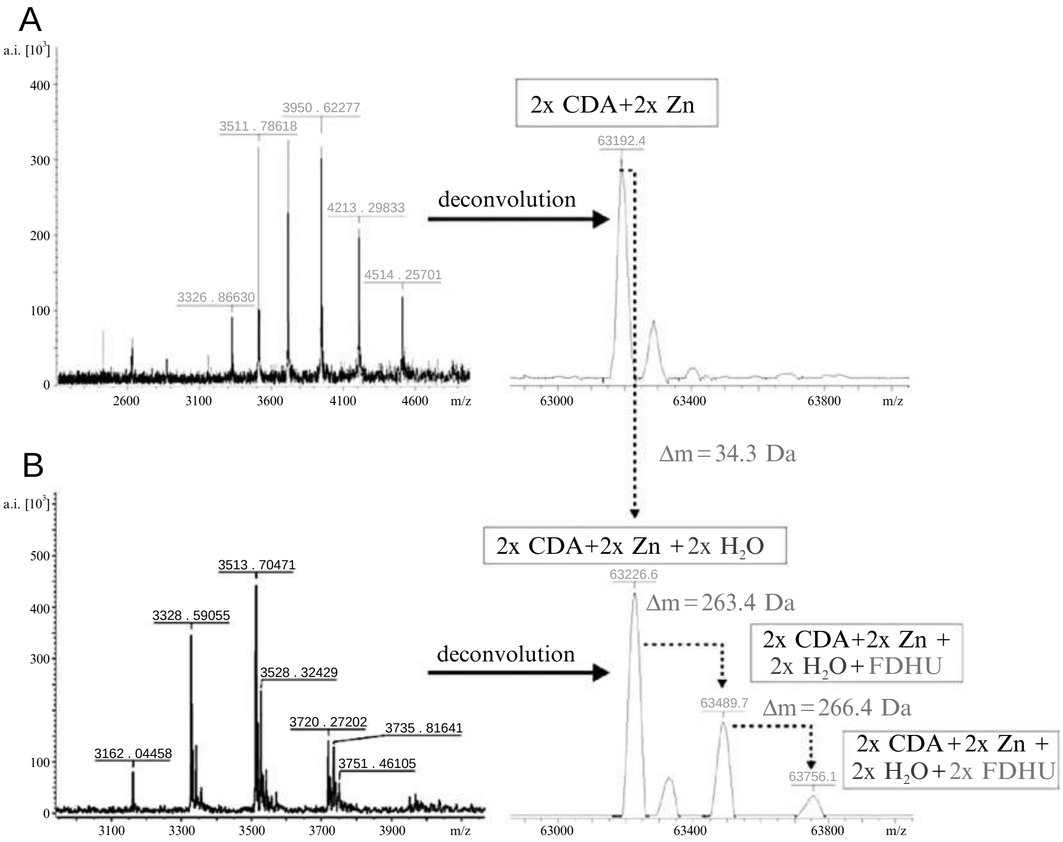


Fig. 4. FTICR-MS analysis of CDA under nondenaturing conditions by Bruker 7T Apex Qq-FTICR in the absence (**a**) and in the presence (**b**) of the transitation state analog FZeb. After deconvolution of the multiply charged ion signals, a mass difference of 35 Da was observed for the complex of dimeric CDA plus two zinc ions after incubation with the transition state analog FZeb, which corresponds to ca. one water molecule per monomer of CDA. In addition, the presence of a second water molecule is revealed by the detection of species at 63,489 and 63,756 Da, corresponding to complexes consisting of dimeric CDA, two zinc ions, two water molecules, and either one or two molecules of the hydrated transition state analog, respectively. Reproduced from ref. *19* with permission from the National Academy of Sciences. (*See Color Plates*)

in the ICR cell. This is especially significant for immonium ion production and detection because immonium ions produced externally to the ICR cell are more likely to be lost during the ion transport from the source or collision cell regions.

All experiments in MS and MS/MS mode are performed using a high-field (7, 9.4, or 12 T) Quadrupole-Fourier Transform Ion Cyclotron Resonance (Qq-FTICR) mass spectrometer (Bruker Daltonics). While our intention was to present protocols generally applicable on a hybrid FTICR spectrometer, portions of these protocols might be specific to the particular instruments used in these studies.

3.3.1. Spraying Solvents

Prepare fresh spraying solvent. A variety of solvents can be used for top-down experiments. The correct choice of solvent usually depends on the protein and on the experiment to be performed: denaturing vs. native.

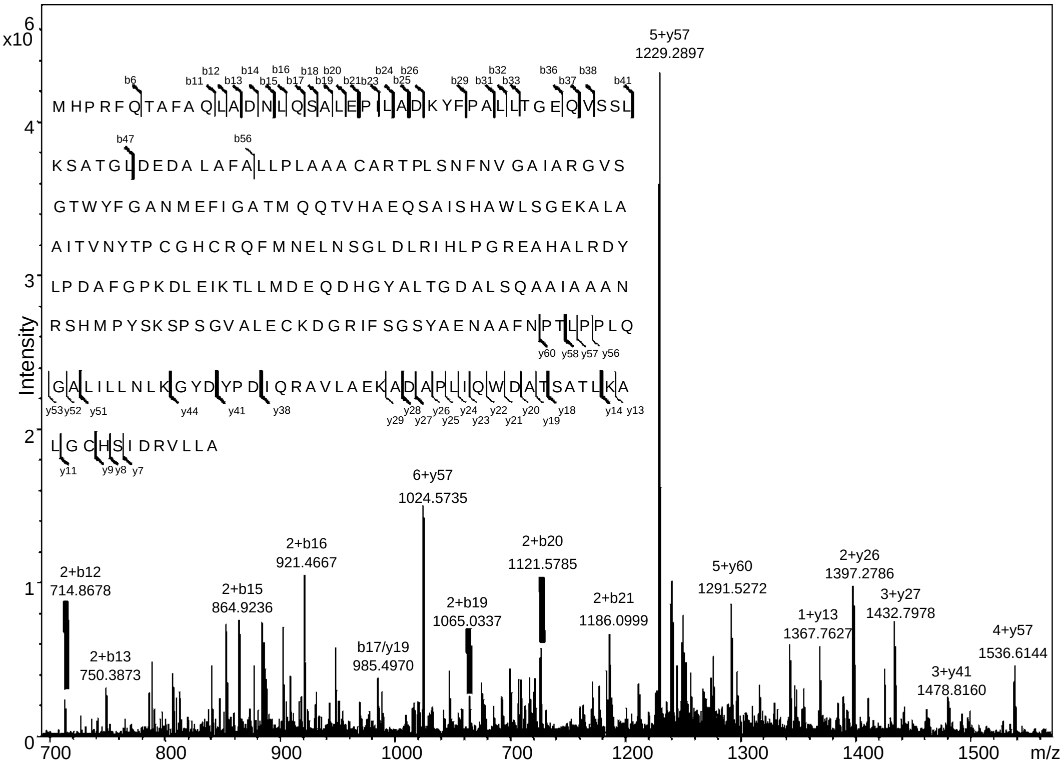


Fig. 5. Tandem mass analysis (CID) of intact CDA under denaturing conditions using offline FFE and Bruker 12T Apex Qq-FTICR. CID was performed on the $(M+29H)^{29+}$ ion (*m/z* 1,088). The spectrum shows deconvoluted masses as determined by the SNAP2 peak picking and deconvolution algorithm of Data Analysis (Bruker Daltonics). Reproducing from ref. *20* with permission from the Wiley-VCH.

1. We have found that for denaturing conditions, methanol or acetonitrile work well with formic acid or acetic acid. The concentration of organic solvent is adjusted to 50–80% and the concentrations of acids vary from 0.05 to 2%. It is extremely important to keep the solvents clean and avoid containers (such as glass containers) that may release a large amount of salts as they will cluster with the proteins and complicate the spectra which will obscure analysis and interpretation. In addition, low-retention Eppendorfs tubes designed for mass spectrometry should be used to avoid sample loss and possible contamination by polymers.
2. For native conditions, a 10 mM ammonium acetate solution with or without acid is appropriate.

3.3.2. Source Mode

The Apex-Qe Qq-FTICR mass spectrometer (Bruker Daltonics) is equipped with a capillary (2 μL/min) and a nanospray (200 nL/min) ESI source.

1. In capillary mode, the analyte is loaded directly.
2. In nanospray mode, a background spray is established and the analyte is loaded through a six-port valve. This allows for

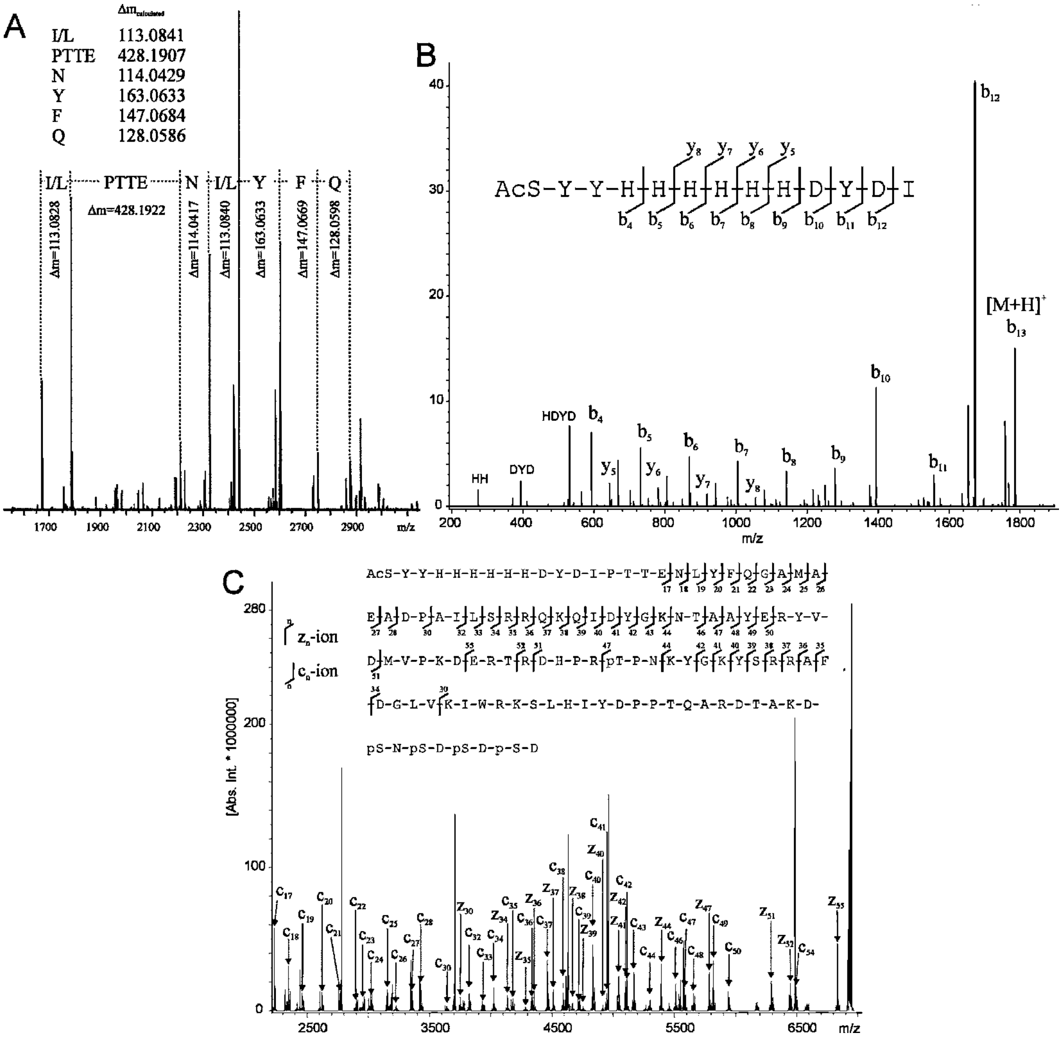


Fig. 6. Top-down analysis of dSLBP-RPD by FTICR-MS. **a** De novo sequencing of the $(M+15H)^{15+}$ ion (*m/z* 927.5) of phosphorylated dSLBP-RPD by CID. The fragment ions did not match the predicted sequence, but de novo sequencing identified modified N terminus. **b** FTICR-MS^3 analysis of dSLBP-RPD by skimmer-induced fragmentation followed by CID. The first MS stage was performed by skimmer-induced fragmentation, and ions with the mass corresponding to the doubly charged b13 ion were selected for CID. BIOTOOLS software assignments show that the N-terminal methionine has been removed (MSYY-HHHHHHDYDI...). **c** ECD analysis of dSLBP-RPD. The phosphorylation site T230 was directly confirmed by the observation of the z-ion z47. N-terminal (b ions by CID, c ions by ECD), C-terminal (y ions by CID, z ions by ECD), and abundant internal fragment ions are assigned. Reproduced from ref. *14* with permission from the National Academy of Sciences.

good spray conditions and saves time as only the loop has to be washed between samples. When the sample is diluted in a high concentration of organic solvents (such as 80%), the background spray is kept at 50% organic solvent because it is more stable.

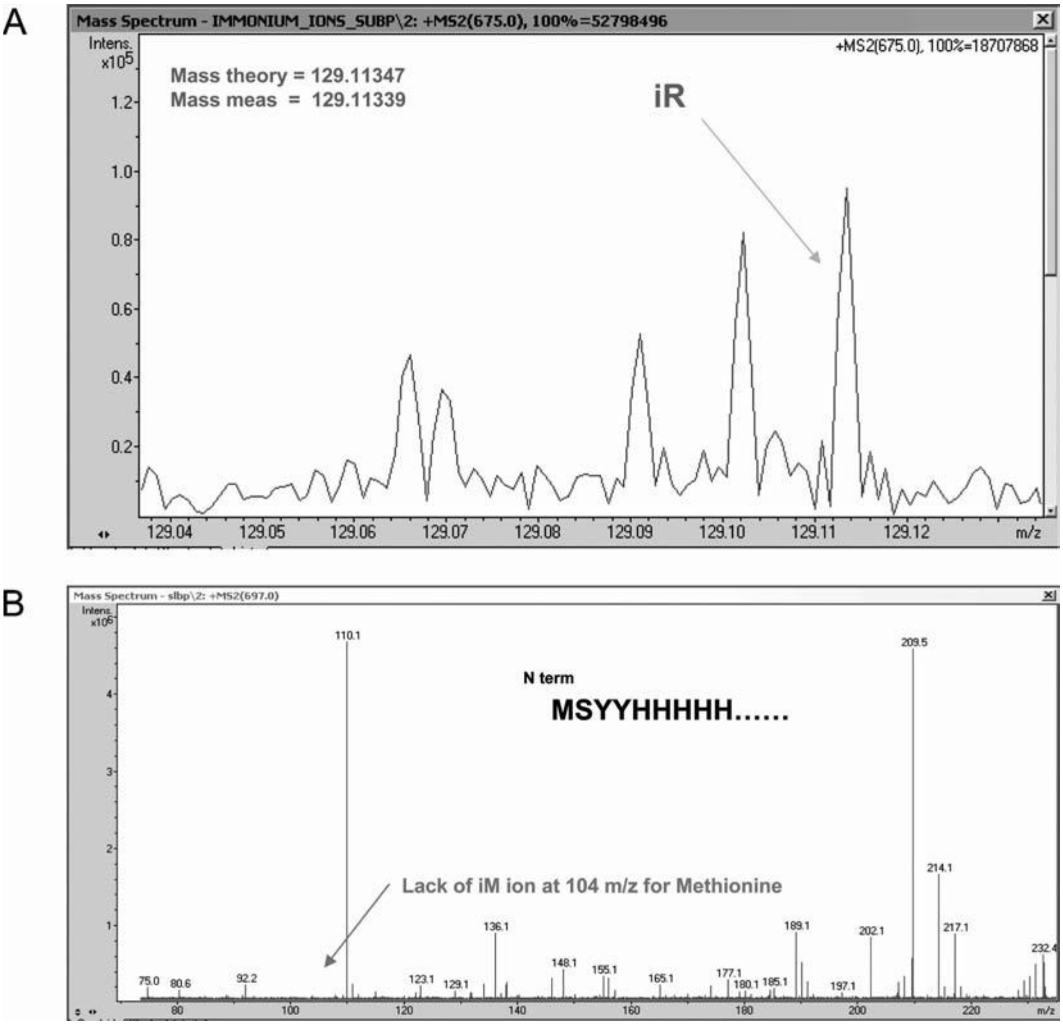


Fig. 7. FTICR-MS^2 and MS^3 experiments employing IRMPD for determination of the amino acid composition. **a** Secondary fragmentation processes produce isobaric ions; thus, the mass resolving power and accuracy of the FTMS is critical for assignment. In this example, the arginine immonium ion at *m/z* 129 is detected along with five other fragments. The strongest peak within the isobaric cluster can be confidently assigned as the arginine immonium ion since the mass deviation from the theoretical mass is only 0.00008 Da. **b** FTICR-MS^3 experiment on dSLBP-RPD used IRMPD to dissociate b-ion obtained by CID. The absence of the methionine immonium ion confirms de novo assignment of the b-ion as described in **Fig. 6b.** (*See Color Plates*)

3.3.3. MS Experiments

1. If running nanospray, establish background spray. If not, proceed to **step 2**.
2. Load standard protein such as ubiquitin at 0.5 μM.
3. Tune the instrument to obtain the best signal. It is critical to tune with the quadrupole isolation turned on as it is easy to lose isolation efficiency when tuning in a broadband mode. Important tuning parameters include the pressure in the source hexapole (~1.5 × 10^{-3} mbar), the skimmer 2 and hexapole DC

voltages, the vertical and horizontal beam steers, and the ion excitation parameters. The protein ions are trapped in the ICR cell by means of a gated trapping: the cell is filled with *N* iterations of ion (e.g., isolated precursor) accumulation from the external collision cell. This experimental design minimizes the storage time at high pressure, which could lead to charge reduction. Usually, the ICR cell is filled 4–8 times before excitation and detection. The trapping voltages and the number of cell fills also have to be tuned for a specific experiment. Hexapole accumulation time and time-of-flight (TOF) need to be adjusted (e.g., longer TOF will favor higher *m/z* ions).

4. Once a satisfying signal is obtained, a spectrum of the calibrant is acquired and the instrument is calibrated using, e.g., the fifth isotope peak of the various ubiquitin charge states.
5. The values of the calibrated file are used to acquire the sample.
6. It is critical to keep the same parameters when acquiring the calibrant spectrum and the sample spectrum. The concentrations of the two also need to be similar because the number of ions in the ICR cell affects the calibration.

3.3.4. MS/MS Experiments: Collision-Induced Dissociation

1. The 12T Apex-Qe Qq-FTICR mass spectrometer is equipped with a mass selective quadrupole Q that allows for isolation of the ion of interest and a collision cell q (hexapole) that allows for CID with argon as the collision gas at a pressure of 1.5×10^{-3} mbar.
2. The charge state of interest is selected through the quadrupole Q with a window width of 10 *m/z*. The time in the hexapole (i.e., accumulation time) can usually be increased as fewer ions are entering the ICR cell.
3. The collision voltage in the hexapole is decreased until the parent ion has almost disappeared and fragments appear.
4. A mass spectrum is acquired. If the protein is large, a large number of spectra is averaged (50–100 spectra) to increase signal-to-noise of the fragments.

3.3.5. MS/MS Experiments: Electron Capture Dissociation

1. The 12T Apex-Qe Qq-FTICR mass spectrometer is equipped with a hollow dispenser cathode.
2. Turn on the cathode slowly (0.1 A/min) up to 1.7 A (*see* **Note 6**).
3. Adjust the bias voltage (1–5 V) and the pulse length (500 ms).
4. Acquire spectrum (average 50–100 spectra).

3.3.6. MS/MS Experiments: Infrared Multiphoton Dissociation

1. The 12T Apex-Qe Qq-FTICR mass spectrometer is equipped with a BaF_2 window and 25 W (CW) CO_2 laser.
2. Isolate the precursor ion of interest using the Qq region.

3. Set the laser power to 35% and irradiation time to 0.1 s.
4. Increase the irradiation time in increments of 0.1 s until fragmentation is observed.
5. Acquire spectrum (average 50–100 spectra).

3.3.7. MS^3 Experiments: CID Combined with Skimmer-Induced Fragmentation

1. Obtain random fragment ions by skimmer-induced fragmentation: the skimmer voltage is increased to indiscriminately fragment ions.
2. A fragment ion is then mass selected in the Qq region and the CID experiment performed as described in **Subheading 3.3.4**.

3.4. Data Analysis

Rich tandem mass spectra typically obtained for proteins and challenges associated with determination of the neutral mass from isotopic clusters representing different protein charge states necessitate sophisticated software algorithms. Listed below are protocols specifically used with Bruker FTICR instrument to determine the molecular mass based on the detection of monosiotopic ions and to assign protein fragment ions (including de novo sequencing tools).

1. Open data in Data Analysis software (Bruker Daltonics; version 3.4, built 179).
2. Determine monoisotopic peaks with the peak picking SNAP2 algorithm. This algorithm reports the monoisotopic mass at all the charge states as well as the deconvoluted monoisotopic masses.
3. If a graphic deconvoluted mass spectrum is desired, one can deconvolute the snapped spectrum. The molecular mass of the protein is readily determined.
4. To assign fragments, the data is exported after SNAP2 mass list determination to Biotools (Bruker Daltonics), which matches theoretical fragments (based on the user's protein) to the observed fragments.

4. Notes

1. HPMC solutions are prone to bacterial growth and should be discarded if not used within 2–3 weeks.
2. Coating with HPMC is carried out to reduce electroendosmosis.
3. Before loading protein, the system should be tested by conducting a control experiment with the pI mix, which verifies proper isoelectric focusing.

4. Long-term exposure of protein to high urea concentrations can lead to carbamylation of lysine residues, resulting in unfavorable heterogeneity in downstream top-down MS analyses. Since carbamylation is inhibited in acidic buffers, we often preacidify each well of the 96-well collection plates with 20 μL of 10% trifluoroacetic acid (final acid concentration after fraction collection is approximately 0.9%). Although this step is not necessary for samples that will be processed soon after plate collection, the acidification does insure that protein fractions are never susceptible to carbamylation over long periods of time and/or multiple freeze/thaw cycles.
5. For analysis of simple protein mixtures, it is often useful to identify the protein-containing fractions by measuring the protein concentration of the collected plates using a Bradford or comparable assay. We typically assay 50 μL of each fraction combined with 150 μL of 20% Bradford reagent in a 96-well format and measure the absorbance of this at 595 nm using a microplate reader.
6. It is extremely important to turn the cathode on and off very slowly to avoid damaging it.

Acknowledgement

This work was supported in part by a Ford Foundation Diversity Fellowship(M.T.), by the Cystic Fibrosis Foundation (CFFT-IBORCHE05U0) and CDC H75/CCH424675 (S.O.-P.) and a grant sponsored by Genome BC and Genome Canada. A portion of this work was done at the UNC-Duke Proteomics Centre, which was supported by a gift from an anonymous donor to support research in proteomics at UNC. The Bruker 12 T FTICR mass spectrometer at the UNC-Duke Michael Hooker Proteomics Centre, UNC-CH, Chapel Hill, NC, was purchased with funding from the North Carolina Biotechnology Center (NCBC 2005-IDG-1015) and from the National Institutes of Health (1-S10-RR019889-01).

References

1. Wysocki, V. H., Resing, K. A., Zhang, Q., and Cheng, G. (2005) Mass spectrometry of peptides and proteins. *Methods* 35, 211–22.
2. Strahl, B. D. and Allis, C. D. (2000) The language of covalent histone modifications. *Nature* 403, 41–5.
3. Johnson, L., Mollah, S., Garcia, B. A., Muratore, T. L., Shabanowitz, J., Hunt, D. F., and Jacobsen, S. E. (2004) Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res* 32, 6511–8.
4. Thomas, C. E., Kelleher, N. L., and Mizzen, C. A. (2006) Mass spectrometric characterization of human histone H3: A bird's eye view. *J Proteome Res* 5, 240–7.

5. Senko, M. W., Speir, J. P., and McLafferty, F. W. (1994) Collisional activation of large multiply charged ions using Fourier transform mass spectrometry. *Anal Chem* 66, 2801–8.
6. Mortz, E., O'Connor, P. B., Roepstorff, P., Kelleher, N. L., Wood, T. D., McLafferty, F. W., and Mann, M. (1996) Sequence tag identification of intact proteins by matching tanden mass spectral data against sequence data bases. *Proc Natl Acad Sci USA* 93, 8264–7.
7. Kelleher, N. L., Lin, H. Y., Valaskvic, G. A., Aaserud, D. J., Fridriksson, E. K., and McLafferty, F. W. (1999) Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *J Am Chem Soc* 121, 806–12.
8. Kelleher, N. L., Taylor, S. V., Grannis, D., Kinsland, C., Chiu, H. J., Begley, T. P., and McLafferty, F. W. (1998) Efficient sequence analysis of the six gene products (7–74 kDa) from the *Escherichia coli* thiamin biosynthetic operon by tandem high-resolution mass spectrometry. *Protein Sci* 7, 1796–801.
9. Liu, Z. and Schey, K. L. (2005) Optimization of a MALDI TOF-TOF mass spectrometer for intact protein analysis. *J Am Soc Mass Spectrom* 16, 482–90.
10. Bogdanov, B. and Smith, R. D. (2005) Proteomics by FTICR mass spectrometry: Top down and bottom up. *Mass Spectrom Rev* 24, 168–200.
11. Hakansson, K., Cooper, H. J., Hudgins, R. R., and Nilsson, C. L. (2003) High resolution tandem mass spectrometry for structural biochemistry. *Curr Org Chem* 7, 1503–25.
12. Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J Am Chem Soc* 120, 3265–66.
13. Bakhtiar, R. and Guan, Z. (2005) Electron capture dissociation mass spectrometry in characterization of post-translational modifications. *Biochem Biophys Res Commun* 334, 1–8.
14. Borchers, C. H., Thapar, R., Petrotchenko, E. V., Torres, M. P., Speir, J. P., Easterling, M., Dominski, Z., and Marzluff, W. F. (2006) Combined top-down and bottom-up proteomics identifies a phosphorylation site in stem-loop-binding proteins that contributes to high-affinity RNA binding. *Proc Natl Acad Sci USA* 103, 3094–9.
15. Zubarev, R. A. (2004) Electron-capture dissociation tandem mass spectrometry. *Curr Opin Biotechnol* 15, 12–6.
16. Bakhtiar, R. and Guan, Z. (2006) Electron capture dissociation mass spectrometry in characterization of peptides and proteins. *Biotechnol Lett* 28, 1047–59.
17. Little, D. P., Speir, J. P., Senko, M. W., O'Connor, P. B., and McLafferty, F. W. (1994) Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Anal Chem* 66, 2809–15.
18. Hakansson, K., Cooper, H. J., Emmett, M. R., Costello, C. E., Marshall, A. G., and Nilsson, C. L. (2001) Electron capture dissociation and infrared multiphoton dissociation MS/MS of an N-glycosylated tryptic peptic to yield complementary sequence information. *Anal Chem* 73, 4530–6.
19. Borchers, C. H., Marquez, V. E., Schroeder, G. K., Short, S. A., Snider, M. J., Speir, J. P., and Wolfenden, R. (2004) Fourier transform ion cyclotron resonance MS reveals the presence of a water molecule in an enzyme transition-state analogue complex. *Proc Natl Acad Sci USA* 101, 15341–5.
20. Ouvry-Patat, S. A., Torresr, M.P., Querk, H., Gelfand, C. A., Schroeder, G. K., Han, M., Elliott, M., Dryhurst, D., Ausio, J., Wolfenden, R., and Borchers, C. H. (2008) Free-flow electrophoresis for top-dwon proteomics by Fourier by transfrom ion cyclotron resonance mass spectrometry. *Proteomics* 8, (in press).

Chapter 13

Capillary Isoelectric Focusing/Reversed Phase Liquid Chromatography/Mass Spectrometry

Cheng S. Lee and Brian M. Balgley

Summary

The vast number of proteins present in the proteome of a typical organism requires that separations be performed on the mixture prior to introduction into a mass spectrometer for protein identification and quantification. An integrated protein separation platform, combining capillary isoelectric focusing (CIEF) with reversed phase liquid chromatography (RPLC), is described to provide high resolving power for the analysis of complex protein mixtures. Thus, the proteins are systematically resolved according to their differences in isoelectric point and hydrophobicity using combined CIEF/RPLC separations. A key feature of the CIEF-based multidimensional separation platform is the elimination of protein loss and dilution in an integrated platform while achieving comprehensive and ultrasensitive analysis of protein profiles within small cell populations or limited tissue samples.

Key words: Capillary isoelectric focusing, Intact proteins, Mass spectrometry, Reversed-phase liquid, chromatography.

1. Introduction

The top-down proteome techniques *(1–4)*, in which intact proteins rather than peptides are measured and sequenced, may allow complete protein characterization far more efficiently than shotgun proteomics using protein digests. Realizing the potential of top-down approaches requires that the processing and separation of intact proteins be brought to a similar level as those routinely achieved in shotgun proteomics. Single dimension protein separations, including isoelectric focusing gel *(5)*, capillary isoelectric focusing (CIEF) *(6)*, and reversed phase liquid

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_13

chromatography (RPLC) *(7)* have been employed for processing protein mixtures prior to MS analysis. However, the use of only a single separation dimension may not provide sufficient peak capacity for the resolution of complex mixtures, putting significant constraints on the detection sensitivity and dynamic range of MS.

We have presented the application of combined CIEF/RPLC separations for concentrating and resolving intact proteins in an integrated platform *(8)*. The grouping of two highly resolving and completely orthogonal separation techniques greatly enhances the overall peak capacity for analyzing complex protein mixtures while eliminating protein loss and dilution in an integrated platform. The proteins resolved and eluted from RPLC are analyzed using electrospray ionization (ESI)-MS for the detection of a large number of protein envelopes and deconvoluted molecular masses.

2. Materials

2.1. Yeast Cell Lysis

1. Lyophilized *Saccharomyces cerevisiae* cells from Sigma (St. Louis, MO).
2. Cell lysis buffer containing 10 mM tris(hydroxymethyl) aminomethane (Tris) pH 8.0, 5 mM magnesium chloride, 0.1 mM dithiothreitol (DTT), and 10% glycerol.
3. A French press from Fisher (Pittsburgh, PA).
4. Solution of DNase (1 mg/mL) from Fisher.

2.2. Protein Denaturation, Reduction, and Alkylation

1. Protein denaturation and reduction buffer: 10 mM Tris-HCl pH 8.0, 100 mM DTT, and 8 M urea.
2. Solution of iodoacetamide (IAM) (500 mM).
3. PD-10 size exclusion columns from Amersham Pharmacia Biotech (Uppsala, Sweden).
4. Protein eluting buffer: 10 mM Tris-HCl pH 8.0 and 4 M urea.
5. Solution of pharmalyte 3–10 (Amersham Pharmacia Biotech).

2.3. Combined CIEF/RPLC Separations

1. Fused-silica capillaries (100 μm i.d./365 μm o.d. and 50 μm i.d./365 μm o.d.) from Polymicro Technologies (Phoenix, AZ).
2. Solutions of 0.5% (v/v) ammonium hydroxide (pH 10.5) and 0.1 M acetic acid (pH 2.5).
3. A CZE 1000R high-voltage power supply from Spellman High-Voltage Electronics (Plainview, NY).
4. Microcross, inline microfilter assemblies, a six-port micoinjection valve, and two six-port microselection valves from Upchurch Scientific (Oak Harbor, WA).

5. A Harvard Apparatus 22 syringe pump (Holliston, MA).
6. An Agilent 1,100 capillary LC pump (Avondale, PA).
7. Methanol solution containing C_{18}-bonded particles (5-μm diameter, 300-Å pores) from Phenomenex (Torrance, CA).
8. Solutions of mobile phase A (water, 0.1% trifluoroacetic acid, and 0.3% formic acid) and mobile phase B (acetonitrile, 0.1% trifluoroacetic acid, and 0.3% formic acid) from Fisher.
9. A quadrupole time-of-flight (q-TOF) mass spectrometer from Micromass (Manchester, United Kingdom).
10. Protein Trawler software from Bioanalyte (Portland, ME).

3. Methods

3.1. Soluble Fraction of Intact Proteins from S. cerevisiae

1. Suspend 3 g of yeast cells in 10 mL of cell lysis buffer.
2. Lysis the cells by a French press twice at 15,000 psi. After lysis, DNase is added to a final concentration of 50 μg/mL for the cleavage and removal of nucleic acids at 37°C for 30 min.
3. Clear the cellular debris by centrifugation for 10 min at 20,000 × *g*, followed by the collection of the soluble protein fraction in the supernatant.

3.2. Protein Sample Preparation

1. Prepare yeast cytosolic proteins by mixing 5 mL of the supernatant with 495 mL of the protein denaturation and reduction buffer. Keep the solution at 37°C under a nitrogen atmosphere for 2 h.
2. Add IAM to a final concentration of 250 mM. Proceed the alkylation reaction at 37°C in the dark for 30 min.
3. Employ a PD-10 size exclusion column for buffer exchange. Elute denatured, reduced, and alkylated proteins using the protein-eluting buffer.
4. Add pharmalyte 3–10 to a final concentration of 1% (v/v). Store the protein solution at −80°C until usage.

3.3. Combined CIEF/RPLC Separations in an Integrated Platform

1. Assemble combined CIEF/RPLC separations in an integrated platform as shown in **Fig. 1** *(9)*.
2. Load a 60-cm long CIEF capillary (100 μm i.d./365 μm o.d.) with the protein solution.
3. Fill the inlet and outlet reservoirs connecting the CIEF capillary with solutions of 0.1 M acetic acid and 0.5% ammonium hydroxide, respectively. Perform protein focusing at electric field strength of 300 V/cm over the entire CIEF capillary (*see* **Note 1**).

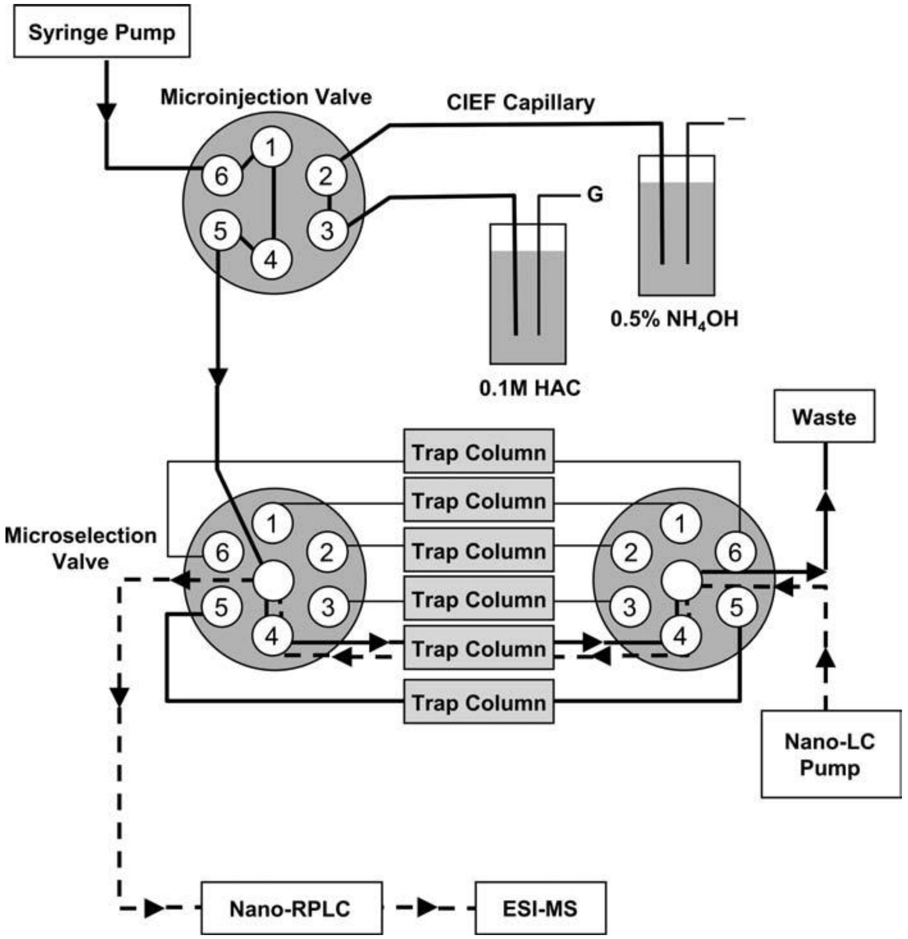

Syringe Pump
Microinjection Valve
CIEF Capillary
1
2
3
4
5
6
G
−
0.1M HAC
0.5% NH_4OH
Waste
Trap Column
Trap Column
Trap Column
Trap Column
Trap Column
Trap Column
Microselection Valve
Nano-LC Pump
Nano-RPLC
ESI-MS

Fig. 1. Schematic of online integration of CIEF with RPLC as a concentrating and multidimensional protein/peptide separation platform. *Solid* and *dashed lines* represent the flow paths for the loading of CIEF fractions and the injection of fractions into a RPLC column, respectively *(9)*.

4. Sequentially and hydrodynamically load focused proteins into an injection loop in a microinjection valve. Inject the loaded proteins into a C_{18} reversed-phase trap column (*see* **Note 2**) using a syringe pump at a flow rate of 2 μl/min through the first six-port microselection valve.
5. Repeat **steps 4** and **5** until the entire CIEF capillary content is sampled into desired number of unique fractions (*see* **Note 3**).
6. Generate a 40-min linear gradient from 10 to 65% acetonitrile at a flow rate of 200 nL/min, using a capillary LC pump. By the use of the second six-port microselection valve, the mobile phase is delivered into the individual trap column, followed by a 13-cm-long C_{18} reversed-phase column (*see* **Note 4**).

7. Repeat **step 6** and sequentially analyze all CIEF fractions collected in the trap columns using RPLC. Monitor the column eluants using ESI-q-TOF MS. Generate electrospray by applying an ESI voltage of 1.8 kV through a platinum electrode housed inside a microcross which is placed in-line with and upstream of the chromatography column. Collect mass spectra from 600 to 1,900 *m/z* using a scan time of 1.0 s. An example result is shown in **Fig. 2**.

8. Analyze protein mass spectra using Protein Trawler software (*see* **Note 5**).

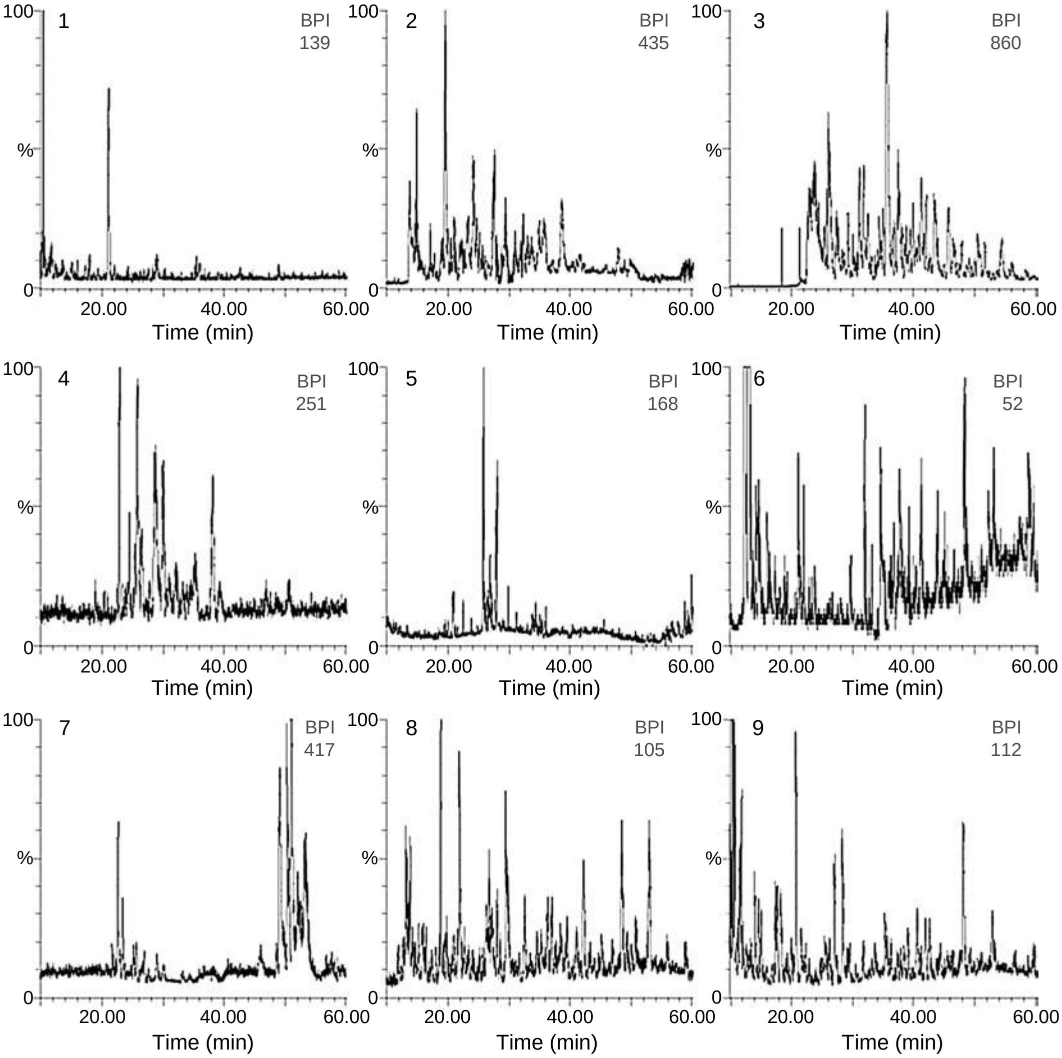


Fig. 2. Base peak chromatograms of a representative CIEF/RPLC multidimensional separation of yeast intact proteins obtained from the soluble fraction of cell lysates. Each *number* represents the sequence of CIEF fractions further analyzed by RPLC from acidic to basic pHs *(9)*.

4. Notes

1. The current decreases continuously as the result of protein focusing. Once the current reduces to ~10% of the original value, usually within 30 min, the focusing is considered to be complete.
2. Trap columns are prepared "in-house." The effluent end of a 8-cm-long fused-silica capillary (100 µm i.d./365 µm o.d.) is connected to an inline microfilter assembly. The C_{18}-bonded particles in methanol are introduced into the capillary by gradually increasing the pressure from 100 to 2,000 psi using a capillary LC pump. The capillary packed with 3 cm of C_{18}-bonded particles is left under pressure for 10 h and then depressurized overnight.
3. By using only two 6-port microselection valves (**Fig. 1**), efforts involve manual and continuous connection and disconnection of individual trap columns during the entire CIEF protein fractionation and RPLC protein separation processes. Thus, there is no need to hold protein fractions in the CIEF capillary. Furthermore, the trap columns containing CIEF fractions, which are not yet subjected to the second dimension RPLC analysis are stored at 4°C.
4. Capillary reversed-phase column is prepared "in-house." The effluent end of a 15-cm-long fused-silica capillary (50 µm i.d./365 µm o.d.) is connected to an inline microfilter assembly. The C_{18}-bonded particles in methanol are introduced into the capillary by gradually increasing the pressure from 100 to 2,000 psi using a capillary LC pump. The capillary packed with 13 cm of C_{18}-bonded particles is left under pressure for 10 h and then depressurized overnight.
5. Approximately 40-min of each RPLC run is divided into eighty 30-s slices. The mass spectra of these 30-s slices are summed and averaged, and Micromass MaxEnt1 is employed as the deconvolution algorithm over a mass range of 5–100 kDa. No baseline subtraction or smoothing is used, as these manipulations are empirically determined to interfere with protein mass deconvolution. Postprocessing of the deconvoluted protein peaks is performed using routines provided with Protein Trawler. Briefly, peaks within 30 Da and occurring within 90-s are only included once and are considered to be the result of bleeding or band broadening of the same protein. Postprocessing also includes the elimination of 1/2 and 2× mass peaks, which are common artifacts of the deconvolution algorithm. An example of deconvolution results is shown in **Fig. 3**.

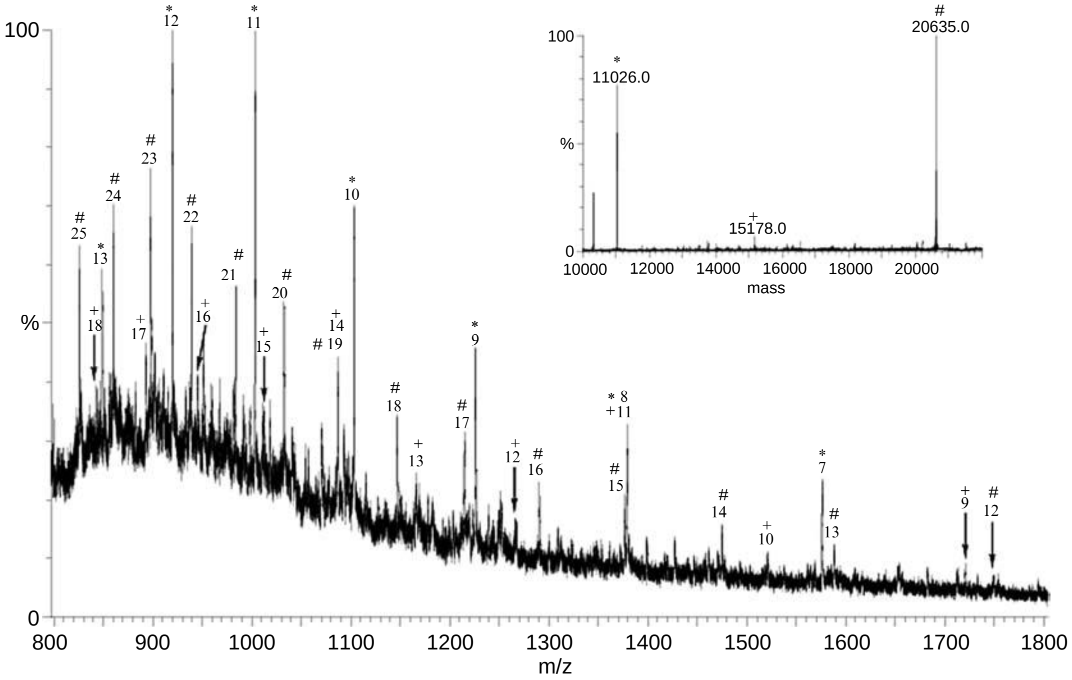


Fig. 3. The mass spectrum of a chromatography peak eluted at 35.7 min for the RPLC separation of the second CIEF fraction. The inset presents deconvoluted protein masses which are labeled together with the corresponding protein envelopes using *asterisk*, *plus*, and *hash,* respectively.

Acknowledgment

We thank the National Cancer Institute (CA103086 and CA107988) and the National Center for Research Resources (RR021239 and RR021862) for supporting portions of this research.

References

1. Amunugama, R., Hogan, J. M., Newton, K. A., and McLuckey, S. A. (2004) Whole protein dissociation in a quadrupole ion trap: identification of a priori unknown modified protein. *Anal. Chem.* 76, 720–727.
2. Meng, F., Du, Y., Miller, L. M., Patrie, S. M., Robinson, D. E., and Kelleher, N. L. (2004) Molecular-level description of proteins from Saccharomyces cerevisiae using quadrupole FT hybrid mass spectrometry for top down proteomics. *Anal. Chem.* 76, 2852–2858.
3. Du, Y., Meng, F., Patrie, S. M., Miller, L. M., and Kelleher, N. L. (2004) Improved molecular weight-based processing of intact proteins for interrogation by quadrupole-enhanced FT MS/MS. *J. Proteome Res.* 3, 801–806.
4. Kelleher, N. L. (2004) Top-down proteomics. *Anal. Chem.* 76, 197A–203A.
5. Loo, J. A., Brown, J., Critchley, G., Mitchell, C., Andrews, P. C., and Ogorzalek Loo, R. R. (1999) High sensitivity mass spectrometric methods for obtaining intact molecular weights from gel-separated proteins. *Electrophoresis* 20, 743–748.
6. Tang, Q., Harrata, A. K., and Lee, C. S. (1997) Two-dimensional analysis of recombinant *E. coli* proteins using capillary isoelectric

focusing electrospray ionization mass spectrometry. *Anal. Chem.* 69, 3177–3182.

7. Lee, S. W., Berger, S. J., Martinovic, S., Paša-Toli , L., Anderson, G. A., Shen, Y., Zhao, R., and Smith, R. D. (2002) Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc. Natl Acad. Sci. USA* 99, 5942–5947.
8. Wang, Y., Balgley, B. M., Rudnick, P. A., Evans, E. L., DeVoe, D. L., and Lee, C. S. (2005) Integrated capillary isoelectric focusing/nano-reversed phase liquid chromatographycoupledwithESI-MSforcharacterization of intact yeast protein. *J. Proteome Res.* 4, 36–42.
9. Chen, J., Balgley, B. M., DeVoe, D. L., and Lee, C. S. (2003) Capillary isoelectric focusing-based multidimensional concentration/separation platform for ultrasensitive proteome analysis. *Anal. Chem.* 75, 3145–3152.

Chapter 14

Integrating Accelerated Tryptic Digestion into Proteomics Workflows

Gordon W. Slysz and David C. Schriemer

Summary

An accelerated protein digestion procedure is described that features a microscale trypsin cartridge operated under aqueous-organic conditions. High sequence coverage digestions obtained in seconds with small amounts of enzyme are possible with the approach, which also supports online integration of digestion with reversed-phase protein separation. The construction and operation of effective digestor cartridges for rapid sample processing are described. For workflows involving chromatographic protein separation an easily assembled fluidic system is presented, which inserts the digestion step after column-based separation. Successful integration requires dynamic effluent titration immediately prior to transmission through the digestor. This is achieved through the co-ordination of the column gradient system with an inverse gradient system to produce steady pH and organic solvent levels. System assembly and operation sufficient for achieving digestion and identification of subnanogram levels of protein are described.

Key words: Proteomics, Rapid protein identification, Online trypsin digestion, Integrated top-down and bottom-up approaches.

1. Introduction

The enzymatic conversion of protein to peptides is an essential component of most proteomic workflows involving protein identification and characterization. Traditional in-solution digestions are slow, requiring several hours and they often perform poorly with dilute protein samples or with resistant, tightly folded proteins. Methods that reduce processing time and lead to more efficient digestion add considerable value to proteomics when high throughput and sensitivity are required *(1, 2)*.

Strategies used to improve digestion focus on both protein substrate and trypsin. Making the substrate more amenable to

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_14

digestion via the reduction and alkylation of disulfide bonds is common in most procedures. This helps unlock the tertiary and quaternary structure of the protein, where greater conformational freedom improves access to all possible sites of hydrolysis *(3, 4)*. This can be assisted through the use of chaotropes such as urea, guanidinium chloride, detergents, and organic solvents *(1, 5)*. Greater rates of hydrolysis are also possible using higher enzyme concentrations at the expense of increased trypsin autolysis products, which in turn can be minimized through immobilization strategies. POROS beads *(6–8)*, monolithic supports *(9)*, membranes *(10)*, nanovials *(11)*, and MALDI plates *(12)* are several materials on which trypsin has been immobilized for proteomic applications. Digestion enhancement techniques involving sonication or microwaves likely promote digestion improvement through a combination of substrate denaturation and trypsin activation *(13–15)*. Improved trypsin activity for in-gel and in-solution digestion formats has benefited from temperatures approaching 60°C *(16, 17)*. Though trypsin turnover rates increase at higher temperatures, its reduced stability has prompted research into the generation of thermostable forms *(18)*.

The method presented here involves organic solvents and immobilized trypsin to enhance rates of hydrolysis in a microreactor-based format. The reactor is small and can be incorporated easily into capillary-based microfluidic system (**Fig. 1**). It houses commercially available immobilized trypsin beads (POROSzyme, Applied Biosystems) and presents a trypsin concentration of approximately 10 mg/mL *(6)*. Digestion is achieved in 45% acetonitrile buffered to pH 8.2. Immobilized trypsin remains maximally active over many hours under these conditions, even under elevated temperatures.

The efficiency of this method is highlighted by the digestion of a modest concentration (250 nM) of transferrin, a highly

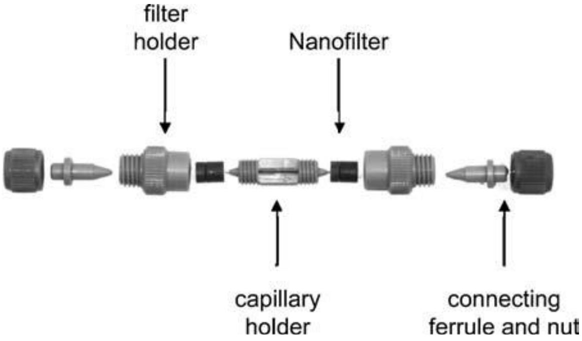


Fig. 1. Photograph of digestor. Fused silica capillary is housed in a column holder. Beads are held in place by a 1-μm titanium nanofilter. The digestor can be easily attached to fluidic lines using low dead volume fittings.

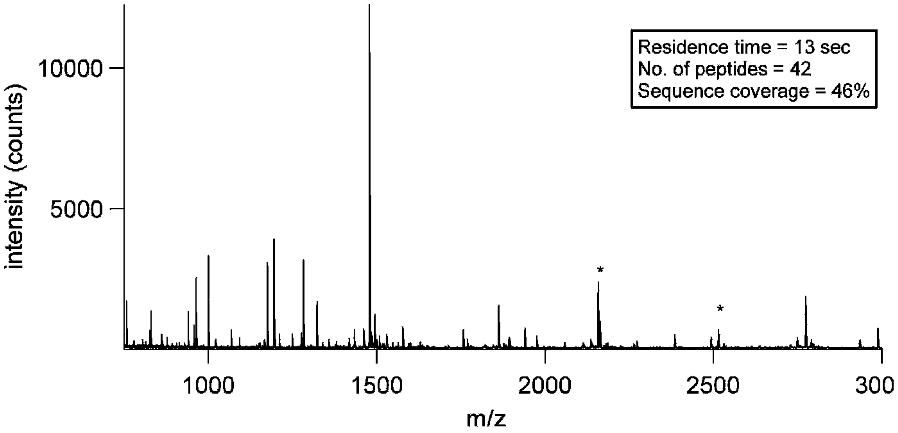


Fig. 2. Peptide mass fingerprint of transferrin generated from cartridge-based trypsin digestion. 250 nM transferrin was dissolved in 20 mM ammonium acetate in 45% ACN and infused through a 150 μm ID × 2.5 cm trypsin cartridge at 1 μL/min. Digest was combined with HCCA matrix (50:50) and spotted on a MALDI plate. *Asterisks* show trypsin autolysis peaks. No reduction and alkylation was performed.

resistant protein containing 19 disulfide bonds *(8)*. More than 40 tryptic peptides were detected following MALDI-TOF detection, as compared to only 8 using typical overnight trypsin digestion (**Fig. 2**). In other studies, near-complete sequence coverage was achieved following digestion of small amounts of myoglobin, cytochrome C, and carbonic anhydrase *(19)*. These results compare well with several other approaches in the literature which show improved digestions of *large* concentrations of protein, but less-than-optimal sequence coverages of when dealing with low concentrations and amounts of protein.

The highly efficient nature of this digestion motivated its insertion directly downstream of a protein separation. This has utility in that branch of gel-free proteomics involving the separation of proteins using multidimensional liquid chromatography *(3, 20)*. Promising results in the analysis of plasma *(21)* and bacterial cell lysates *(22, 23)* illustrate the potential of these protein-based ("top-down") separations. However, protein identifications are usually made at the peptide level, requiring a "bottom-up" approach. In most protein-based LC systems, eluted proteins are collected as fractions and undergo multiple handling steps prior to analysis. Each sample-handling step negatively affects yield and adds to sample analysis times *(24)*. Insertion of an efficient digestor directly downstream of the protein-based separation simplifies the workflow and although generating a small "dilution penalty," it supports sensitive and rapid bottom-up analysis *(19, 25)*.

This chapter describes the construction and operation of the digestor, and how to insert it into a reversed-phase protein separation for subnanogram limits of detection.

2. Materials

2.1. Cartridge-Based Digestion

2.1.1. Digestor Construction

1. Fused silica capillary, various internal diameters (100–250 µm ID, 360 µm OD).
2. POROSzyme immobilized trypsin (Applied Biosystems).
3. Cartridge assembly (2.5 cm length), for trapping the beads and housing the fused silica capillary (**Fig. 1**). The entire cartridge assembly (part C-1700) and its parts can be purchased from Upchurch Scientific. Parts include the fused silica capillary holder, filter holders, NanoFilters, and connecting ferrule. For custom-length cartridges, beads can be trapped using an inline NanoFilter assembly (part M-537).
4. Hamilton gastight syringe (100 µL) for packing trypsin beads. Use a microinjection port assembly (part M-432) to connect syringe to fused silica capillary.

2.1.2. Solutions and Flow

1. Digestion buffer: 20 mM ammonium acetate, pH 8.2, in 45% acetonitrile (HPLC grade) (*see* **Notes 1 and 2**).
2. Harvard syringe pump and Hamilton syringe for infusion of protein sample through the digestor.

2.2. Integrated Digestion in an LC/MS System

2.2.1. LC-Digestor-MS Construction

1. Two binary pumps, each capable of delivering nanoflow gradients. Options: splitless nanofluidic systems from Upchurch Scientific or Eksigent.
2. One splitless nanofluidic module (Upchurch Scientific), capable of steady 100 nL/min flow rates. This is used to reacidify the gradient/countergradient flow before introduction into the mass spectrometer. A syringe pump is an alternative, though flow-rate stability on such pumps is not high.
3. Two mixing tees (Part P-888, Upchurch Scientific), one for combining the gradient/countergradient streams and one for combining this blended with a reacidification stream prior to MS analysis.
4. 50-µm ID-fused silica capillary (Polymicro Technologies, Phoenix, AZ) for all connections between gradient systems, columns, and the mass spectrometer.
5. Reversed phase column for protein separation. Example: C4 beads (300 Å, 5 µm) packed in-house in 75-µm ID-fused silica capillary (*see* **Note 3**)
6. Optional: Column heater (model CH-30, Fiatron Systems, Milwaukee, WI) set at 37°C, for housing reversed-phase and immobilized trypsin columns, transfer lines, and mixers.

2.2.2. Solutions Required for LC-Digestor-MS System (see Note 4)

1. Gradient 1, mobile phase A (G1-A): 0.05% formic acid in water (HPLC grade).
2. Gradient 1, mobile phase B (G1-B): 0.05% formic acid in 90% ACN (HPLC grade).
3. Gradient 2 (countergradient), mobile phase A (G2-A): 20 mM ammonium acetate in water, titrated with ammonium hydroxide (*see* **Subheading 3.2.2**).
4. Gradient 2 (countergradient), mobile phase B (G2-B): 20 mM ammonium acetate in 90% ACN, titrated with ammonium hydroxide (*see* **Subheading 3.2.2**).

2.3. Mass Spectrometry

1. A tandem mass spectrometer configured for nanospray applications. We have used both a QSTAR Pulsar *i* QqTOF mass spectrometer and an Agilent XCT Ultra ion trap mass spectrometer.

3. Methods

3.1. Cartridge-Based Tryptic Digestion

3.1.1. General Operation

Efficient online trypsin digestions are achieved by infusing the protein sample, dissolved in digestion buffer at a rate of 1 μL/min through a 2.5 cm × 150 μm trypsin cartridge. **Figure 2** shows the results of an online digestion of 250 nM transferrin, a highly resistant protein containing 19 disulfide bonds, without prior disulfide reduction. Over 40 peptides were detected, comparing well to traditional in-solution methods that resulted in only 8 peptides *(8)*.

1. Pack a dilute slurry (>10:1 v/v of aqueous buffer to beads) of immobilized trypsin beads into the fused silica capillary, housed in a column assembly (*see* **Note 5**).
2. Connect trypsin reactor to the syringe pump. Equilibrate with digestion buffer at a flow rate of 1–2 μL/min for 20 min to reduce trypsin autolysis products (*see* **Note 6**).
3. Prepare protein sample in the digestion buffer. Load into syringe and flow at rates 0.1–20 μL/min, depending on desired level of digestion (*see* **Note 7**). Alternatively, smaller volumes of sample can be injected on cartridge through a suitable injection valve.
4. The protein digest can be collected as a fraction, or the outflow of the digestor can be reacidified using a second syringe pump and directly interfaced to an electrospray mass spectrometer.There is a relationship between digestion efficiency, flow rate and reactor dimensions. A high-quality

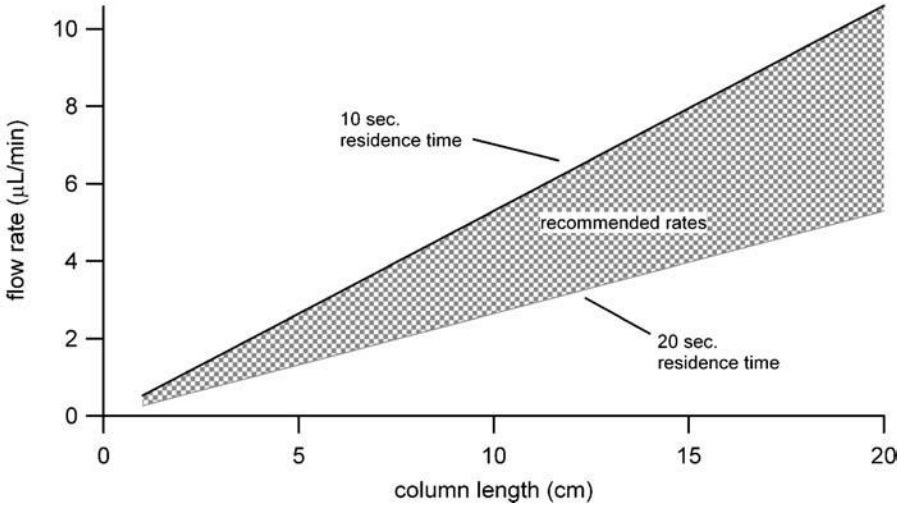


Fig. 3. Length selection for a 150-μm ID cartridge. Excellent digestions are generated using cartridges with lengths that give residence times greater than 10 s. Suggested lengths for generating thorough digestions while minimizing digestor size are shown in the *hatched area*.

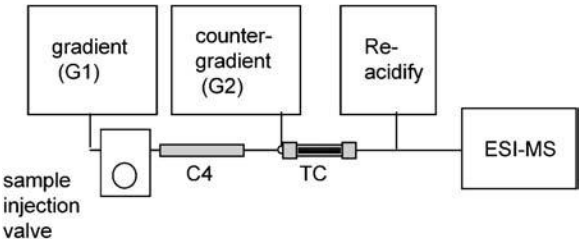


Fig. 4. The NanoLC-trypsin-MS system. Intact protein is loaded onto the C4 reversed-phase column and eluted under gradient conditions (G1). The countergradient (G2) stream titrates the effluent for optimal digestion by the trypsin cartridge (TC). The flow is reacidified and analyzed by a mass spectrometer using electrospray ionization.

digestion is achieved at 1 μL/min through a 150 μm × 2.5 cm trypsin cartridge. Other configurations are also possible, but require adapting the flow rate. Excellent digestion results from a 10–20 s reactor residence time and with the aid of **Fig. 3** the appropriate configuration can be readily assembled *(25)*. This very short residence time provides the basis for successful incorporation downstream of protein chromatography, the operation of which will be described below.

3.2. NanoLC-trypsin-MS set-up

3.2.1. General Description

The NanoLC-trypsin-MS system is configured as in **Fig. 4**. Proteins are captured on a C4 reversed-phase column and eluted under conventional gradient conditions. To produce optimal conditions for digestion, the column effluent requires titration to pH 8.2 and dilution to constant organic solvent composition. This is accomplished using a second gradient connected at the head of the trypsin column. Postdigestion, the peptides are reacidified using an isocratic pump prior to electrospray mass spectrometry.

3.2.2. Gradient and Countergradient Solutions

Preparation of the gradient solutions (G1-A, G1-B) and countergradient solutions (G2-A, G2-B) is guided by the requirement of pH 8.2 for equal-volume combinations of G1 and G2. This can be achieved in the following manner:

1. Prepare G1-A and G1-B (*see* **Subheading 3.2.2**).
2. Prepare a 10× ammonium acetate stock solution (200 mM). Titrate 20 mL of this solution with 10% ammonium hydroxide (recommended starting volume of 4.2 mL). Prepare G2-A and G2-B from this stock (*see* **Subheading 2.2.2**). Combine equal volumes of all four solutions and measure pH. If the pH is not within 8–8.2, titrate the ammonium acetate stock solution as required, and repeat. Since 10% ammonium hydroxide is volatile and ammonium acetate is hygroscopic, this empirical approach is more useful than specifying accurate amounts of reagent.

3.2.3. Microfluidics

1. Make all connections using narrow-bore (50 μm) fused silica capillary. For experiments requiring analysis of sub-pmol levels of protein, passivating the inner surface of fused silica capillary may be required to minimize excessive nonspecific binding of proteins *(19, 26)* (*see* **Note 8**). This is particularly necessary for elements of the flowpath that experience solutions with pH >4.
2. Insert the reversed phase column between the two gradient chromatography systems as outlined in **Fig. 4**. Establish flow on each gradient system to 300 nL/min, using 50% A compositions in both pumps.
3. Upon equilibration, insert the trypsin microreactor directly into the mixing tee downstream of the reversed phase column (*see* **Note 9**). Attach the reactor outlet to the mass spectrometer through the second tee, which brings in a stream of acidified acetonitrile to promote ionization (recommended rate: 100 nL/min).
4. Option: House the reversed-phase column, immobilized trypsin column, mixing unions, and fused silica capillary within a column heater set at 37°C.

3.2.4. Operation

Gradient/Countergradient Co-ordination

Maintaining a constant flow of 45% acetonitrile through the immobilized trypsin column is key to obtaining high digestion efficiency. Therefore, the gradient and countergradient timetables must be coordinated. Consider a simple linear gradient of 5–90% G1-B for reversed phase chromatography. This should be twinned with a countergradient from 95 to 10% G2-B. Because of the gradient delay due to the reversed-phase column, the onset of the countergradient needs to be delayed for proper matching. This delay time is easily measured by monitoring total ion current changes due to solvent. One such method of delay time measurement is as follows:

1. Replace the immobilized trypsin reactor with a union or a dummy reactor.
2. Set the flow at 5% G1-B in the reversed-phase gradient and 95% G2-B in the countergradient, and equilibrate the flow path. Simultaneously switch the reversed-phase gradient to 95% G1-B and the countergradient to 5% G2-B.
3. The solvent change from the countergradient reaches the mass spectrometer first. The ion current typically drops in a highly aqueous mixture. When the reversed-phase gradient reaches the mass spectrometry, the ion current returns to its original level. The time difference between these two changes provides an accurate measurement of the required countergradient delay.

Trypsin Column Insertion and Equilibration (*see* Note 9)

1. Establish start compositions for both gradient pumps (e.g., reversed-phase gradient at 5% G1-B and the countergradient at 95% G2-B) and set the desired flow rate (e.g., 300 nL/min).
2. Insert a fresh trypsin reactor and upon observation of flow through the device, attach to the reacidification stream.
3. Equilibrate at least 20 min at the desired flow rate to reduce trypsin autolysis (*see* **Note 6**).

Protein Analysis

1. Upon equilibration, inject protein dissolved in G1-A solution.
2. Elute proteins from reversed-phase column using an appropriate gradient for the column of choice.
3. Start the countergradient after the predetermined delay.

3.2.5. Performance (see Note 10*)*

Proteins are separated, but are detected by the mass spectrometer as peptides. Therefore, integrating over a feature in the extracted ion chromatogram will generate the peptide mass fingerprint for that protein (**Fig. 5**). To help gauge performance, **Table 1** provides a list of peptides generated from 1 pmol bovine carbonic anhydrase injected in the above-configured system and detected by a Qstar Pulsar i QqTOF instrument.

Separating at the protein level offers the potential to resolve proteins existing in different states. For example, **Fig. 6** shows how the nanoLC-trypsin-MS method could be used to observe free and ubiquitin-modified protein superoxide dismutase as contaminants in a solution of four standard proteins. Extracted

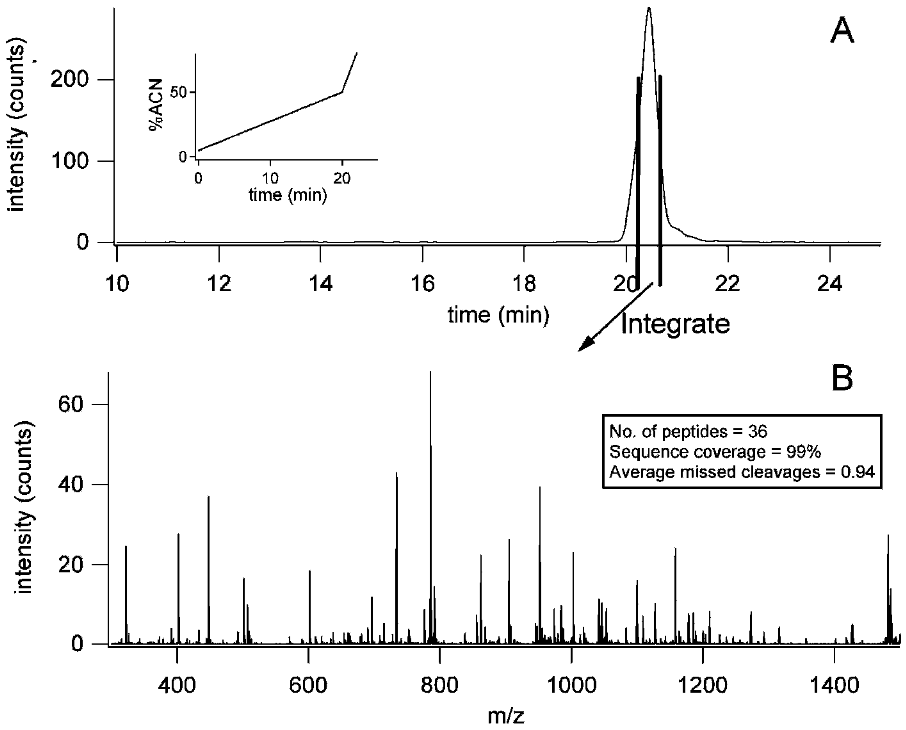


Fig. 5. NanoLC-trypsin-MS analysis of 1 pmol carbonic anhydrase (bovine) as performance standard. **a** Protein chromatogram from representative peptide ion current; inset: gradient for C4-based separation; **b** peptide mass fingerprint at the retention time for carbonic anhydrase.

Table 1
Bovine carbonic anhydrase peptides detected after nanoLC-trypsin-MS analysis of 1 pmol

Residues	m.c.[a]	Peptide sequence
1–8	0	Ac-VLDALDSIKTK
1–26	2	Ac-SHHWGYGKHNGPEHWHKDFPIANGER
9–17	0	HNGPEHWHK
9–26	1	HNGPEHWHKDFPIANGER
18–26	0	DFPIANGER
27–35	0	QSPVDIDTK
36–56	0	AVVQDPALKPLALVYGEATSR
36–57	1	AVVQDPALKPLALVYGEATSRR
57–79	2	RMVNNGHSFNVEYDDSQDKAVLK
57–88	3	RMVNNGHSFNVEYDDSQDKAVLKDGPLTGTYR

(continued)

Table 1
Continued

Residues	m.c.[a]	Peptide sequence
58–75	0	MVNNGHSFNVEYDDSQDK
58–79	1	MVNNGHSFNVEYDDSQDKAVLK
58–88	2	AVLKDGPLTGTYR
76–88	1	DGPLTGTYR
80–88	0	LVQFHFHWGSSDDQGSEHTVDR
89–110	0	LVQFHFHWGSSDDQGSEHTVDR
89–111	1	LVQFHFHWGSSDDQGSEHTVDRK
89–112	2	LVQFHFHWGSSDDQGSEHTVDRKK
112–147	2	YAAELHLVHWNTK
113–125	0	YAAELHLVHWNTK
113–147	1	YAAELHLVHWNTKYGDFGTAAQQPDGLAVVGVFLK
113–157	2	YAAELHLVHWNTKYGDFGTAAQQPDGLAVVGVFLKVG-DANPALQK
126–147	0	YGDFGTAAQQPDGLAVVGVFLK
126–157	1	YGDFGTAAQQPDGLAVVGVFLKVGDANPALQK
126–168	3	YGDFGTAAQQPDGLAVVGVFLKVGDANPALQKVLDAL-DSIKTK
148–166	1	VGDANPALQKVLDALDSIK
148–168	2	VGDANPALQKVLDALDSIKTK
171–223	1	STDFPNFDPGSLLPNVLDYWTYPGSLTTPPLLESVTWIV-LKEPISVSSQQMLK
224–250	1	FRTLNFNAEGEPELLMLANWRPAQPLK
224–252	2	FRTLNFNAEGEPELLMLANWRPAQPLKNR
226–250	0	TLNFNAEGEPELLMLANWRPAQPLK
226–252	1	TLNFNAEGEPELLMLANWRPAQPLKNR
253–255	0	QVR
256–259	0	GFPK

[a] Missed cleavages

ion chromatograms of superoxide dismutase-specific and ubiquitin-specific peptides each resulted in two peaks, with one peak from ubiquitin (peak "B" of Fig. 6) overlapping with a peak from superoxide dismutase (peak "C" of Fig. 6). This suggests that superoxide dismutase existed both in the free state and bound to ubiquitin.

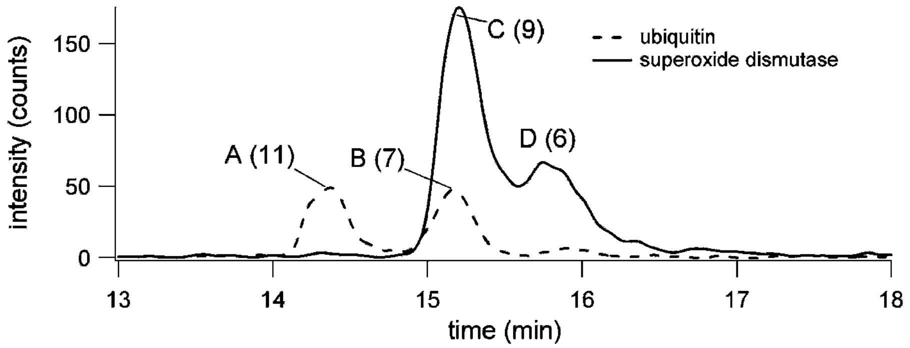


Fig. 6. Ubiquitylation detection using LC-trypsin-MS. Analysis of a protein mixture revealed two contaminants – ubiquitin and superoxide dismutase (SOD). Extracted ion chromatograms of ubiquitin-specific peptides resulted in two peaks, labeled "A" and "B." Extracted ion chromatogram of SOD-specific peptides shows peaks "C" and "D." The number of peptides detected in elution-time-based peptide mass fingerprints is indicated in *parentheses*. These data suggest that SOD was both free (peak "D") and labeled (peak "C") with ubiquitin. Free ubiquitin was also present (peak "A").

4. Notes

1. The pH is the most critical factor in achieving good digests. Organic solvent influences pH measurement, but are presented as uncorrected values in this chapter. To duplicate performance, pH adjustments should be made on 45% acetonitrile solutions. A typical pH probe can be used, calibrated using an aqueous solution as a reference *(27)*.
2. Ammonium acetate is preferred over ammonium bicarbonate. The pH of ammonium bicarbonate drifts day-to-day, even when stored at 4°C. The pH of ammonium acetate remains steady over several weeks at 4°C and although the buffering capacity is not high, it is sufficient for the application.
3. C4 columns were packed in-house using a pressurized cell (Brechbuhler, Inc., Spring, TX). Commercial capillary columns for protein separation can also be used.
4. Solutions for the gradient/countergradient systems should be stored in sealed glass Pyrex containers with Teflon caps, for example. Solution G2-B is particularly volatile and if not adequately sealed, the pH upon titration will drift downward.
5. POROSzyme beads average 20 μm in diameter and have a macroporous structure *(6)*. As a result, trypsin reactors have low backpressures and can be easily packed by hand using a Hamilton syringe into the empty cartridge. To pack, preassemble the cartridge with a longer silica capillary (e.g., 15 cm by 150 μm ID) and connect to the syringe via the injection port adapter (part M-432, Upchurch). After packing to the desired bed depth, cut to length and insert the top frit of the cartridge.

6. When new trypsin reactors are first used under mixed solvent digestion conditions, substantial autolysis products are generated. This is believed to arise from both trypsin bleed and some measure of increased fluidity in the protein stationary phase under mixed solvent conditions. In our system, signals decreased to relatively low levels over a period of 20 min. However, when generating digests from low levels of protein substrate, autolysis peaks may still dominate.
7. Proteins denature in mixed aqueous-organic solvents *(28)*. However, protein solubility in aqueous-organic mixtures may suffer *(29)* and denatured proteins have a greater propensity toward precipitation. Substrate concentration should be maintained at lower levels as a result (less than 10 μM), or alternative organic solvents used (e.g., 2-propanol). If solvent-induced precipitation is an issue, the online version of our digestion method should be used (*see* **Subheading 3.2**).
8. Untreated fused silica capillary strongly retains protein under neutral pH. The passivation method of Belder et al. *(26)* involving in situ generation of cross-linked poly(vinyl alcohol) reduces adsorption and leads to detection limits as low as 5 fmol *(19)*. Other strategies involving dynamic or permanent capillary coatings may also be useful (e.g., addition of dilute EDTA). Many such methods have been recently reviewed *(30, 31)*. Coated capillaries are also commercially available.
9. Trypsin reactors will tolerate fluctuations in organic solvent composition, but not fluctuations in pH. Ensure that both pumps always deliver equal volumes as dynamic titration occurs at the head of the reactor. The trypsin reactor may recover after a pulse of acid or base, but requires several minutes for an acidic pulse, while the alkaline pulse results in a much slower and incomplete recovery. When optimizing flow conditions, simply remove the reactor and insert after equilibration.
10. With new trypsin columns, the first injection and analysis of protein sample usually produced poor results (relatively low signal). This "first-run effect" is likely due to nonspecific binding within the system. Second runs typically showed excellent results.

References

1. Klammer, A. A., and MacCoss, M. J. (2006) Effects of modified digestion schemes on the identification of proteins from complex mixtures. *J. Proteome Res.* 5, 695–700.
2. Lopez-Ferrer, D., Canas, B., Vazquez, J., Lodeiro, C., Rial-Otero, R., Moura, I., and Capelo, J. L. (2006) Sample treatment for protein identification by mass spectrometry-based techniques. *TrAC-Trends Anal. Chem.* 25, 996–1005.
3. Stults, J. T., and Arnott, D. (2005) Proteomics. *Methods Enzymol* 402, 245–289.

4. Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) Disulfide bonds and protein folding. *Biochemistry* 39, 4207–4216.
5. Lundell, N., and Schreitmuller, T. (1999) Sample preparation for peptide mapping – A pharmaceutical quality-control perspective. *Anal. Biochem.* 266, 31–47.
6. Hsieh, Y. L. F., Wang, H. Q., Elicone, C., Mark, J., Martin, S. A., and Regnier, F. (1996) Automated analytical system for the examination of protein primary structure. *Anal. Chem.* 68, 455–462.
7. Wang, S. H., and Regnier, F. E. (2001) Proteolysis of whole cell extracts with immobilized enzyme columns as part of multidimensional chromatography. *J. Chromatogr. A* 913, 429–436.
8. Slysz, G. W., and Schriemer, D. C. (2003) On-column digestion of proteins in aqueous-organic solvents. *Rapid Commun. Mass Spectrom.* 17, 1044–1050.
9. Peterson, D. S., Rohr, T., Svec, F., and Frechet, J. M. J. (2002) Enzymatic microreactor-on-a-chip: Protein mapping using trypsin immobilized on porous polymer monoliths molded in channels of microfluidic devices. *Anal. Chem.* 74, 4081–4088.
10. Cooper, J. W., Chen, J. Z., Li, Y., and Lee, C. S. (2003) Membrane-based nanoscale proteolytic reactor enabling protein digestion, peptide separation, and protein identification using mass spectrometry. *Anal. Chem.* 75, 1067–1074.
11. Litborn, E., Emmer, A., and Roeraade, J. (1999) Chip-based nanovials for tryptic digest and capillary electrophoresis. *Anal. Chim. Acta* 401, 11–19.
12. Harris, W. A., and Reilly, J. P. (2002) On-probe digestion of bacterial proteins for MALDI-MS. *Anal. Chem.* 74, 4410–4416.
13. Lopez-Ferrer, D., Capelo, J. L., and Vazquez, J. (2005) Ultra fast trypsin digestion of proteins by high intensity focused ultrasound. *J. Proteome Res.* 4, 1569–1574.
14. Pramanik, B. N., Mirza, U. A., Ing, Y. H., Liu, Y. H., Bartner, P. L., Weber, P. C., and Bose, M. K. (2002) Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: A new approach to protein digestion in minutes. *Protein Sci.* 11, 2676–2687.
15. Sun, W., Gao, S. J., Wang, L. J., Chen, Y., Wu, S. Z., Wang, X. R., Zheng, D. X., and Gao, Y. H. (2006) Microwave-assisted protein preparation and enzymatic digestion in proteomics. *Mol. Cell. Proteomics* 5, 769–776.
16. Havlis, J., Thomas, H., Sebela, M., and Shevchenko, A. (2003) Fast-response proteomics by accelerated in-gel digestion of proteins. *Anal. Chem.* 75, 1300–1306.
17. Finehout, E. J., Cantor, J. R., and Lee, K. H. (2005) Kinetic characterization of sequencing grade modified trypsin. *Proteomics* 5, 2319–2321.
18. Sebela, M., Stosova, T., Havlis, J., Wielsch, N., Thomas, H., Zdrahal, Z., and Shevchenko, A. (2006) Thermostable trypsin conjugates for high-throughput proteomics: synthesis and performance evaluation. *Proteomics* 6, 2959–2963.
19. Slysz, G. W., Lewis, D. F., and Schriemer, D. C. (2006) Detection and identification of sub-nanogram levels of protein in a nanoLC-trypsin-MS system. *J. Proteome Res.* 5, 1959–1966.
20. Roe, M. R., and Griffin, T. J. (2006) Gel-free mass spectrometry-based high throughput proteomics: Tools for studying biological response of proteins and proteomes. *Proteomics* 6, 4678–4687.
21. Linke, T., Ross, A. C., and Harrison, E. H. (2006) Proteomic analysis of rat plasma by two-dimensional liquid chromatography and matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *J. Chromatogr. A* 1123, 160–169.
22. Simpson, D. C., Ahn, S., Pasa-Tolic, L., Bogdanov, B., Mottaz, H. M., Vilkov, A. N., Anderson, G. A., Lipton, M. S., and Smith, R. D. (2006) Using size exclusion chromatography-RPLC and RPLC-CIEF as two-dimensional separation strategies for protein profiling. *Electrophoresis* 27, 2722–2733.
23. Sharma, S., Simpson, D. C., Tolic, N., Jaitly, N., Mayampurath, A. M., Smith, R. D., and Pasa-Tolic, L. (2007) Proteomic profiling of intact proteins using WAX-RPLC 2-D separations and FTICR mass spectrometry. *J. Proteome Res.* 6, 602–610.
24. Ethier, M., Hou, W. M., Duewel, H. S., and Figeys, D. (2006) The proteomic reactor: A microfluidic device for processing minute amounts of protein prior to mass spectrometry analysis. *J. Proteome Res.* 5, 2754–2759.
25. Slysz, G. W., and Schriemer, D. C. (2005) Blending protein separation and peptide analysis through real-time proteolytic digestion. *Anal. Chem.* 77, 1572–1579.
26. Belder, D., Deege, A., Husmann, H., Kohler, F., and Ludwig, M. (2001) Cross-linked

poly(vinyl alcohol) as permanent hydrophilic column coating for capillary electrophoresis. *Electrophoresis* 22, 3813–3818.

27. Espinosa, S., Bosch, E., and Roses, M. (2000) Retention of ionizable compounds on HPLC. 5. pH scales and the retention of acids and bases with acetonitrile-water mobile phases. *Anal. Chem.* 72, 5193–200.
28. Griebenow, K., and Klibanov, A. M. (1996) On protein denaturation in aqueous-organic mixtures but not in pure organic solvents. *J. Am. Chem. Soc.* 118, 11695–11700.
29. Ruckenstein, E., and Shulgin, I. L. (2006) Effect of salts and organic additives on the solubility of proteins in aqueous solutions. *Adv. Colloid Interface Sci.* 123, 97–103.
30. Liu, J. K., and Lee, M. L. (2006) Permanent surface modification of polymeric capillary electrophoresis microchips for protein and peptide analysis. *Electrophoresis* 27, 3533–3546.
31. Doherty, E. A. S., Meagher, R. J., Albarghouthi, M. N., and Barron, A. E. (2003) Microchannel wall coatings for protein separations by capillary and chip electrophoresis. *Electrophoresis* 24, 34–54.

Chapter 15

Hydrogen/Deuterium Exchange Mass Spectrometry

Xuguang Yan and Claudia S. Maier

Summary

Amide hydrogen/deuterium (H/D) exchange of proteins monitored by mass spectrometry has established itself as a powerful method for probing protein conformational dynamics and protein interactions. The method uses isotope labeling to probe the rate at which protein backbone amide hydrogens undergo exchange. Backbone amide hydrogen exchange rates are particularly sensitive to hydrogen bonding; hydrogen bonding slows the exchange rates dramatically. Exchange rates reflect on the conformational mobility, hydrogen bonding strength, and solvent accessibility in protein structure. Mass spectrometric techniques are used to monitor the exchange events as mass shifts that arise through the incorporation of deuterium into the protein. Global conformational information can be deduced by monitoring the exchange profiles over time. Combining the labeling experiment with proteolysis under conditions that preserve the exchange information allows for localizing exchange events to distinct regions of the protein backbone and thus, the study of protein conformation with medium spatial resolution. Over the past decade, H/D exchange mass spectrometry has evolved into a versatile technique for investigating conformational dynamics and interactions in proteins, protein–ligand and protein–protein complexes.

Key words: Deuterium, Conformation, Folding, Mass spectrometry, Electrospray.

1. Introduction

The application of hydrogen/deuterium exchange labeling approaches in combination with mass spectrometry for studying protein dynamics, structures, and interactions was first demonstrated in 1991 by Katta and Chait *(1)*. The approach is based on the fact that hydrogen bonding between amino acid side chains and protein backbone amides retard exchange whereas amides that are located in regions that are not stabilized by hydrogen bonding will exchange more readily *(2- 3)*. Also, amides buried in

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_15

the protein core and/or located in interaction surfaces are protected against hydrogen exchange due to the limited solvent exposure, and will therefore show slower exchange kinetics *(4-5)*. Mass spectrometry is used for determining the mass shifts induced by the exchange of hydrogens by deuteriums (or vice versa) with time, as a consequence of changed environments, structure, and complex formation *(6-7)*.

Generally, the level of exchanged hydrogens reflects the flexibility, solvent accessibility, and hydrogen bonding interactions in protein structures. While the rate of exchange of amide hydrogens varies over a large range (i.e., 10^8), depending mostly on their solvent accessibility and hydrogen bonding status, hydrogens within the functional groups exchange too fast to be measured by most techniques. The information on protein backbone structure and conformational dynamics is therefore provided by the presence of the hydrogen on the amide nitrogens. Amide hydrogen exchange can be either acid- or base-catalyzed *(2, 8)*. The exchange reaction is slowest at pH 2–3, which is about 10^3–10^4 times slower then the rate around pH 7 *(9)*. In addition, the exchange rate is temperature dependent, decreasing by approximately a factor of ten as the temperature is reduced from 25 to 0°C. Consequently, under pH 2–3 and at 0°C (commonly referred to as "quench conditions") the half-life for amide hydrogen isotopic exchange in an unstructured polypeptide is 30–90 min, depending on the solvent shielding effect caused by the side chains *(10)*. This sensitivity to pH and temperature is the basis for conducting exchange under a given set of conditions and arresting it under another, thereby allowing the sample to be processed for mass spectrometric analysis *(6, 11)*. Labeled proteins are analyzed intact by LC-ESI-MS to obtain global exchange profiles (**Fig. 1**). The deuterium levels in proteins and peptides are obtained from the increase in mass after exchange-in experiments. Alternatively, exchange-out approaches can be designed that use deuterated protein samples as starting materials and yield exchange profiles showing a reduction in mass over exchange-out time *(12)*.

H/D exchange labeling approaches are frequently combined with proteolysis under conditions that maintain the labeling information in order to obtain exchange maps with medium spatial resolution (**Fig. 1**). This method was introduced by Smith and coworkers in 1993 and has since been applied to numerous protein structural studies, protein folding and unfolding studies, and protein interaction studies *(6)*. By implementing peptic proteolysis, peptides are generated under conditions that retain the labeling information. The centroids of the isotopic clusters are used to determine the deuterium content. The use of hydrogen exchange labeling studies and mass spectrometric

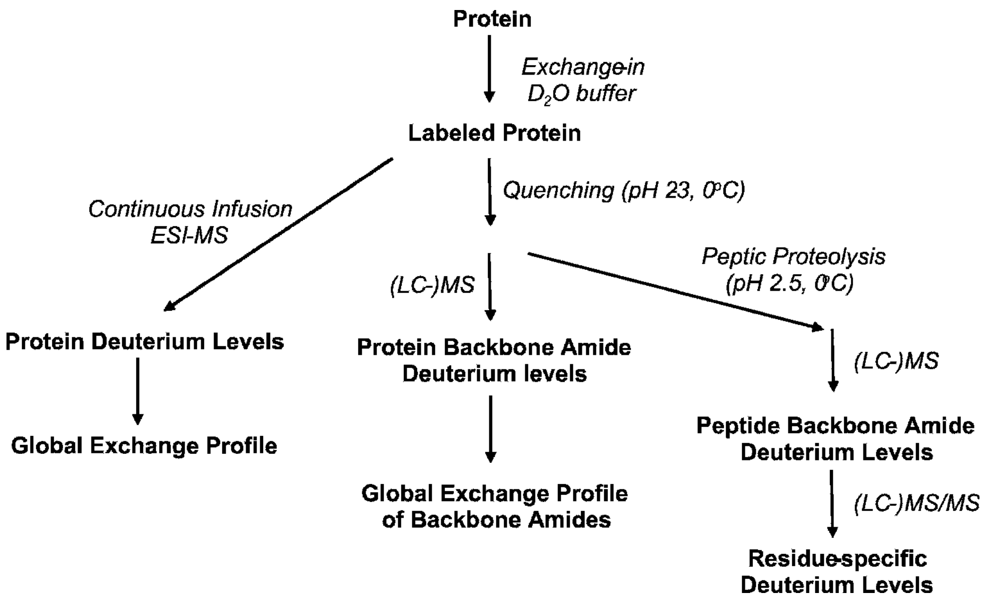

Protein
Exchange-in D_2O buffer
Labeled Protein
Continuous Infusion ESI-MS
Quenching (pH 2.3, 0°C)
(LC-)MS
Peptic Proteolysis (pH 2.5, 0°C)
Protein Deuterium Levels
Global Exchange Profile
Protein Backbone Amide Deuterium levels
Global Exchange Profile of Backbone Amides
(LC-)MS
Peptide Backbone Amide Deuterium Levels
(LC-)MS/MS
Residue-specific Deuterium Levels

Fig. 1. General procedures used to study H/D exchange by mass spectrometry.

gas phase fragmentation techniques to determine the deuterium content at individual amino acid residues and, thus, enable the extraction of residue-specific exchange information is still a relatively immature field under intensive investigation *(13–16)*.

2. Materials

2.1. Solvents and Chemicals

1. D_2O (Aldrich).
2. CHAPS (3-[(3-cholamidopropyl)dimethyl-ammonio]-1-propanesulfonate) (Pierce).
3. 9-*cis*-retinoic acid (Aldrich).
4. ACN (acetonitrile, HPLC-grade, Fisher Scientific).
5. TFA (trifluoroacetic acid, Aldrich).
6. TCEP (tris(2-carboxyethyl)phosphine hydrochloride, Pierce).
7. Urea-d_4 (Aldrich), DCl (Aldrich).
8. Guanidinium chloride-d_6 (GuHCl-d_6, Aldrich).
9. H_2O buffer: H_2O solution of 0.5 M KCl, 0.5 mM CHAPS, 1 mM TCEP, potassium phosphate buffer (50 mM, pH 7.4).
10. D_2O buffer: D_2O solution of 0.5 M KCl, 0.5 mM CHAPS, 1 mM TCEP, potassium phosphate buffer (50 mM, pH 7.4)

2.2. Proteins

1. *Growth factor receptor-bound protein 2 (Grb2).* Grb2 SH2 domain was expressed in *E. coli* strain DH5α as soluble N-terminal glutathione-S-transferase fusion protein with a cleavage sequence for thrombin *(17)*. The full length Grb2 protein was expressed and purified as described by Guilloteau et al. *(18)*.
2. *E. coli thioredoxin and S-alkyl derivatives. E.coli* thioredoxin (oxidized) was from Promega (Madison, WI). The preparations of the thioredoxin derivatives are described in Kim et al. *(19, 20)*.
3. *The retinoid X receptor (RXR) ligand binding domain.* The protein encompassing the ligand-binding domain (LBD) of RXRα was expressed in *E. coli* as His-tagged construct containing a thrombin cleavage site and purified as described by Egea et al. *(21)*.
4. *Preparation of fully deuterated proteins.* Fully deuterated proteins can be prepared according to three general protocols that use (a) heat denaturation, (b) acid denaturation, and (c) chaotropic agents to promote unfolding of the protein. It depends on the protein which method will work best. Typical conditions to prepare fully deuterated proteins involve the denaturation of the protein with 8 M urea-d_4 or 4–6 M guanidinium chloride-d_6 (all solution prepared in D_2O). Some proteins will unfold readily in 0.1N DCl. In order to ensure high level of deuteration, the respective procedure may have to be repeated several times (*see* **Note 2**). Refolding of the protein is initiated by dilution or removing of chaotropic denaturants by ultrafiltration or size exclusion chromatography.

The protocol for the preparation of fully deuterated RXRα LBD protein is detailed below:

1. The RXRα protein (45 μM) is prepared in H_2O buffer.
2. Prepare a D_2O buffer and D_2O buffer containing 8 M urea-d_4; lyophilize and re-dissolve in D_2O three times to ensure a high level of deuteration.
3. Initiate denaturation by adding an aliquot of the RXR protein stock solution (in H_2O buffer) into D_2O buffer containing 8 M urea-d_4; carry out a 1:15 (v/v) dilution.
4. Allow for deuterium exchange for 5 h at room temperature.
5. Use a Biomax membrane ultrafiltration unit (Millipore Corp. Bedford, MA) for buffer exchange from urea-d_4 containing D_2O buffer to D_2O buffer;
6. The renatured and fully deuterated RXR homodimer is then retrieved in D_2O buffer.

2.3. Mass Spectrometry

For the Grb2 studies, mass spectral data were acquired on an Electrospray Ionization time-of-flight instrument (LC-T Classic, Micromass/Waters, Manchester, UK). The thioredoxin studies were performed on a Sciex triple quadrupole instrument or a ThermoFinnigan LCQ Classic-type quadrupole ion trap instrument. The mass spectral data for the RXRα LBD studies were obtained on a LCQ Classic instrument.

3. Methods

3.1. Determining Global H/D Exchange Profiles

3.1.1. Global Exchange Profiles by Direct Infusion

Time course measurements of H/D exchange-in experiments can be directly monitored by ESI-MS using direct infusion. Exchange is initiated by dilution of a protonated protein solution into a deuterated exchange buffer. Deuterium in-take is monitored as time-dependent increase in molecular mass. Using this method, multiply deuterated $[M+nD]^{n+}$ ions are observed. This approach is advantageous for proteins that might be difficult to be prepared in their fully deuterated form. However, because of the incompatibility of the electrospray ionization process with high salt concentrations, the solvent conditions adequate for direct infusion experiments are limited to buffers employing low salt concentrations and volatile salts. Deuterated solutions containing low concentration (10–100 mM) ammonium acetate are commonly used for direct infusion ESI-MS.

For example, to study the changes in structural dynamics of the Grb2 protein upon ligand binding, H/D exchange-in profiles of the SH2 domain and full length Grb2 protein were acquired in 65 mM ammonium acetate D_2O solutions (**Fig. 2**). These studies are outlined in detail in de Mol et al. *(17)*. As SH2 domain ligand, a peptide derived from the Shc adapter protein (position Y427 in human Shc) Ac-PSpYVNVQN-NH2 (1,041 Da) was used. The exchange profiles for determining the changes in conformational mobility upon ligand binding for the Grb2 SH2 domain as well as the full length protein in absence and presence of the peptide ligand is depicted in **Fig. 2** (*see* **Notes 3, 4, and 8**). The conformational unfolding dynamics of thioredoxin *(22)* and two S-alkylated derivatives *(20)* have been studied by direct infusion H/D exchange-in mass spectrometry (**Fig. 3** and **Note 6**).

Alternatively, exchange-out experiments can be performed in which the fully deuterated protein is the starting material and exchange is initiated by diluting the fully deuterated protein in protonated solvents. Using exchange-out approaches, protonated species $[M+nH]^{n+}$ are observed by mass spectrometry and the exchange profiles monitor the exchange behavior of the

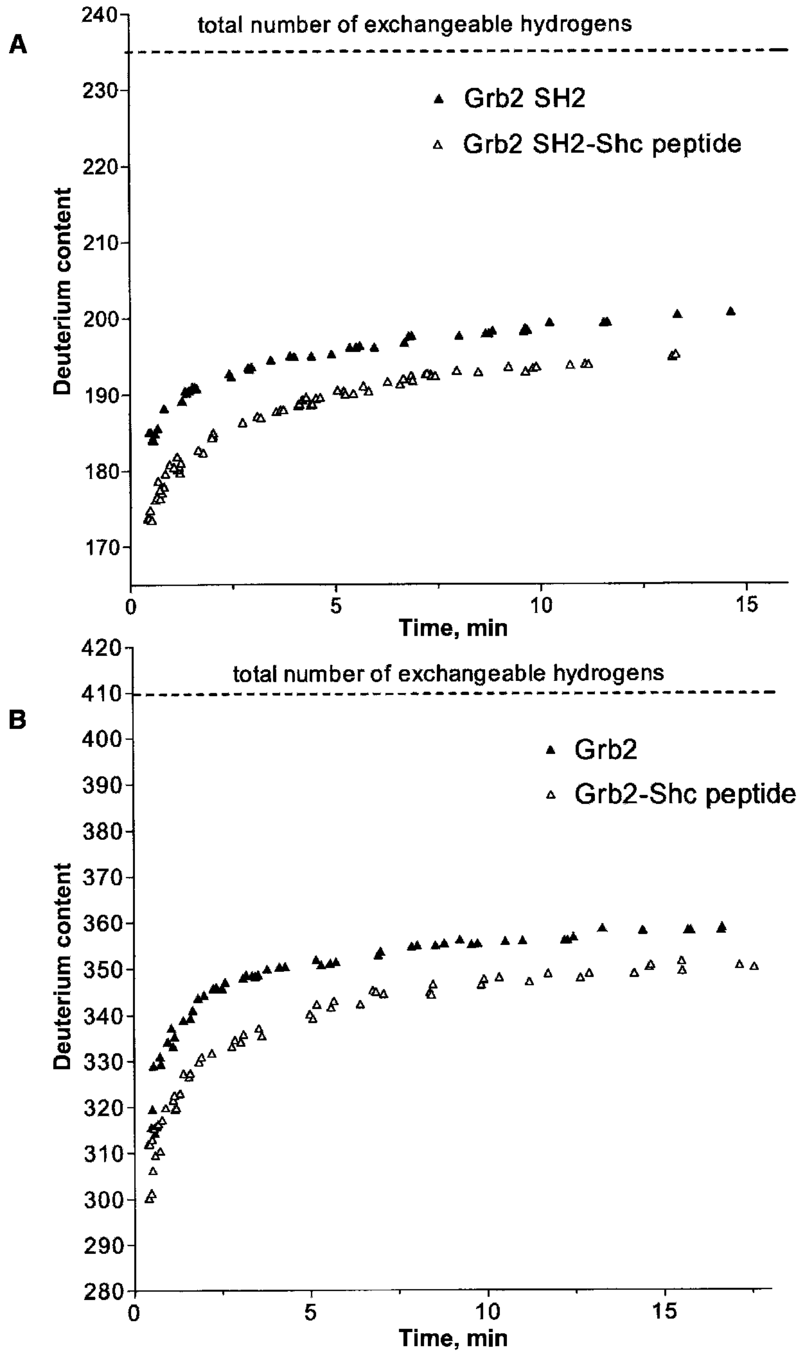


Fig. 2. Global exchange profiles of the Grb2 SH2 domain and the full length Grb2 protein in presence (*open triangles)* and absence (*filled triangles*) of its phosphopeptide ligand. In order to obtain the deuterium incorporation into the protein–ligand complex, the protein–peptide complex was dissociated in the gas phase by nozzle-skimmer dissociation. Time course analysis of deuterium incorporation reveals conformational stabilization upon ligand binding for the Grb2 SH2 domain as well as for the full-length protein. (Adapted from ref. *17).*

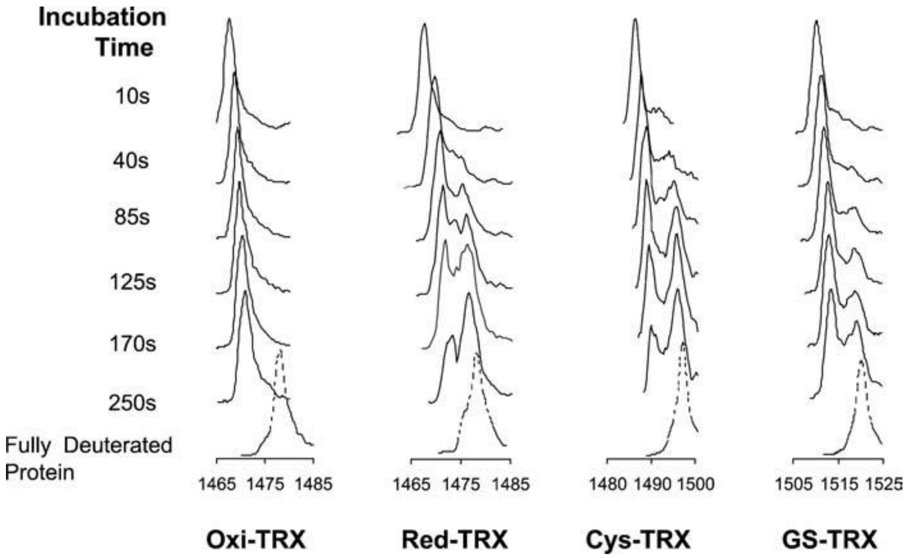


Fig. 3. H/D exchange experiment under conditions that favor EX1-type exchange studied by direct infusion ESI-MS. Evolution of the eightfold charged ion peaks of *E. coli* thioredoxin and S-alkylated derivatives is depicted. H/D exchange-in experiments are performed in 1% AcOD/D_2O at 50°C. Time points refer to H/D exchange periods. *Dashed* ion peaks represent the fully deuterated controls. Adapted from ref. *20*.

backbone amide deuterons because exchange of side-chain deuterons is several orders of magnitude faster than exchange from backbone amides. The approach is particularly powerful for study protein folding pathways *(12, 23)* and protein interfaces *(4-5)*.

3.1.2. Global Exchange Profiles Monitored by LC-ESI-MS

The best way to initiate hydrogen exchange experiments is to dilute a concentrated protein solution into deuterated buffer solutions (*see* **Notes 1**, **4**, and **5**). Deuterium incorporation at different time points, t, is then determined after quenching by LC-ESI-MS (**Fig. 4**). A trap column (optional) in combination with a short reverse phase column works in most cases and allows online desalting and concentration of protein samples (*see* **Note 7**). Side chain labels back-exchange readily under the LC conditions commonly employed; information regarding the extent of backbone amide labeling can be deduced from measured protonated species, $[M+nH]^{n+}$. To minimize artifactual back exchange of backbone amides, the solvent delivery tubing, injection valve and trap/column are immersed in an ice bath. Deuterium incorporation will be obtained at specified time point, t, $D(t) = m_t - m_{0\%}$, where m_t is the corrected mass at time t and $m_{0\%}$ is the mass of the nondeuterated peptide. If desired, correction for artifactual back exchange which occurs during the LC run can be achieved according to

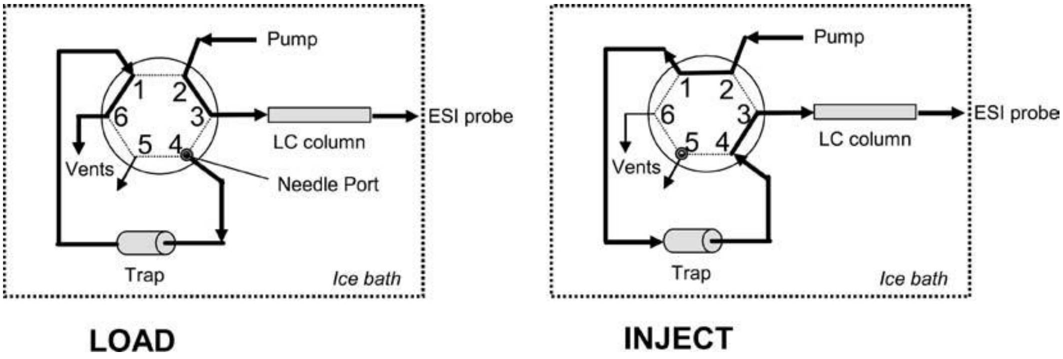


Fig. 4. LC setup used for measuring H/D exchange of backbone amides by LC-ESI MS.

$$\%D(t) = 100 \times (m_t - m_{0\%}) / (m_{100\%} - m_{0\%}),$$

where $m_{100\%}$ is the mass of the fully deuterated protein *(6)*.

The experimental procedures for determining the exchange profiles for RXRα homodimer protein after ligand (i.e., 9-*cis*-retinoic acid) binding were as follows *(24)*:

1. The purified RXRα homodimer (45 μM homodimer, 35 μL) was incubated at 25°C in H_2O buffer with and without 9-*cis*-retinoic acid (10 mM, 1 μL), for 30 min.
2. Start timer! Deuterium exchange was initiated by tenfold dilution of the protein in D_2O buffer (pD 7.0) at 25°C with and without 9-*cis*-retinoic acid ligand, i.e., 20 μL RXRα in H_2O buffer was diluted into 200-μL D_2O buffer.
3. Hydrogen/deuterium exchange was allowed to proceed for various periods.
4. At specified time, a 20-μL aliquot was taken and added to a prechilled tube containing 20 μL 0.42% phosphoric acid to quench the reaction (to reach final pH 2.5). Stop timer!
5. The quenched sample was immediately frozen in liquid nitrogen for later LC-ESI-MS analysis.
6. The number of deuterium incorporated in the protein equals the mass difference between the measured mass at time *t* and molecular mass of the protonated species. Figure 5 summarizes the data from two independently performed exchange-in time course experiments. The results are plotted as the number of deuterons found in the protein vs. labeling time.

3.2. Determining Local Hydrogen/Deuterium (H/D) Exchange Profiles

3.2.1. Establishing Peptic Peptide Maps

The protein fragmentation or"bottom-up" approach for probing local protein structural changes is essentially a two-step protocol. First, the protein is digested with pepsin under conditions that will be used for the deuterium-labeled protein samples; i.e., "quench conditions" (0°C, pH 2.5) and relative short digestion periods in combination with high enzyme-to-substrate (E:S)

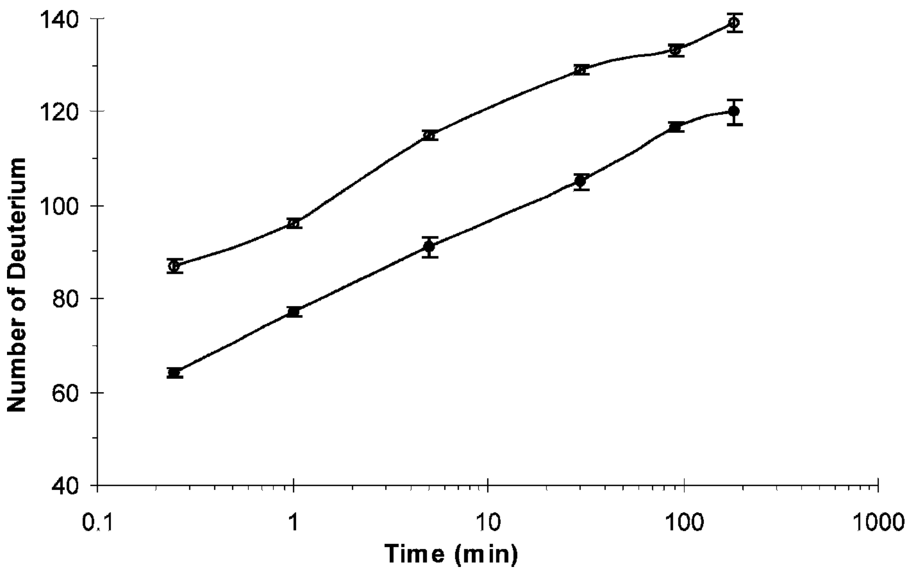


Fig. 5. Global exchange profile for RXR obtained by LC-ESI MS. **a** Number of deuterium incorporated plotted against a linear timescale; **b** Same data but plotted against logarithmic scale. Deuterium levels at amide linkages in the free human RXRα LBD homodimer (*open circle*) and the protein–ligand complex with 9-*cis*-retinoic acid (*solid circle*) shown as a function of deuterium exchange-in time (15 sec–180 min). The error bar is the standard deviation for two experiments. Adapted from ref. *24* .

ratios using (immobilized) pepsin; the peptic peptides are then identified by MS/MS. This experiment is required because pepsin cleavage sites cannot be accurately predicted, although the reproducibility of the cleavage pattern for a particular protein seems to be fairly high. In addition, it is necessary to optimize the digestion conditions for a particular protein sample to ensure coverage of the entire protein backbone.

A variation of this digestion procedure is the use of immobilized pepsin on agarose beads instead of pepsin in solution. The advantage of using immobilized pepsin is that the pepsin beads can be conveniently removed by centrifugation. Therefore, the pepsin amount can be increased to achieve digestion that is more complete and higher sequence coverage.

The inclusion of a denaturant, such as guanidinium chloride, commonly increases the digestion efficiency and, thus, significantly improves the overall sequence coverage. Although denaturants and higher salt concentration are incompatible with the ESI process, these reagents can be used with the online system (**Fig. 4**), because a C_{18}-trap assembled into the loop of the switching valve allows for removal of denaturants prior to MS analysis.

Likely, some of the proteins and protein assemblages will contain disulfide bridges, which could potentially hamper efficient proteolysis or confound the exchange information embedded in the peptic peptides. Disulfide exchange-mediated reduction of disulfides is not compatible with the quenching conditions necessary to maintain the labeling information. Protocols are available that use phosphine-based reagents, e.g., tris(2-carboxyethyl phosphine) hydrochloride (TCEP), to reduce disulfides under acidic conditions that preserve the deuterium label information *(25)*.

A protocol for the use of immobilized pepsin is listed below. The amounts given are sufficient for three digestions. To verify the reproducibility of the peptic digestion at least duplicate analysis is recommended.

1. Keep vials, buffer, and immobilized pepsin bead suspension always chilled using an ice bath.
2. Aliquot 20-μL protein solution (~50 μM) into a prechilled vial; add 20 μL prechilled phosphate buffer (0.2 M, pH 2.5).
3. Buffer exchange 150 μL of immobilized pepsin into phosphate buffer (0.2 M, pH 2.5) using three times 500 μL; each time quick-spin on a picofuge; finally resuspend in 150 μL.
4. Start timer for next step.
5. Add a 45 μl-immobilized pepsin aliquot to each of the prechilled vials containing the protein solution; pH should be 2.5!
6. Stop reaction after 3 min by quick-spin and transfer supernatant into a prechilled vial; immediately analyze by LC-MS/MS or store by flash-freezing in liquid nitrogen.

3.2.2. Exchange Experiments and Offline Peptic Digestion in Combination with LC-ESI-MS

Proteins are typically equilibrated under physiological conditions (e.g., 10 mM phosphate buffer). The equilibrated protein solutions are then diluted 20-fold into 10 mM D_2O-phosphate buffer and incubated for specified periods. Exchange is quenched by acidifying with ice-cold 0.1 M HCl (pH 2.5). The labeled protein is subsequently digested with pepsin at 0°C. All operations should be carried out as quickly as possible. Vials and syringes have to be prechilled. To minimize back exchange of deuterium labels during HPLC analysis, the solvent delivery tube, injection valve, and the analytical columns are immersed in an ice bath.

If necessary, samples can be stored by flash-freezing in liquid nitrogen after peptic digestion for later analysis.

Magtran software *(26)* can be used to generate the centroid MS data for deuterated peptides. Deuterium incorporation for peptides is determined from the difference of the deuterated peptide mass and the undeuterated peptide mass (**Figs. 6** and 7).

The following protocol was used for the peptic proteolysis of RXRα *(19)*:

1. Deuterium exchange reaction of a protein (50 μM) was carried out at 25°C in a phosphate buffer (10 mM, pD 6.5) or other buffers that are close to neutral pH.

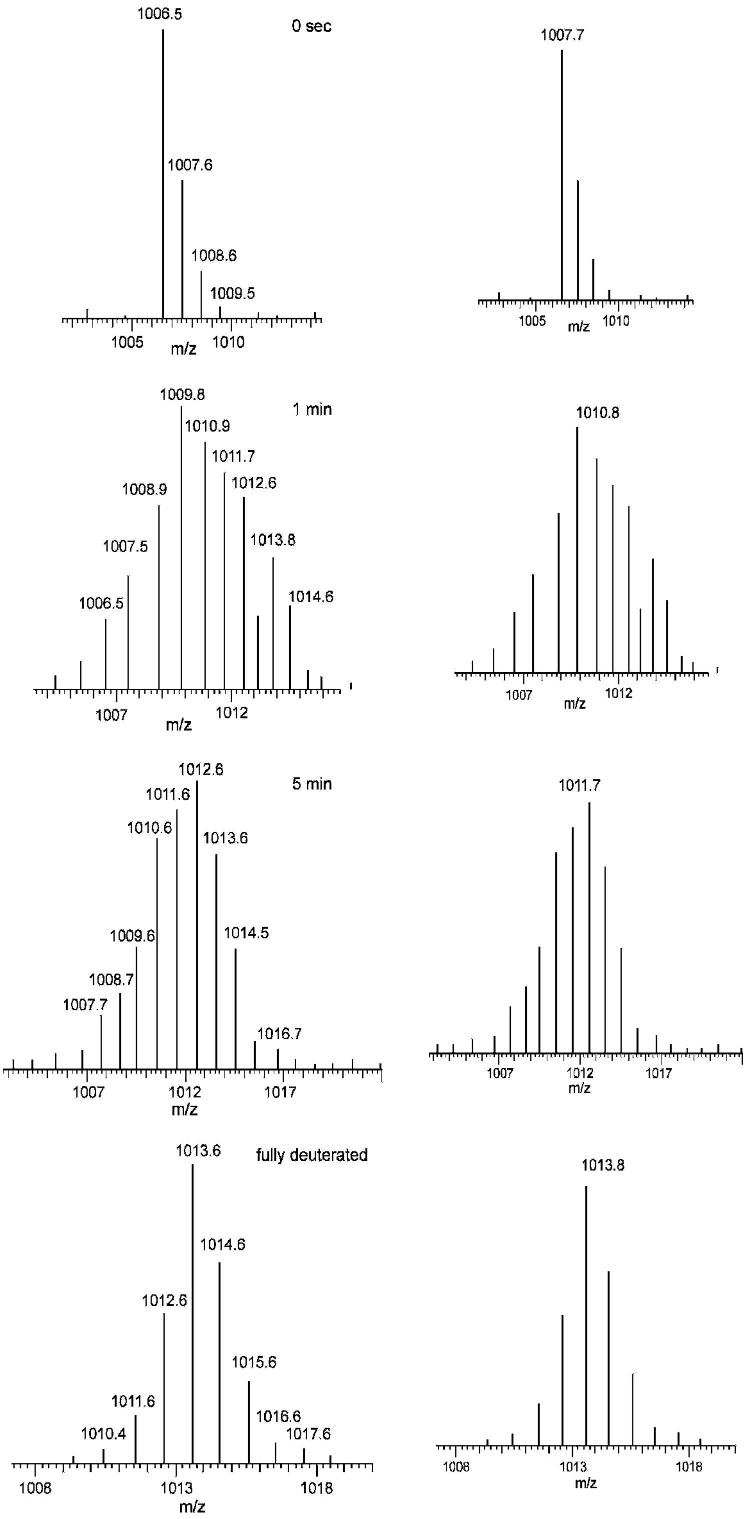


Fig. 6. H/D exchange mass spectra for the singly charged ion of the peptic peptide aa 271–279 from RXRα LBD when bound to its ligand 9-*cis*-retinoic acid. The partial sequence of this peptide is AADKQLFTL (eight backbone amides). *Left side*: Mass spectra of the peptide obtained after 0-, 1-, 5-min deuterium labeling and of the fully deuterated peptide. *Right side*: Magtran software was used to obtain the centroid masses.

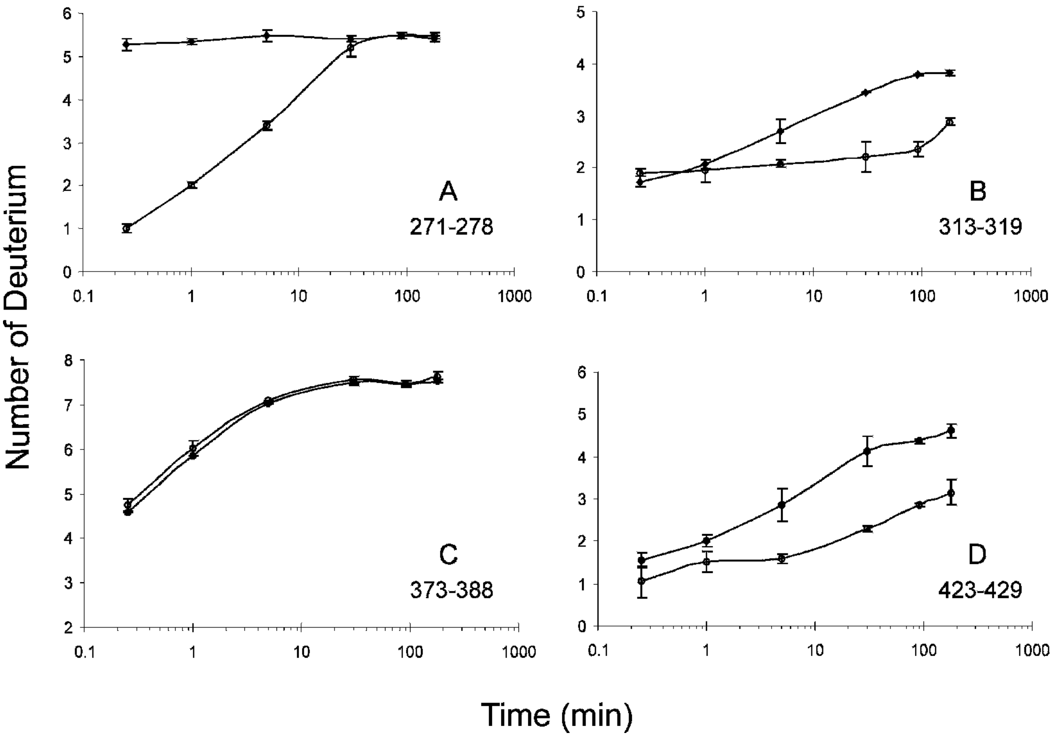


Fig. 7. Time course of deuterium incorporation for peptic peptides measured by LC-ESI MS. Deuterium levels found at amide linkages as a function of time in four representative peptic fragments of human RXRα LBD, aa 271–278 (**a**), 313–319 (**b**), 373–388 (**c**) and 423–429 (**d**). The *solid circles* represent the peptide deuterium content in the free RXRα LBD and the *open circles* indicate deuterium content in the protein-9-*cis*-retinoic acid complex. The error bar is the standard deviation for two experiments. Adapted from ref. *24*.

2. After a certain period, an aliquot was taken (20 μL) and immediately quenched by acidification with 0.1 M HCl to pH 2.5 in a prechilled reaction vial.
3. Have a freshly prepared and chilled pepsin solution (3 mg/mL) in 5% formic acid/H_2O ready for immediate use.
4. Digest protein sample by adding 5 μL pepsin for 3 min at 0°C (E:S ~1:2 molar ratio), while shaking (using a vortex shaker).
5. Flash-freeze in liquid nitrogen for subsequent analysis or immediately analyze aliquot by LC-ESI-MS.

Several groups have reported the implementation of pepsin columns for online proteolysis. The efficacy of different pepsin columns for online digestion and subsequent LC analysis is described in *(27)* and *(28)*.

3.3. H/D Exchange Data Analysis

3.3.1. Determining the Exchange Regime

H/D exchange data will be evaluated to determine which exchange behavior is observed. A dual pathway model for hydrogen exchange of peptide amide hydrogens is most often considered. One pathway describes exchange directly from the folded state via local transient opening reactions or "breathing" motions. It is assumed that peptide amide hydrogens located at or near the surface of the protein (or in close proximity to solvent channels) exchange via this pathway. In the other pathway, exchange is preceded by transient reversible unfolding reactions that are accompanied by the breaking of hydrogen bonds. Unfolding of proteins can occur via (a) global or whole-molecule transitions, (b) subglobal transitions, or (c) local fluctuations. The dominant mechanism of this dual pathway depends on the experimental conditions, i.e., the pH, temperature, or chemical denaturants present. The Linderstrøm-Lang model connects the second mechanism with the observed exchange rates *(7, 29)*

$$F_H \qquad \overset{D_2O}{\underset{k_f}{}} \qquad F_H$$

$$F_H \qquad \overset{k_1}{\underset{k_{\square 1}}{}} \qquad U_H \; \square \; \overset{D_2O}{\underset{k_{rc}}{\square}} \; \square\square \; UD \qquad \overset{k_{\square 1}}{\underset{k_1}{}} \qquad F_D$$

The observed exchange reaction will be second order if the refolding rate of the transient opening reaction is much greater than the rate with which the intrinsic chemical exchange occurs, i.e., $k_{-1} >> k_{rc}$. Thus, the experimentally observed rate constant is given by $k_{obs} = k_{rc}(k_1 / k_{-1})$. This type of exchange is termed as EX2-type mechanism and exchange via this mechanism is typically observed under neutral pH and in the absence of denaturants. In this mechanism, refolding (i.e., large k_{-1}) is more likely than the exchange reaction and the exchange is random or uncorrelated. For EX2 behavior, the mass spectra of the respective peptide in a time course study will show a single weighted average peak whose mass increases over time. The rate constant is then calculated from the increase in $\%D$ as a function of time.

If the rate of the chemical reaction is much greater than refolding, i.e., $k_{rc} >> k_{\square 1}$, then exchange occurs via so-called EX1-type mechanism. The observed exchange rate is directly related to the protein unfolding rate, i.e., the observed exchange rate constant is given by $k_{obs} = k_1$ and correlated motion can be detected by mass spectrometry. For the EX1 limit, two population of peptides will be observed, completely exchanged (high mass speak) and not exchanged (low mass peak), and as time increases, the area under the high mass peak will increase and the area under the low mass peak will decrease (**Fig. 3**); therefore, the unfolding rate, k_1, can be determined from the rate of increase of area of the higher mass peak.

3.3.2. Determining Exchange Rate Constants

Deuterium contents of labeled peptic peptides are obtained by comparing the masses of the undeuterated peptides with the masses of the deuterated peptides. Deuterium incorporation is plotted as the function of the incubation time; the H/D exchange rate constant can be calculated using the relationship:

$$\%D = \sum_{i=1}^{n} A_i(1 - e^{-k_i t}),$$

where A_i is the amplitude and k_i is the rate constant for deuterium exchange in the *i*th kinetic phase. The distribution of rate constants for isotope exchange at peptide amide bonds can be obtained by fitting the exchange data using, e.g., Scientist (Micromath), Sigma Plot (Jandell), or Origin (Microcal Software, Inc.). For instance, in the conformational studies of the recombinant protein M-CSFβ *(30)*, the exchange data obtained for the peptic peptides were fitted to a series of four exponential functions that allowed the grouping of the exchange rate constants into four categories: very fast ($k_1 > 10$ min^{-1}), fast ($1 > k_2 < 0.1$ min^{-1}), slow ($0.1 > k_3 < 0.01$ min^{-1}), and slowest ($k_4 < 0.001$ min^{-1}).

4. Notes

1. Lack of hydrogen exchange protection for a given protein system may indicate that the protein system is not properly folded. One of the many possible reasons is that the solvent conditions chosen for studying H/D exchange are inappropriate, e.g., wrong pH range, incorrect salt concentrations, absence of a stabilizing metal ion or cofactor that may promote folding, etc. It is advisable to check correct folding and functional characteristics by independent spectroscopic techniques and functional assays prior to H/D exchange experiment.
2. In order to judge the experimental procedure with a so-called fully deuterated control, one has to remember that the isotope content of the deuteration solution has to be considered. For example, for the Grb2 protein, the total number of exchangeable hydrogens is 410 (by adding up all the side chain and backbone exchangeable hydrogens). If the deuteration buffer contains 5% residual H_2O, ~20–21 sites will be occupied by protons (1% of the total number of exchangeable sites).
3. For direct infusion measurements using ESI-MS, the fused silica tubing that connects the syringe with the ESI source and the ionization chamber should be equilibrated with the deuterated solvent at least 15–20 min prior to the measurements of $[M+nD]^{n+}$ species.

4. To study protein–ligand or protein–protein complexes, it is important to equilibrate the constituents of the complex prior to H/D experiment. Commonly, exchange is initiated by dilution; it is therefore worth checking if the noncovalent complex is maintained under the diluted conditions (check K_D) and adjust the concentration of the constituents during dilution appropriately.

5. While the sensitivity of ESI-MS considerably improved over the last decade, approximately 100 μL of a 0.5–5 μmol protein solution is still advantageous for a time-course study. It is advisable to do at least two measurements (duplicates) for each time point.

6. To study exchange in the EX1-type regime, thermal unfolding conditions were applied. To study the dynamic processes by direct infusion ESI-MS, the solvent delivery line, the reaction vessel, and the injection valve were immersed in a water bath equilibrated at the desired temperature. Thermal denaturation and H/D exchange-in occur in the reaction vessel. In order to quench H/D exchange-in and initiate refolding, the capillary that connected the reaction vessel with the ESI source was immersed in an ice bath. This setup also insures that the same charge states $[M+nH]^{n+}$ can be used for determining deuterium incorporation levels. The experimental conditions are described in detail in *(20, 22)*.

7. A typical instrumental LC setup for H/D exchange experiments consists of a C_{18}-trap (0.5 mm ID × 2 mm; ~30 μL/min flow rate) followed by a C_{18}-reverse phase column (0.1 mm ID × 50 mm; ~0.8–1 μL/min flow rate) as analytical separation column. Miniaturizing of the LC setup improves the sensitivity of the approach as described by Wang and Smith *(27)*.

8. When analyzing the exchange characteristics of noncovalent protein–peptide complexes by direct infusion ESI-MS, resulting mass spectra typically show numerous multiply charged ions. To analyze the deuterium incorporation into the protein, disruption of the protein–peptide complex is often necessary. Nozzle-skimmer dissociation of the protein–ligand complex is typically achieved by applying high cone voltages (e.g., between 75 and 150 V on the LC-T type instrument). To ensure accurate mass calibration of the instrument under given conditions the free protein samples should be measured under the same conditions that were used for the analysis (and dissociation) of the protein–peptide complex (e.g., elevated cone voltages of 75–150 V).

References

1. Katta, V., Chait, B. T. Conformational changes in proteins probed by hydrogen-exchange electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* 5, 214–217 (1991).
2. Englander, S. W., Kallenbach, N. Hydrogen exchange and structural dyamics of proteins and nucleic acids. *Q. Rev. Biophys.* 16, 521–655 (1984).
3. Englander, S. W., Sosnick, T. R., Englander, J. J., Mayne, L. Mechanisms and uses of hydrogen exchange. *Curr. Opin. Struct. Biol.* 6, 18–23 (1996).
4. Mandell, J. G., BaergaOrtiz, A., Akashi, S., Takio, K., Komives, E. A. Solvent accessibility of the thrombin–thrombomodulin interface. *J. Mol. Biol.* 306, 575–589 (2001).
5. Mandell, J. G., Falick, A. M. & Komives, E. A. Identification of protein–protein interfaces by decreased amide proton solvent accessibility. *Proc. Natl Acad. Sci. USA* 95, 14705–14710 (1998).
6. Zhang, Z., Smith, D. L. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Sci.* 2, 522–531 (1993).
7. Smith, D. L., Deng, Y., Zhang, Z. Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J. Mass Spectrom.* 32, 135–146 (1997).
8. Eigen, M. Proton transfer, acid-base catalysis, and enzymatic hydrolysis. *Angew. Chem.* **3**, 1–19 (1964).
9. Eriksson, M. A. L., Haerd, T., Nilsson, L. On the pH dependance of amide proton exchange rates in proteins. *Biophys. J.* 69, 329–339 (1995).
10. Bai, Y., Milne, J. S., Mayne, L., Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins* 17, 75–86 (1993).
11. Englander, J. J., Rogero, J. R., Englander, S. W. Protein hydrogen exchange studied by the fragment separation method. *Anal. Biochem.* 147, 234–244 (1985).
12. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., Dobson, C. M. Detection of transient protein folding populations by mass spectrometry. *Science* 262, 896–899 (1993).
13. Demmers, J. A., Rijkers, D. T., Haverkamp, J., Killian, J. A., Heck, A. J. Factors affecting gas-phase deuterium scrambling in peptide ions and their implications for protein structure determination. *J. Am. Chem. Soc.* 124, 11191–11198 (2002).
14. Kim, M. Y., Maier, C. S., Reed, D. J., Deinzer, M. L. Site-specific amide hydrogen/deuterium exchange in *E. coli* thioredoxins measured by electrospray ionization mass spectrometry. *J. Am. Chem. Soc.* 123, 9860–9866 (2001).
15. Kaltashov, I. A., Eyles, S. J. Crossing the phase boundary to study protein dynamics and function: Combination of amide hydrogen exchange in solution and ion fragmentation in the gas phase. *J. Mass Spectrom.* 37, 557–565 (2002).
16. Jorgensen, T. J., Gardsvoll, H., Ploug, M., Roepstorff, P. Intramolecular migration of amide hydrogens in protonated peptides upon collisional activation. *J. Am. Chem. Soc.* 127, 2785–2793 (2005).
17. de Mol, N. J.et-al.. , Changes in structural dynamics of the Grb2 adaptor protein upon binding of phosphotyrosine ligand to its SH2 domain. *Biochim. Biophys. Acta* 1700, 53–64 (2004).
18. Guilloteau, J. P.et-al.. , Purification, stabilization, and crystallization of a modular protein: Grb2. *Proteins* 25, 112–119 (1996).
19. Kim, M. Y., Maier, C. S., Ho, S., Deinzer, M. L. Intramolecular interactions in chemically modified Escherichia coli thioredoxin monitored by hydrogen/deuterium exchange and electrospray ionization mass spectrometry. *Biochemistry* 40, 14413–14421 (2001).
20. Kim, M. Y., Maier, C. S., Reed, D. J., Deinzer, M. L. Conformational changes in chemically modified *Escherichia coli* thioredoxin monitored by H/D exchange and electrospray ionization mass spectrometry. *Protein Sci.* 11, 1320–1329 (2002).
21. Egea, P. F.et-al.. , Effects of ligand binding on the association properties and conformation in solution of retinoic acid receptors RXR and RAR. *J. Mol. Biol.* 307, 557–576 (2001).
22. Maier, C. S., Schimerlik, M. I., Deinzer, M. L. Thermal denaturation of Escherichia coli thioredoxin studied by hydrogen/deuterium exchange and electrospary ionization mass spectrometry: Monitoring a two-state protein unfolding transition. *Biochemistry* 38, 1136–1143 (1999).
23. Matagne, A.et-al.. , Thermal unfolding of an intermediate is associated with non-Arrhenius kinetics in the folding of hen lysozyme. *J. Mol. Biol.* 297, 193–210 (2000).
24. Yan, X., Broderick, D., Leid, M. E., Schimerlik, M. I., Deinzer, M. L. Dynamics and ligand-induced solvent accessibility changes in human

retinoid X receptor homodimer determined by hydrogen deuterium exchange and mass spectrometry. *Biochemistry* 43, 909–917 (2004).

25. Yan, X., Zhang, H., Watson, J., Schimerlik, M. I., Deinzer, M. L. Hydrogen/deuterium exchange and mass spectrometric analysis of a protein containing multiple disulfide bonds: Solution structure of recombinant macrophage colony stimulating factor-beta (rhM-CSFbeta). *Protein Sci.* 11, 2113–2124 (2002).
26. Zhang, Z., Marshall, A. G. A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *J. Am. Soc. Mass Spectrom.* 9, 225–233 (1998).
27. Wang, L., Smith, D. L. Downsizing improves sensitivity 100-fold for hydrogen exchange-mass spectrometry. *Anal. Biochem.* 314, 46–53 (2003).
28. Wu, Y., Kaveti, S., Engen, J. R. Extensive deuterium back-exchange in certain immobilized pepsin columns used for H/D exchange mass spectrometry. *Anal. Chem.* 78, 1719–1723 (2006).
29. Hvidt, A., Nielsen, S. O. Hydrogen exchange in proteins. *Adv. Protein Chem.* 21, 287–386 (1966).
30. Zhang, Y. H., Yan, X., Maier, C. S., Schimerlik, M. I., Deinzer, M. L. Structural comparison of recombinant human macrophage colony stimulating factor beta and a partially reduced derivative using hydrogen deuterium exchange and electrospray ionization mass spectrometry. *Protein Sci.* 10, 2336–2345 (2001).

Chapter 16

Mass Spectrometry Detection and Characterization of Noncovalent Protein Complexes

Sheng Yin and Joseph A. Loo

Summary

Protein noncovalent complexes compose most of the essential biological machines in the cell. The characterization of protein complexes and assemblies using mass spectrometry (MS) has significant advantages over many other biophysical methods because of the inherent sensitivity and resolution of MS. The applicability of MS coupled with electrospray ionization (ESI) for the measurement of large proteins and protein complexes has been furthered by the development of sensitive analyzers, such as the time-of-flight (TOF) analyzer. Moreover, sample preparation has a very important role for such studies, as it could significantly affect the results of mass spectrometry experiments. We discuss the experimental variables for the ESI-MS detection of noncovalent protein complexes by featuring two different protein systems: yeast enolase is a 93-kDa homodimeric complex, and α-synuclein is a small, 14-kDa protein implicated in the pathogenesis of Parkinson's disease and binds noncovalently to endogenous polyamines, such as spermine.

Key words: Mass spectrometry, Electrospray ionization, Noncovalent protein complex, Enolase, α-Synuclein, Spermine.

1. Introduction

The structural determination of protein complexes can play an important role in the fundamental understanding of biochemical pathways. The application of electrospray ionization mass spectrometry (ESI-MS) for studying noncovalent complexes has utility in biomedical research. ESI-MS has been applied as a drug screening and chemical biology tool in the pharmaceutical industry because of its ability to measure protein–ligand associations *(1, 2)*. By measuring or knowing the masses of the individual binding partners, the stoichiometry of the binding partners in the

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_16

complex can be determined from the molecular mass measurement of the intact complex, even for multiligand heterocomplexes *(1, 3–5)*. The ability of ESI to ionize macromolecules without disrupting covalent bonds and maintaining the weak noncovalent interactions is the key distinguishing feature of ESI for the study of such biological complexes. Although the molecular masses of large proteins and protein complexes generally exceeds the mass-to-charge (m/z) ratio of most mass spectrometric analyzers, ESI generates multiply charged molecules (i.e., $z > 1$), such that analyzers of m/z range 10,000 or less can be used for most large protein complexes. (Complexes larger than 500,000 Da may require analyzers with m/z limits greater than 10,000 *(6)*.)

Often, the noncovalent binding of protein–protein and protein–ligand complexes is well defined in solution. For many classes of complexes, a correlation is found between the binding affinity and the surface area that is buried upon complexation. Moreover, the magnitudes of noncovalent binding in water are determined largely by hydrophobic effects. However, it is often the electrostatic interactions that govern whether noncovalent interactions can be detected by ESI-MS, as electrostatic interactions are strengthened in vacuum. As a result, we have been able to measure protein–ligand complexes with solution equilibrium constants (K_D) even in the millimolar (10^{-3} M) regime. ESI-MS has demonstrated its potential for detecting protein–protein and protein–ligand complexes. This requires that there is some fidelity between the structure and energetics of the gas phase and solution phase complexes. In our laboratory, we have used ESI-MS to measure a variety of noncovalent protein complexes, ranging in size from a few kDa to over 0.5 MDa *(6)*. We describe the critical experimental steps to prepare the protein solutions required for the successful measurement of noncovalent protein complexes. In general, extraneous salts and buffers found often in biological preparations need to be removed and replaced with an ESI-compatible buffer that is volatile, such as ammonium acetate or ammonium bicarbonate. Noncovalent protein complexes are generally observed only from solutions maintained at near-physiological pH (pH 6–8). Some complexes in biology can be observed at the extremes of the pH scale (less than 4 and greater than 9) *(7)*. However, these pH values are considered denaturing for a majority of noncovalent protein complexes.

2. Materials

2.1. Protein Sample Preparation

1. Protein solutions are desalted in 100 mM ammonium acetate (NH_4OAc) adjusted to pH 6–7 with acetic acid.

2. Final protein solutions analyzed by ESI-MS under native solution conditions generally contain 10 mM NH_4OAc or 10 mM ammonium bicarbonate adjusted to pH 6–8 with acetic acid or formic acid.
3. Protein solutions are desalted and concentrated using centrifugal membrane filtration devices. Typical filtration devices are Centricon and Microcon filtration tubes (Millipore, Billerica, MA) for smaller volume quantities, and the Amicon Ultra-4 centrifugal filter devices for larger volume applications. Follow the manufacturer's protocols for application of the filtration devices. The centrifuge instrument used in our laboratory is the Eppendorf centrifuge 5804R with an Eppendorf swing bucket rotor (A-4-44).
4. Samples discussed are yeast enolase (Sigma Chemical, St. Louis, MO), horse myoglobin (Sigma Chemical), recombinant human α-synuclein (rPeptide, Athens, GA), and spermine (Sigma Chemical).

2.2. ESI Mass Spectrometry

1. The ESI mass spectrometer used is a QqTOF mass spectrometer (Applied Biosystems/Sciex QSTAR Pulsar XL, Concord, ON, Canada) with a *m/z* range of 40,000.
2. A nanoESI source operating at low analyte flow conditions was used with the QqTOF instrument. Borosilicate glass nanospray emitters that are coated with Au/Pd to allow for 10–50 nL/min spray operation were obtained from Proxeon (Odense, Denmark).
3. A metal sleeve used to restrict pumping in the initial region of the Q0-focusing quadrupole was used to allow for improved trapping and transmission of the noncovalent complexes.

3. Methods

It is important to prepare the protein solutions properly for the successful MS measurement of the noncovalent complexes. Buffer solutions composed of either ammonium acetate or ammonium bicarbonate are preferred for ESI-MS measurements because they are volatile; as a result, observed adduct formation is minimal. Protein solutions should be desalted to remove extraneous components. Centrifugal membrane filtration devices can be used to desalt, buffer exchange, and concentrate protein solutions prior to ESI-MS measurements. The mass spectra shown were acquired using a quadrupole time-of-flight (QTOF) mass spectrometer. The ESI-QTOF and ESI-TOF systems are especially popular for the detection of noncovalent protein complexes because of their excellent sensitivity, modest mass resolution, and large *m/z* range required (*m/z* > 6,000) for measuring larger protein complexes. However, other types of mass spectrometers, such as quadrupole, ion trap, OrbiTrap, and Fourier-transform ion cyclotron resonance (FT-ICR) systems have been demonstrated to be applicable for such measurements.

Two examples of noncovalent protein complexes are demonstrated. Enolase is a metalloenzyme involved in the glycolytic pathway, catalyzing the dehydration of 2-phospho-d-glycerate. The yeast enzyme is a 93-kDa homodimeric complex (monomer 46.7 kDa) *(8)*. The pathological hallmark of the neurodegenerative disorder, Parkinson's disease (PD), is the presence of intracellular inclusions, called Lewy bodies, in the dopaminergic neurons of the substantia nigra. Filamentous α-synuclein (AS, 14,460 Da) protein is the major component of these deposits and its aggregation is believed to play an important role in PD. Binding of AS to natural polycations, such as spermine (202.2 Da), enhances aggregation of the AS protein.

3.1. Sample Preparation for ESI-MS of Noncovalent Protein Complexes

1. Prepare a stock solution of the protein of interest with a concentration approximately 1–5 mg/mL in 10 mM ammonium acetate (NH_4OAc) buffer (pH 6.5). The protein concentration will depend on the size and solubility of the protein. A concentration of 100 μM is a good target for further dilution experiments.
2. Centrifugal membrane filtration devices (e.g., Centricon and Microcon) are used to desalt and concentrate the protein solutions. Use the proper molecular weight cutoff filter membranes such that the expected molecular weight of the protein complex is greater than the MWCO size. This will allow smaller molecular weight salts to be filtered away from the solution. (*see* **Notes 1–3**)
3. A pre-rinse is necessary to prepare the filtration devices prior to addition of the protein. Add 2 mL of 100 mM NH_4OAc into the centrifugal tubes and centrifuge at 2,000 × *g*–4,000 × *g* at 4°C for 5 min.
4. After rinsing, mix the protein sample with a solution of 100 mM NH_4OAc buffer to a volume of 4 mL into each filtration tube. Centrifuge at least three times at 4,000 × *g* at 4°C for 10 min; the filtrate should be discarded and replaced with an equal volume of 100 mM NH_4OAc. Keep in mind that the final volume in Amicon centrifugal tubes will be no less than 50 μL, and the final concentration of the protein complex for MS experiments should be at least 5 μM.
5. After desalting with the 100 mM NH_4OAc buffer three times, rinse the sample one additional time with 10 mM NH_4OAc buffer to replace the higher concentration buffer.

6. The final volume according to the scale on the centrifugal tube should be approximately 50 µL. If the volume is significantly greater, continue centrifugation at 4,000 × *g*, 4°C.
7. Use a pipette to carefully withdraw the cleaned sample solution from the bottom of the centrifuge tube into a storage tube, and record the final volume to estimate the protein concentration.(*see* **Note 4**)

3.2. Measuring Noncovalent Protein–Protein Complexes by ESI-MS

1. The molecular weight of the protein monomer should be measured by ESI-MS prior to measurement of the noncovalent protein complex. Using an aliquot of the cleaned, desalted protein, add sufficient volumes of acetonitrile and formic acid to fully denature the protein. Typically, a solution of 50% acetonitrile (*v/v*) and 0.1% formic acid (pH less than 3) is sufficient to completely denature the protein.
2. The ESI mass spectrum of yeast enolase under denaturing solution conditions (**Fig. 1**) shows multiple charging to greater than 70+, with protein ions observed between *m/z* 600 and 1,600. The molecular weight can be calculated using the mass spectrometer's software for mass spectral deconvolution. For enolase, the monomer molecular weight is approximately 46.7 kDa.
3. Obtaining an ESI mass spectrum of denatured proteins can be performed using either conventional low solution flow rate ESI sources (1–10 µL/min) by direct infusion or

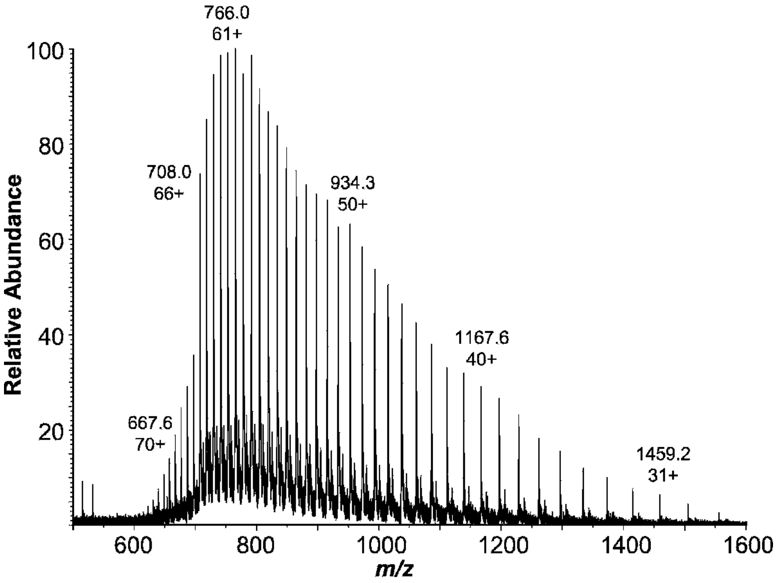


Fig. 1. ESI QqTOF mass spectrum of yeast enolase (5 µm) dissolved in 50% acetonitrile, 0.1% formic acid. Under denaturing solution conditions, multiply charged molecules of the enolase monomer are detected.

nanoelectrospray sources with borosilicate nanospray needles (50–200 nL/min).

4. To measure the molecular weight of the noncovalent protein complex, a protein solution in 10 mM NH_4OAc or ammonium bicarbonate (NH_4HCO_3), pH 6–8, should be used (i.e., no addition of organic modifiers or acids).
5. Best sensitivity for detecting the noncovalent protein complex is obtained using nanoelectrospray sources with borosilicate nanospray needles (50–200 nL/min) because of the smaller droplets formed.
6. Often, multiply charged ions for the noncovalent protein complex are found at much higher *m/z* than that observed for the denatured protein. **Figure 2** shows the ESI mass spectrum of yeast enolase in ammonium acetate. The protein dimer (93.4 kDa) charges to 21+ and its ions are found between *m/z* 4,000 and 5,000.
7. The lens potentials in the atmospheric pressure/vacuum interface will need to be adjusted properly to optimize sensitivity for the protein complexes. Each mass spectrometer and ESI interface from each commercial vendor optimizes differently. However, generally, the potentials on the nozzle and skimmer elements found in the ESI interface (or their equivalents) need to be increased to properly effect ion desolvation. As the potentials are increased, the peak widths for the protein complex ions should narrow because solvent and salt mol-

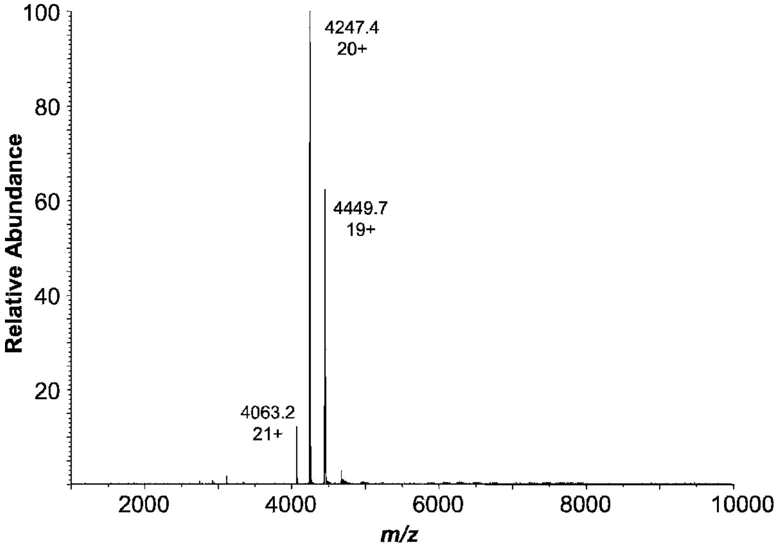


Fig. 2. ESI QqTOF mass spectrum of yeast enolase (5 μm) in 10-mm NH_4OAc (pH 6.5) showing multiply charged ions for the intact 93.4 kDa dimer.

ecules that adduct onto the ions are removed. There will be a potential, however, that desolvation is maximized but protein complex dissociation starts to occur. (*see* **Note 5**)

3.3. Measuring the Solution Binding Constant for a Protein–Ligand Complex by ESI-MS

1. Prepare the target protein solution as discussed in **Subheading 3.1**. A buffer of 10 mM NH_4OAc (pH 6.5) is the typical solution that is used for such ESI-MS experiments. For protein–small molecule ligand studies, the protein is normally desalted/cleaned separately, and the small molecule ligand is added or titrated separately into the protein solution.
2. Prepare a solution containing the stock desalted protein and excess ligand in 10 mM NH_4OAc (pH 6.5) and measure an ESI mass spectrum of the noncovalent protein–ligand complex using nanoelectrospray sources with borosilicate nanospray needles, as discussed in **Subheading 3.2**. Optimize the lens potentials in the atmospheric pressure/vacuum interface to maximize sensitivity for the protein–ligand complex, while also minimizing protein–ligand gas phase dissociation. The optimized lens settings should be used for all subsequent titration experiments for the given protein–ligand system.
3. Start with a protein solution concentration of approximately 5 μM in 10 mM NH_4OAc and add sufficient ligand to a concentration less than the protein concentration. Record the ESI mass spectrum over the full *m/z* range that covers both the mass of the ligand and the ions for the protein and protein–ligand complex.
4. The concentration of the stock solution for the ligand should be sufficiently concentrated to minimize the volume added to the protein solution for the titration experiment. For example, for a 5 μM protein solution, the stock ligand solution could be 100 μM.
5. Keep the protein concentration constant and titrate increasing amounts of the ligand to the solution. Record ESI mass spectra for each ligand concentration studied. The intensity of the ions for the protein–ligand complex should increase with increasing ligand concentration. **Figure 3** shows the ESI mass spectrum of AS protein complexed with spermine to form a 1:1 noncovalent complex.
6. To construct a protein–ligand binding titration plot, the following principles are assumed (the method for measuring a binding constant is similar to that reported by Zhang et al. *(9)*):
    (a) The sum of the intensities (I) of the multiply charged protein ions or protein–ligand complex ions reflects the

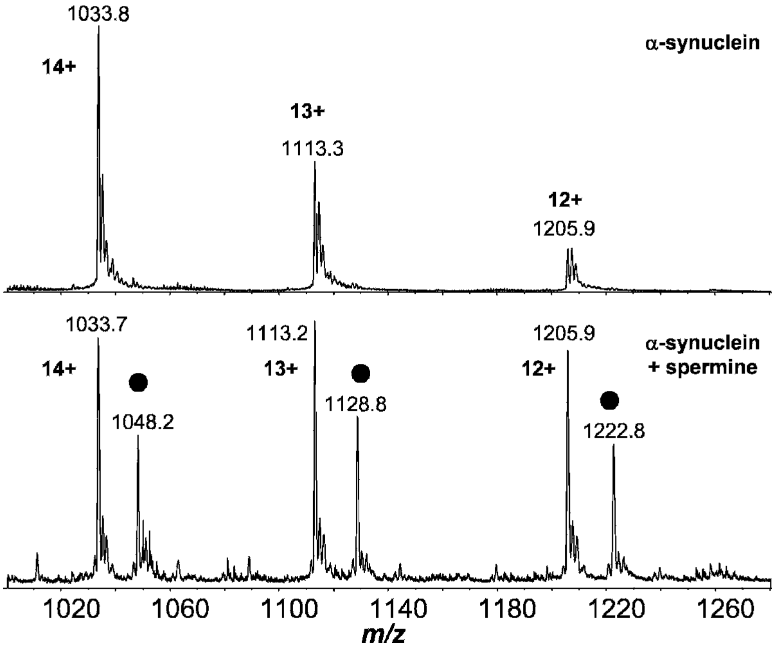


Fig. 3. ESI QqTOF mass spectra of α-synuclein (5 μm) in 10-mm NH_4OAc (pH 6.5) (*top*) and with 10× concentration of spermine (50 μm) (*bottom*). Peaks labeled with the *filled circles* represent the 1:1 AS:spermine complex.

solution concentrations of the unbound protein and protein–ligand complex, respectively.

(b) The solution concentration of the unbound protein plus the concentration of protein–ligand complex equals the total concentration of the protein in solution.

(c) For a 1:1 protein–ligand interaction, Protein + Ligand ↔ Protein–Ligand, and the dissociation constant is $K_d = [\text{Protein}_{\text{free}}]*[\text{Ligand}_{\text{free}}]/[\text{protein–ligand}]$.

(d) $[\text{protein–ligand}] = [\text{protein}_{\text{total}}]*I_{\text{complex}}/(I_{\text{protein}} + I_{\text{complex}})$.

(e) $\text{Protein}_{\text{free}}] = [\text{Protein}_{\text{total}}] - [\text{protein–ligand}]$.

(f) $[\text{Ligand}_{\text{free}}] = [\text{Ligand}_{\text{total}}] - [\text{protein–ligand}]$.

7. Plot the *x*-axis as the concentration of $[\text{Ligand}_{\text{free}}]$.
8. Plot the *y*-axis as $[\text{protein–ligand}]/[\text{protein}_{\text{free}}]$
9. The slope of the line measured should be the solution K_d of the protein–ligand interaction. **Figure 4** shows such a plot for the titration of spermine ligand to the AS protein. A K_d of 0.36 mM is calculated from the ESI-MS titration. NMR experiments by Fernandez et al. *(10)* measured a K_d of 0.61 ± 0.03 mM for AS-spermine.

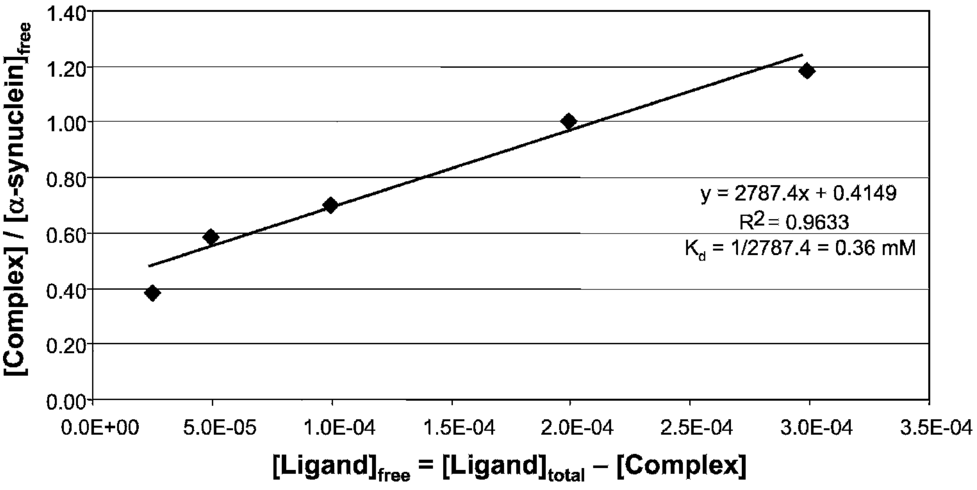


Fig. 4. ESI qqtofms titration experiments of α-synuclein (AS) with various concentrations of spermine in 10-mm NH_4OAc (pH 6.5). From the slope of the line, a solution K_d of 0.36 mm is calculated for the 1:1 AS:sperimine complex.

4. Notes

1. For protein sample preparation, be sure to check the manufacturer's protocol for proper use of the centrifugal membrane filtration device. Some solvents and chemicals are not compatible with the membranes used in the filtration devices.
2. The final concentration of the protein sample should typically be around or higher than 5 μM. For samples prepared from cell culture, it may be difficult to obtain sufficient quantities of purified protein. The centrifugal filtration device could help to concentrate sample concentration.
3. For preparing protein–ligand solutions, often the ligand sample itself contains high concentration of salts. For this case, one could mix the protein sample and ligand prior to centrifugal filtration, and desalt the protein–ligand mixture.
4. Often, it is advantageous to store cleaned, desalted protein stock solutions into 10–20 μL aliquots, and store them in a -80°C freezer prior to use.
5. Each type of mass spectrometer and each protein complex have their unique set of experimental conditions for optimal sensitivity and resolution. The greatest variability may be found in the potentials used for declustering or desolvating the ions for the protein complexes (without dissociating the complex). Some ESI interfaces require increased pressure in the interface to collisionally cool and focus the complex ions. Interfaces that utilize a metal or glass ion transport tube have variable

operating temperatures; protein complexes often require lower tube temperatures. A good protein complex standard to optimize instrument performance is the horse myoglobin system, a complex between the apoprotein (16,952 Da) and protoporphyrin IX (heme, 616 Da). Multiply charged ions with 7–10 positive charges between m/z 1,700 and 2,510 should be observed with good sensitivity and resolution *(5)*.

Acknowledgments

We are grateful for the support for the UCLA Functional Proteomics Center provided by the W. M. Keck Foundation, and the funding of our research by the National Institutes of Health (RR 20004 to JAL).

References

1. Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* 16, 1–23.
2. Loo, J. A., Hu, P., McConnell, P., and Mueller, W. T. (1997) A study of Src SH2 domain protein–phosphopeptide binding interactions by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* 8, 234–43.
3. Heck, A. J. R., and van den Heuvel, R. H. H. (2004) Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* 23, 368–89.
4. Sobott, F., Hernandez, H., McCammon, M. G., Tito, M. A., and Robinson, C. V. (2002) A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal. Chem.* 74, 1402–07.
5. Loo, J. A. (2000) Electrospray ionization mass spectrometry: A technology for studying noncovalent macromolecular complexes. *Int. J. Mass Spectrom.* 200, 175–86.
6. Loo, J. A., Berhane, B., Kaddis, C. S., Wooding, K. M., Xie, Y., Kaufman, S. L., and Chernushevich, I. V. (2005) Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. *J. Am. Soc. Mass Spectrom.* 16, 998–1008.
7. Loo, J. A., Holler, T. P., Foltin, S. K., McConnell, P., Banotai, C. A., Horne, N. M., Mueller, W. T., Stevenson, T. I., and Mack, D. P. (1998) Application of electrospray ionization mass spectrometry for studying human immunodeficiency virus protein complexes. *Proteins: Struct., Funct., Genet.* Suppl. 2, 28–37.
8. Loo, J. A. (2000) Probing protein–metal ion interactions by electrospray ionization mass spectrometry: Enolase and nucleocapsid protein. *Int. J. Mass Spectrom.* 204, 113–23.
9. Zhang, S., Pelt, C. K. V., and Wilson, D. B. (2003) Quantitative determination of noncovalent binding interactions using automated nanoelectrospray mass spectrometry. *Anal. Chem.* 75, 3010–18.
10. Fernández, C. O., Hoyer, W., Zweckstetter, M., Jares-Erijman, E. A., Subramaniam, V., Griesinger, C., and Jovin, T. M. (2004) NMR of α-synuclein–polyamine complexes elucidates the mechanism and kinetics of induced aggregation. *EMBO J.* 23, 2039–46.

Chapter 17

Chemical Cross-Linking for Protein–Protein Interaction Studies

Xiaoting Tang and James E. Bruce

Summary

Most proteins function through protein complex assemblies. Defining and mapping protein complex networks are crucial elements in the fundamental understanding of biological processes. The ability to measure protein–protein interactions in biological systems has undergone significant advances in the past decade due to emergence and growth of numerous new molecular biology and mass spectrometry technologies. Chemical cross-linking, along with yeast two-hybrid, fluorescence resonant energy transfer (FRET), and co-immunoprecipitation have become important tools for detection and characterization of protein–protein interactions. Individual protein members in a noncovalent complex assembly remain in close proximity which is within the reach of the two reactive groups of a cross-linker. Thus cross-linking reactions have potential for linking two interacting proteins which exist in close proximity. In general, chemical cross-linking experiments are carried out by first linking the interacting proteins through covalent bonds followed by a series of well-established protocols – SDS-PAGE, in-gel digestion, and shotgun LC/MS/MS for identification of the cross-linked proteins. These approaches have been employed for both mapping topology of protein complex *in vitro* and determining the protein interaction partners *in vivo*.

Key words: Chemical cross-linking, Protein–protein interactions, Topology, Protein complex, Mass spectrometry, Protein identification.

1. Introduction

Cross-linking is a process that links two proteins or two regions in a protein through covalent bonds, which are formed via chemical reactions between the end reactive groups of cross-linkers and the functional groups on proteins. The common reactive groups employed in the cross-linkers include *N*-Hydroxysuccinimide (NHS) esters (amine-reactive), maleimides (sulfhydryl-reactive),

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_17

and nonspecific photoreactive groups such as aryl azides, benzophenones, diazirines, etc. The most widely used cross-linkers are those homobifunctional cross-linkers containing two NHS esters, since (a) NHS esters have high reaction efficiency; (b) NHS esters yield stable amide bonds with primary amines in proteins upon reactions; (c) most proteins have many lysine residues; and (d) the cross-linking reactions occur under physiological conditions (pH 7–8), where most protein complexes exist in their native state.

Chemical cross-linking methods in combination with mass spectrometry have become increasingly important tools for mapping low-resolution topology of proteins and for studying protein–protein interactions *(1–4)*. These approaches are either employed to study spatial relationships in purified protein complexes *(5–9)* or coupled with immunoaffinity purification methods to probe the potential binding partners for a specific protein of interest in living cells *(10–13)*. Mass spectrometric analysis of cross-linked protein complex allows identification of the protein interacting partners and in principle, could lead to identification of interaction regions between proteins provided that the labeling sites are identified. However, the identification of the sites of interaction or cross-linker labeling sites are highly challenging due to the complexity inherent with cross-linking approaches; this has been only achieved for purified protein complex systems when relatively large quantity of pure proteins are available. For *in vivo* applications, most cross-linking approaches are primarily used to determine the interacting partners for a specific protein of interest with the aid of immunoaffinity purifications, SDS-PAGE separations, Western blot visualizations, and mass spectrometric analysis.

2. Materials

2.1. Cross-Linking Reaction

1. Deionized water (*see* **Note 1**).
2. Dry dimethylsulfoxide (DMSO), HPLC grade.
3. Cross-linking reagents: NHS ester-based homobifunctional cross-linkers, e.g., bis[sulfosuccinimidyl] suberate (BS^3), dithiobis[succinimidyl propionate] (DSP), dissuccinimidyl tartrate (DST), etc. (Pierce, Rockford, IL) (*see* **Note 2**).
4. Biological samples: Purified protein complex and cell culture (*see* **Note 3**).
5. Reaction buffer: Phosphate buffered saline (PBS) (0.1 M sodium phosphate, 0.15 M NaCl, pH 7.4), or other nonprimary amine containing buffer such as Hepes, and bicarbonate/carbonate at pH 7.4 (*see* **Note 4**).

6. Quench buffer: 1 M Tris (pH 7.5), or other primary amine containing buffer such as 1.0 M glycine in PBS and 1 M NH_4HCO_3.
7. Cell lysis buffer: 50 mM Tris (pH 7.5), 150 mM NaCl, 10% glycerol, 1% NP-40, 5 mM EDTA, protease inhibitor cocktail (Sigma, St. Louis, MO).

2.2. Sds-page

1. 4× sample buffer: 200 mM Tris-HCl (pH 6.8), 8% SDS, 0.4% bromophenol blue, and 40% glycerol.
2. 10× running buffer: 0.25 M Tris, 1.9 M glycine, and 4% SDS.
3. Precast polyacrylamide gels with various percentages (Bio-Rad, Hercules, CA).
4. Prestained molecular weight markers (Bio-Rad, Hercules, CA).
5. Coomassie blue staining solution (Bio-Rad, Hercules, CA).

2.3. In-Gel Solution

1. Destaining solution: methanol (50%, v/v), acetic acid (5%, v/v), and water (45%, v/v).
2. Digestion buffer: 50 mM NH_4HCO_3, pH 7.8.
3. Dehydration solution: acetonitrile, HPLC grade.
4. Reducing and alkylating reagent: 10 mM DTT and 50 mM iodoacetamide.
5. Trypsin (Sequencing grade, Promega, Madison, WI) solution: 20 μg/mL dissolved in the digestion buffer. Prepare freshly before use.

2.4. Western Blot

1. Transfer buffer: 25 mM Tris, 190 mM glycine, 20% (v/v) methanol, and 0.05% (w/v) SDS.
2. Nitrocellulose membrane (Millipore, Bedford, MA).
3. 10× Tris-buffered saline with Tween (TBS-T) buffer: 1.37 M NaCl, 27 mM KCl, 250 mM Tris-HCl, pH 7.4, 1% Tween-20.
4. Blocking buffer: 5% (w/v) nonfat dry milk in TBS-T.
5. Primary antibody (*see* **Note 5**).
6. Secondary antibody, e.g., if the primary antibody is raised from mouse, antimouse IgG conjugated to horse radish peroxidase (Sigma, St. Louis, MO) can be used.
7. Enhanced chemiluminescent (ECL) reagents (Pierce, Rockford, IL).
8. Stripping buffer: 62.5 mM Tris-HCl, pH 6.8, 2% (w/v) SDS.

2.5. Immunoprecipitation

1. Seize X immunoprecipitation kits including protein A or G agarose support and cross-linking (DSS) reagent (Pierce, Rockford, IL) (*see* **Note 6**).
2. Equilibration buffer: 20 mM Tris (pH 7.5) 100 mM NaCl, 0.1 mM EDTA.

3. Wash buffer: 20 mM Tris (pH 7.5) 150 mM NaCl, 10% glycerol, 1% Triton X-100, 2 mM EDTA.
4. Elution buffer: 100 mM glycine (pH 3)
5. Neutralization buffer: 1 M Tris (pH 8)

3. Methods

The *N*-Hydroxysuccinimide ester (NHS ester) is the most commonly used reactive group for cross-linking because of its high reactivity and specificity. The reaction of NHS ester with the ε-amine in the side chain of lysine (K) or the α-amine of N-terminus at pH 7–9 results in the release of *N*-hydroxysuccinimide and formation of stable amide bond. However, NHS esters can also react with water or ammonia under the same conditions and the rate of hydrolysis or aminolysis increases with increasing pH and temperature *(14)*. Most NHS ester-based cross-linkers are water-insoluble and are thus, introduced to the reaction media by diluting the stock solution of the reagent dissolved in pure organic solvent. However, the solubility of cross-linker can be greatly increased with the inclusion of charged-sulfonate group on NHS ester. Water-insoluble cross-linkers are mostly used for intramembrane or intracellular cross-linking applications since they can penetrate the lipid bilayers of the cell membranes; while water-soluble cross-linkers are utilized to cross-link the cell surface or extracellular proteins. Overall, selection of the type of cross-linker and reaction conditions is determined by the specific requirements for each individual application.

Cross-linking approaches have been applied to map the low-resolution three-dimensional structures for a particular protein, determine spatial relationships between proteins in a multiprotein assembly, and detect potential interacting partners for a target protein. In general, these applications can be divided into two primary categories (a) *in vitro* application using purified proteins/protein complexes and (b) *in vivo* application in combination with immunoaffinity purification methods. Here we describe both types of applications with NHS ester-based cross-linkers as examples.

3.1. In Vitro Application

3.1.1. Cross-Linking Reaction

1. Dissolve cross-linker in dry DMSO at 100 mM (*see* **Note 7**).
2. Prepare the purified protein complex in the reaction buffer. If the protein complex or its individual protein component is available in pure power forms, prepare 1 mM stock solution in water for each protein and then dilute each into 10–20 μM in the reaction buffer placed in the same tube. Equilibrate the mixture at room temperature (RT) for 30 min. If the

protein complex is purified through immunoprecipitation or other procedure, perform a protein assay for the final elution solution with BCA protein assay kit (Pierce, Rockford, IL) or Bradford assay kit (Bio-Rad, Hercules, CA). Exchange the elution buffer with the reaction buffer using gel filtration spin column (Chroma Spin Columns, Clontech, Mountain View, CA) or dialysis. Prepare a final concentration of 5–10 mg/mL proteins in the reaction buffer (*see* **Note 8**).

3. To determine the optimal concentration of cross-linker, add 1, 5, 10, and 50 μL of 100 mM cross-linker solution to the reaction tube and make 1 mL of final reaction mixture. The final concentration of the cross-linker is 0.1, 0.5, 1.0, and 5.0 mM, respectively. Incubate the mixture with shaking for 1 h on ice.
4. Quench the reaction by adding 50 μL of 1 M Tris (pH 7.5). Incubate for another 5 min at RT.
5. A separate control sample is prepared in parallel without the use of any cross-linker (*see* **Note 9**).

3.1.2. SDS-PAGE Isolation of Cross-Linked Products

1. Take an aliquot of reaction and control mixture solution, add 4× sample buffer, and then heat at 70°C for 10 min.
2. Load proteins onto appropriate percentage SDS-PAGE gel for separation (*see* **Note 10**).
3. A prestained protein molecular weight marker is used to determine the stop time for the electrophoresis process. After separation, take the gels out and wash extensively with water.
4. Stain the gels with Commassie blue and then destain with water.
5. The cross-linked products can be identified by comparing the cross-linked sample lane with the control lane. The cross-linked product is usually observed as a new band and higher molecular weight band distinctive only to the cross-linked sample lane. The amount of the cross-linked product is indicated by the intensity of the band.
6. After the optimal concentration of cross-linker is determined, the same procedure described above can be used to investigate other reaction conditions, e.g., the reaction time, pH, temperature, etc.

3.1.3. In-Gel Digestion

1. Wear gloves and take special care for each step to avoid introducing any keratin contaminations from dust, hair, and hands (*see* **Note 11**).
2. Repeat the cross-linking reactions under the optimized conditions and prepare gel-separated cross-linked products in as large quantity as possible (*see* **Note 12**).

3. Excise the band of interest, split the gel slice evenly to 3–4 pieces and then place in a siliconized microfuge tube.
4. Remove excess Coomassie Blue stain by washing the gel pieces with 1 mL of destaining buffer at RT for 1 h. Repeat if needed.
5. Dehydrate the gel slices in 200 μL of acetonitrile for 5 min. Repeat one more time. The fully dehydrated gels appear white and crystalline.
6. Reduce proteins in the gel pieces in 100 μL of 10 mM DTT for 30 min at RT and then alkylate in 100 μL of 50 mM iodoactemide at 37°C in dark for 30 min.
7. Repeat **step 5** to wash and dry the gels.
8. Prepare trypsin in 20 ng/μL by reconstituting 20 μg trypsin in 1,000 μL cold digestion buffer.
9. Add 50–100 μL of the trypsin solution to cover the gel pieces and rehydrate the gels on ice for 5–10 min to allow gels to fully swell (*see* **Note 13**).
10. Remove excess trypsin solution and add 20–30 μL of digestion buffer. Incubate overnight at 37°C (*see* **Note 14**).
11. Take the supernatant out to another tube for MS analysis (*see* **Note 15**).

3.2. In vivo Application

3.2.1. Intact Cell Cross-Linking

1. Harvest cells and wash cells with ice-cold reaction buffer 3–5 times to remove culture media (*see* **Note 16**).
2. Resuspend the pelleted cells in the reaction buffer at ~5 × 10^6 cells/mL concentration.
3. Add 1, 10, and 100 μL of 100 mM cross-linker to 1 mL of cell suspension to determine the optimal cross-linker concentration.
4. Incubate the mixture on ice for 1 h with mild shaking.
5. Stop the reaction with 50 μL of quench buffer and incubate for another 5 min at RT.
6. A separate control sample is prepared in parallel without the use of cross-linker.

3.2.2. Cell Lysis

1. Wash the labeled cells with ice-cold reaction buffer 3–5 times to remove excess unreacted cross-linkers. Resuspend the cell pellets in lysis buffer and lyse cells with vortexing or sonication.
2. Centrifuge the cell lysates for 15 min at 15,000 rpm to pellet cell debris.
3. Collect the supernatant and perform protein assay using BCA protein assay kit (Pierce, Rockford, IL) or Bradford assay (Bio-Rad, Hercules, CA) according to the manufacture's protocol.

3.2.3. SDS-PAGE and Western Blot

1. Load an aliquot of cross-linked samples and control sample to SDS-PAGE gel. Run gel electrophoresis until the designated proteins reach the bottom of the gel with the aid of prestained marker.
2. Transfer proteins from the gel to a nitrocellulose membrane with a Trans-Blot® semidry transfer cell (Bio-Rad, Hercules, CA).
3. After transferring, the membrane is blocked with blocking buffer overnight at 4°C followed by incubating with the primary antibody at 1:5,000 dilution for 1 h at RT.
4. Wash the membrane with TBS-T buffer three times, and then incubate with secondary antibody for 1 h at RT. Wash with TBS-T buffer for another three times
5. Membrane is visualized with ECL reagents.
6. The membrane can be stripped and then reprobed with different primary antibodies to verify the existence of other interacting proteins. Heat 50–100 mL of stripping buffer to 70°C and then add 100 mM β-mercaptoethanol and the membrane. Incubate for 30 min with shaking. Wash with 0.1% (w/v) BSA in TBS-T three times and block again in the block buffer. The membrane is then ready to be reprobed with a different primary antibody as described above.
7. Optimize other conditions such as reaction time, pH, and temperature using the procedures described above.

3.2.4. Immunoprecipitation for Small-Scale Analysis

1. Prepare the antibody (antitarget protein or antitag on target protein) beads by coupling 2 mg antibody to 1 mL of protein A or G agarose beads with Seize X immunoprecipitation kits (Pierce, Rockford, IL) according to manufacture's protocol.
2. Prepare cross-linked cell lysates under the optimized conditions.
3. Cell lysates are precleared with 50 μL of protein A or G agarose beads for 10 min at 4°C.
4. Incubate precleared cell lysates with antibody beads for 2 h at 4°C. Use 10–20 μL of beads for 1 mg of total proteins.
5. Wash antibody beads with 1 mL of wash buffer five times.
6. Following the final wash, carefully remove all traces of the supernatant by aspiration.
7. Add 100 μL of elution buffer and incubate for 15 min at 37°C. Transfer the elution solution to another tube.
8. Neutralize the elution solution with 5 μL of 1 M Tris (pH 8).
9. Add 4× sample buffer and heat at 70°C for 10 min prior to loading to SDS-PAGE.
10. Run SDS-PAGE and in-gel digestion as described above.

3.2.5. Immunoaffinity Chromatography for Large-Scale Analysis

1. Load 1 mL of antibody beads into an empty chromatography column (Poly-Prep columns, Bio-Rad, Hercules, CA).
2. Wash with 1 mL of 100 mM glycine (pH 3) three times.
3. Pass 10 mL of equilibration buffer through the column.
4. Prepare cross-linked cell lysates under the optimized conditions in large scale (e.g., use 10^9 cells).
5. Cell lysates are precleared with 1 mL of protein A or G agarose beads for 10 min at 4°C.
6. Load 20 mL of precleared lysates onto the column (*see* **Note 17**).
7. Pass 20 mL of ice-cold equilibration buffer through the column. Repeat twice.
8. Wash the column with the 20 mL of ice-cold wash buffer three times.
9. Elute the bound material with 6 mL of 100 mM glycine (pH 3).
10. Neutralize the eluate with 300 μL of 1 M Tris (pH 8).
11. Concentrate the eluate solution using Centricon ultrafiltration column (Millipore, Bedford, MA) according to the manufacturer's protocol (*see* **Note 18**). Reduce to a small volume of ~50 μL.
12. Add 4× sample buffer and heat at 70°C for 10 min prior to loading to SDS-PAGE (*see* **Note 19**).
13. Perform SDS-PAGE and in-gel digestion as described above.

4. Notes

1. All water used in this study is deionized water that has a resistivity of 18-MΩ and total organic content of less than five parts-per-billion.
2. All NHS ester-based cross-linkers are moisture sensitive. Store desiccated at 4°C. Equilibrate the sample vial to RT before opening to avoid moisture condensation. It takes 20–30 min for equilibration.
3. This protocol is written for general applications of *in vitro* and *in vivo* cross-linking methods. Minor modifications may be needed for different biological samples. For *in vitro* application, some protein complex can be obtained through reconstitution of lyophilized powder in the appropriate buffer (e.g., ribonuclease S complex *(9)*). In some cases, each

individual protein member for a protein complex is available in the purified form and the protein complex can be reconstituted by mixing all the members in the appropriate buffer (e.g., calmodulin and melittin *(15)*). For the protein complex not available in pure forms, immunoprecipitation method is used to purify the complex if an antibody is available for a member of the complex or the tag engineered in a member of the complex (e.g., Nup84p complex *(16)*). For *in vivo* application, cells can be harvested at 80% confluence (for mammalian cells) or at mid-log phase (for bacterial cells).

4. Avoid any reaction buffer that contains primary amines such as Tris, glycine, etc. If the buffer composition is not compatible with cross-linking, use gel filtration or dialysis to exchange buffer with reaction buffer. Mild pH, low-salt, and detergent-free buffer conditions are preferred to maintain the native structure of the protein complex.
5. The selected primary antibody can be reactive against target protein, a tag engineered on target protein, or a member of the protein complex.
6. The primary purpose of immobilizing antibody to protein A or G agarose beads is to reduce antibody contamination. Other comparable antibody immobilizing method can also be used.
7. Always prepare fresh cross-linker solution. Do not use frozen stocks. Water-soluble cross-linker can be prepared in water. However, preparation of stock solution in DMSO increases the stability of cross-linker.
8. Freezing solutions of protein complexes is not advised, since freeze-thaw process often results in partial denaturation. Rather, protein complex can be stored at –80°C after cross-linking.
9. Control sample is used to distinguish cross-linked products from non-cross-linked products.
10. The percentage of the gel is determined by the size of expected cross-linked products. Note that cross-linked proteins have larger molecular weight than uncross-linked proteins, thus low-percentage gels (8–10%) are often used.
11. Keratins are commonly observed contaminations for protein identification using in-gel digestion and mass spectrometric analysis. Severe contamination can lead to false or no identification especially when cross-linked product is in low quantity.
12. For the purpose of mass spectrometric analysis of cross-linked products, separation of large quantity of proteins is preferred. In this case, the entire reaction mixture can be concentrated using TCA precipitation or Centricon ultrafiltration column (Millipore, Bedford, MA) before loading to the gel.

13. The purpose of this step is to allow trypsin to penetrate into the gels with no autolysis.
14. Remove the excess trypsin to reduce the contamination peaks resultant from the autolysis of trypsin. After overnight digestion, more than 80% peptides are extracted to the digestion buffer. Keep the volume of digestion buffer as small as possible to ensure the peptide concentration is sufficiently high for subsequent MS analysis.
15. Mass spectrometry-based protein identification is achieved either by peptide mass fingerprinting of three or more peptides or tandem MS fragmentation of a few peptides from an enzymatic digestion. The instruments suitable for protein identification include MALDI-TOF-MS, ESI-Ion Trap-MS, ESI-Q-TOF-MS, ESI-FTICR-MS, etc.
16. Some culture media have protein additives such as bovine serum albumin. Some have components not compatible to cross-linking reactions. Thus a thorough intact cell wash is necessary in many cases to remove all the interferences.
17. Let cell lysates pass through the column slowly by gravity. Do not let the beads dry out at any point.
18. Volume reduction is necessary to ensure that sufficient amount of proteins is loaded on to SDS-PAGE. To reduce sample losses due to adsorption of protein onto the ultrafiltration membrane, the membrane can be precoated by concentrating a BSA solution to 1 mg/mL and then washing the membrane three times.
19. Recover some hydrophobic proteins that might stick to the ultrafiltration membrane by washing it with 20 μL of 4× sample buffer.

Acknowledgments

This research was supported by the Office of Science (BER), US Department of Energy, Grant No. DE-FG02–04ER63924 and NIH-NCRR, Grant No. 1 R01 RR023334-01A1.

References

1. Back, J. W., de Jong, L., Muijsers, A. O., and de Koster, C. G. (2003) Chemical cross-linking and mass spectrometry for protein structural modeling. *J Mol Biol* 331, 303–13.
2. Trakselis, M. A., Alley, S. C., and Ishmael, F. T. (2005) Identification and mapping of protein–protein interactions by a combination of cross-linking, cleavage, and proteomics. *Bioconjugate Chem* 16, 741–50.
3. Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. *Mass Spectrom Rev* 25, 663–82.

4. Vasilescu, J., and Figeys, D. (2006) Mapping protein–protein interactions by mass spectrometry. *Curr Opin Biotechnol* 17, 394–9.
5. Trester-Zedlitz, M., Kamada, K., Burley, S. K., Fenyo, D., Chait, B. T., and Muir, T. W. (2003) A modular cross-linking approach for exploring protein interactions. *J Am Chem Soc* 125, 2416–25.
6. Chang, Z., Kuchar, J., and Hausinger, R. P. (2004) Chemical cross-linking and mass spectrometric identification of sites of interaction for UreD, UreF, and urease. *J Biol Chem* 279, 15305–13.
7. Petrotchenko, E. V., Olkhovik, V. K., and Borchers, C. H. (2005) Isotopically coded cleavable cross-linker for studying protein–protein interaction and protein complexes. *Mol Cell Proteomics* 4, 1167–79.
8. Kalkhof, S., Ihling, C., Mechtler, K., and Sinz, A. (2005) Chemical cross-linking and high-performance fourier transform ion cyclotron resonance mass spectrometry for protein interaction analysis: Application to a calmodulin/target peptide complex. *Anal Chem* 77, 495–503.
9. Tang, X., Munske, G. R., Siems, W. F., and Bruce, J. E. (2005) Mass spectrometry identifiable cross-linking strategy for studying protein–protein interactions. *Anal Chem* 77, 311–8.
10. Chen, Y., Ebright, Y. W., and Ebright, R. H. (1994) Identification of the target of a transcription activator protein by protein–protein photocrosslinking. *Science* 265, 90–2.
11. Wells, J., and Farnham, P. J. (2002) Characterizing transcription factor binding sites using formaldehyde crosslinking and immunoprecipitation. *Methods* 26, 48–56.
12. Schmitt-Ulms, G., Hansen, K., Liu, J., Cowdrey, C., Yang, J., DeArmond, S. J., Cohen, F. E., Prusiner, S. B., and Baldwin, M. A. (2004) Time-controlled transcardiac perfusion cross-linking for the study of protein interactions in complex tissues. *Nat Biotechnol* 22, 724–31.
13. Vasilescu, J., Guo, X., and Kast, J. (2004) Identification of protein–protein interactions using in vivo cross-linking and mass spectrometry. *Proteomics* 4, 3845–54.
14. Anjaneyulu, P. S., and Staros, J. V. (1987) Reactions of *N*-hydroxysulfosuccinimide active esters. *Int J Pept Protein Res* 30, 117–24.
15. Schulz, D. M., Ihling, C., Clore, G. M., and Sinz, A. (2004) Mapping the topology and determination of a low-resolution three-dimensional structure of the calmodulin–melittin complex by chemical cross-linking and high-resolution FTICRMS: Direct demonstration of multiple binding modes. *Biochemistry* 43, 4703–15.
16. Rappsilber, J., Siniossoglou, S., Hurt, E. C., and Mann, M. (2000) A generic strategy to analyze the spatial organization of multiprotein complexes by cross-linking and mass spectrometry. *Anal Chem* 72, 267–75.

Chapter 18

Tissue Analysis with High-Resolution Imaging Mass Spectrometry

A.F. Maarten Altelaar and Ron M.A. Heeren

Summary

Molecular mass spectrometric images of tissue sections facilitate precise location of biomolecules, drugs, their metabolites and the biomolecular aberrations they target. Technological developments are rapidly expanding the capabilities of imaging mass spectrometers. Speed, resolution, sensitivity, and sample preparation protocols are no longer limiting factors in its application.

Key words: Imaging mass spectrometry, Surface analysis, SIMS, MALDI, Tissue analysis.

1. Introduction

The molecular landscape of living systems is continuously changing *(1–3)*. The objective of many modern microscopic technologies is to generate snapshots of this molecular landscape. Imaging mass spectrometry (IMS) can visualize many different classes of molecules directly on histopathological tissue sections without the use of molecular labels *(4)*. A single IMS experiment thus provides different snapshots of the molecular landscape. This unique capability distinguishes IMS from the battery of microscopic imaging techniques used in the life sciences. Key to the success of this mass spectrometry-based technology is a proper preparation of the sample surface to make it suitable for biomolecular surface analysis *(5,6)*.

Unlike conventional microscopic imaging techniques, mass spectrometry-based imaging requires the generation of ions from the surface under study. There are essentially two methods that

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_18

can generate molecular ions directly from surfaces (a) using an energetic beam of primary charged particles referred to as SIMS imaging *(7)* and (b) using a laser beam where photons deliver the energy required for desorption and ionization referred to as LDI-imaging. If these two techniques are applied to an unmodified surface, a significant amount of rovibrational energy is deposited in the top-layer of the sample. This results in extensive fragmentation of the surface molecules. A layer of matrix crystals can be applied to a tissue surface to reduce this effect and to allow the generation of intact molecular ions directly. These matrix-aided techniques are referred to as matrix-enhanced SIMS (ME-SIMS) *(8,9)* or matrix-assisted laser desorption and ionisation (MALDI) *(10,11)*, respectively. The role of the matrix crystals is believed to be fourfold: first it serves to isolate analyte molecules by dilution which prevents aggregation, second it acts as a strong absorber of the laser or primary ion energy through electronic or vibrational excitation, third it facilitates a nondestructive transition of the analyte molecules to the gas phase, and fourth it promotes the ionization of the analyte molecules by charge transfer.

ME-SIMS imaging and MALDI-IMS are two complementary chemical imaging techniques that generate molecular snapshots of biological surfaces. Combined they can visualize distributions of elements *(12–14)*, small molecules, such as pharmaceutical compounds *(15,16)*, peptides *(17,18)*, and proteins *(19–21)* at different levels of spatial resolution on complex surfaces. Where SIMS analysis typically achieves ;1 μm spatial resolution *(7)* in conventional MALDI-IMS, pixel sizes of 25–250 μm *(4,21)* are standard. In order to be able to image molecular distributions at cellular lengthscales using MALDI-IMS, a new approach has been developed *(22)*. Using a stigmatic mass spectrometric microscope (mass microscope), which decouples the obtainable spatial resolution from the dimensions of the footprint of the ionization beam, features of ~4 μm can be distinguished. Here the desorbed ions retain their original spatial distribution from the tissue surface during their time-of-flight (ToF) separation and are imaged by a two-dimensional position sensitive detector. As can be seen from **Fig. 1**, the spatial detail observed in the stigmatic total-ion-count (TIC) image is very high compared to the microprobe data. The insets in the figure show a magnification of a specific area in the tissue, which results in a 100% image ratio (i.e., the real size) for the stigmatic image. Here it can be seen that spatial resolution at cellular lengthscales can be achieved using the mass microscope.

Two complementary methods are described, matrix-enhanced secondary ion imaging mass spectrometry (ME-SIMS) and matrix-assisted laser desorption and ionization imaging mass spectrometry (MALDI-IMS).

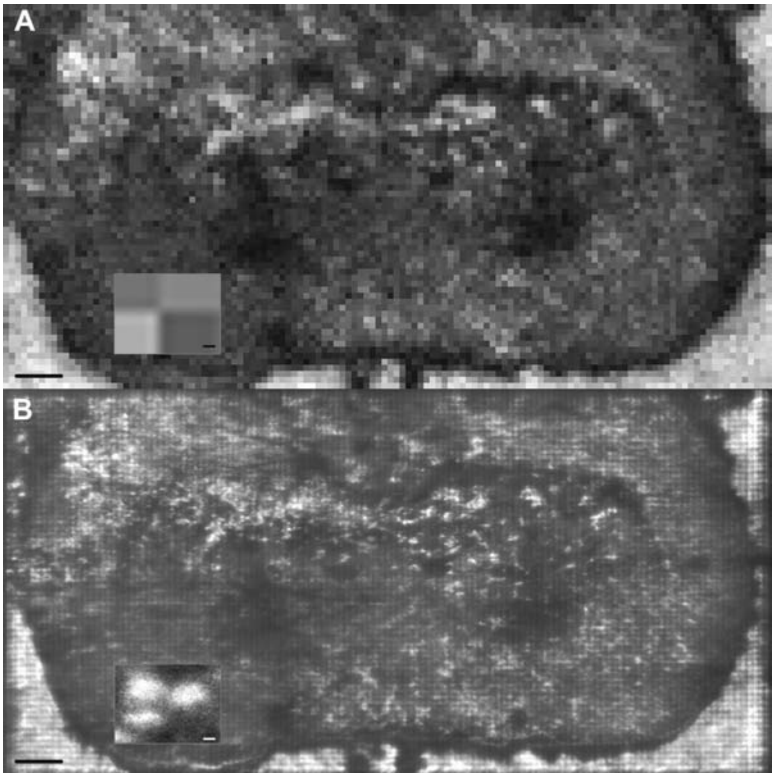


Fig. 1. Total-ion-count (TIC) image of a rat brain tissue section obtained with (**a**) microprobe MALDI-IMS and (**b**) microscope MALDI-IMS. Insets show a magnification of the area on the tissue within the *red square*. Scale bar in the images are 1 mm and in the insets are 25 μm.

2. Materials

2.1. Tissue Sections

1. Dissection microscope allowing low magnification (10×/40×) (CETI, Antwerpen, Belgium).
2. Dissected tissue is frozen immediately in liquid isopentane, cooled on dry ice and stored at −80°C until use (*see* **Note 1**).
3. When embedding is needed (cutting of very small organs can result in damage of the tissue sections) embed the organs in compatible embedding material, right after dissection. Compatible embedding materials are nonpolymer containing solutions like 10% Gelatine (De Twee Torens, Delft, The Netherlands).
4. Cryomicrotome; Leica CM 3000 cryostat (Leica Microsystems, Nussloch, Germany).

5. Tissue-Tek O.C.T. Compound (Sakura Finetek USA).
6. Conductive glass slides; 25 × 50 × 1.1 mm unpolished float glass, SiO_2 passivated/indium-tin-oxide coated, $R_s = 6 \pm 2\ \Omega$ (Delta Technologies, Stillwater, MN).
7. Dry ice for transport.

2.2. Tissue Preparation

1. Washing solution; Ethanol (Biosolve, The Netherlands) and HPLC grade water (Riedel-de Haën, Zwijndrecht, The Netherlands) are mixed in a 70:30 ratio and placed on ice.
2. Matrix solution; α-cyano-4-hydroxycinnamic acid (HCCA) or 2,5-dihydroxybenzoic acid (DHB) (Fluka, Zwijndrecht, The Netherlands) is dissolved in a 50:50 solution of HPLC grade water with either ethanol or acetonitrile (Biosolve, The Netherlands), containing 0.1% trifluoracetic acid (TFA) (Sigma-Aldrich, Zwijndrecht, The Netherlands). Store solutions at 4°C, DHB is light sensitive.
3. Gold coating: On top of samples prepared for MALDI analysis, a ~4 nm layer of gold is deposited using a Quorum Technologies (Newhaven, East Sussex, UK) SC7640 sputter coater equipped with a gold target (SC510-314A).

2.3. Matrix Deposition

1. *Home-built electrospray deposition (ESD) setup.* In this setup, a syringe pump (KD Scientific) pumps matrix solution (10–50 μL/min) from a gastight syringe (Hamilton) through a stainless steel electrospray capillary (O.D. 220 μm, I.D. 100 μm) maintained at 3–5 kV (0–6 kV power supply, Heinzinger). The capillary is mounted on an electrically isolated manual translation stage (Thorlabs) in a vertical orientation. The stage is fitted with a digital micrometer (Mitutoyo) for accurate positioning of the needle tip with respect to the grounded sample plate. The sample plate is mounted on a *x*,*y* moveable table (Thorlabs).
2. Home-built pressure driven spray setup (SprayStation, Fig. 2). In this setup, the matrix solution (stored in pressure resistant glass bottles (SchottDuran)) is pressure driven (1 bar) toward a stainless steel spray valve (TS5540, Unitek Eapro) with adjustable nozzle. The opening of the nozzle is controlled by a combination of a "flow control knob", a return spring and air pressure (5 bar). The position of the knob determines the pressure on the spring, closing the nozzle, while the air pressure is used to counter the pressure of the knob and open the nozzle. A solvent spray is realized by the opening of the nozzle in combination with an atomizing airflow (1 bar) coming in from the side of the spray valve, just above the nozzle. The spray valve is mounted on a Perseptive Biosystems Symbiot for control of the *x*,*y*,*z* movement. The whole setup is placed in a closed Perspex box equipped with a hygrostat, humidifier,

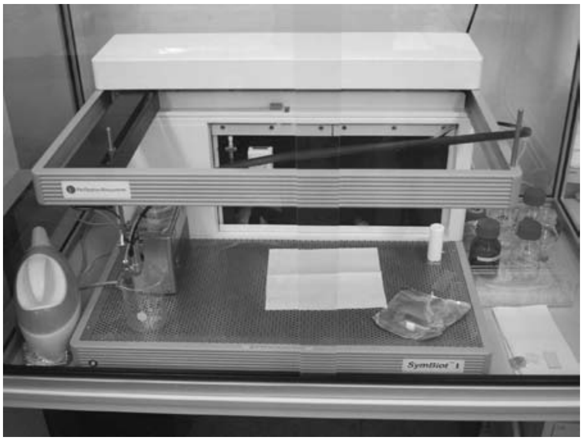


Fig. 2. Photograph of the in-house developed air driven spray station.

and fan to control the humidity during matrix deposition. The nozzle is cleaned after every spray cycle by placing the whole spray valve in a sonicating washing solvent and, after a set number of spray cycles, by flushing through the nozzle with 50% ethanol washing solvent (1 bar).

3. *TLC sprayer*. 10 mL Chromatography sprayer (Sigma-Aldrich, Zwijndrecht, The Netherlands).

2.4. Mass Spectrometry

1. Physical Electronics (Eden Prairie, MN) TRIFT-II mass spectrometer equipped with a MCP/phosphor screen/CCD camera (LaVision, Germany) optical detection combination and a MALDI ionization source (described in detail by Luxembourg et al. *(22)*). The instrument is equipped with a DP214 digitizer card with 1 GHz bandwidth and 2 GS/s sampling rate (Acqiris, Switzerland) for readout of the MCP signals.
2. Applied Biosystems 4700 proteomics MALDI TOF/TOF analyzer equipped with a 200 Hz Nd:YAG laser.
3. Physical Electronics (Eden Prairie, MN) TRIFT-II (triple focusing time-of-flight) SIMS (ToF-SIMS) equipped with an ^{115}In^{+}liquid metal ion gun.

3. Methods

IMS experiments described here are ME-SIMS and MALDI experiments. The SIMS experiments are conducted on a Physical Electronics TRIFT-II equipped with an ^{115}In^{+} liquid metal ion gun (SIMS). Stigmatic MALDI experiments are conducted on a

Physical Electronics TRIFT-II equipped with a phosphor screen/CCD camera optical detection combination and a MALDI ionization source. The microprobe MALDI experiments were conducted on an Applied Biosystems 4700 proteomics analyzer operated using the 4000 Series Explorer software. For imaging experiments, an extra software package (4000 Series Imaging) is required, which is freely available form the http://www.maldi-msi.org website.

In our stigmatic imaging approach, the 150 × 200 μm homogeneous laser pulse irradiates the sample surface. The desorbed ions pass an immersion lens/transfer lens combination followed by a high-speed blanker and are detected at a microchannel plate (MCP)/position sensitive detector. In this manner, a mass-to-charge (*m/z*) separated series of molecular images is generated, allowing simultaneous recording of a microscope and microprobe dataset in a single experiment. The position sensitive detector consists of a CCD camera phosphor screen assembly, where snapshots of the ions reaching the detector are taken. These snapshots are used to construct larger area stigmatic ion images.

Sample handling, from dissection to measurement, is of crucial importance for extraction of useful data from any tissue section. Sample preparation has to be done fast, to prevent (degenerative) alterations in the sample, clean, and with as little lateral diffusion of the analytes as possible. The time between sectioning and freezing as well as between defrosting and measurement is a key factor since observations by Svensson et al. *(23)* suggested postmortem changes of susceptible peptides and proteins within minutes. These changes can be minimized by snap freezing of the dissected tissue and defrosting only just before sample preparation starts.

For matrix deposition different methods are available, depending on the spatial resolution required. For high-spatial resolution IMS, ESD or pressure driven spray is used. Key issues for the matrix deposition method are optimal incorporation of analyte into matrix crystals with a defined size and with minimal lateral diffusion. These requirements can be met if the matrix arrives at the tissue surface in very small droplets before all solvent has evaporated.

3.1. Preparation of the Samples

1. Small organs are embedded in 10% gelatin at 30°C directly after dissection and frozen at −80°C. Never embed the tissue in polymer containing cryopreservative solution like Tissue-Tek O.C.T. Compound (Sakura Finetek USA).
2. The dissected organs are placed on the sample holder in the cryostat and secured by a very small amount of Tissue-Tek (*see* **Note 2**).
3. Cut the tissue sections 10-μm thick using the cryomicrotome at −20°C. Ten micrometers of thick tissue sections provide

enough material for meaningful molecular analysis without giving problems concerning insulation.

4. Pick up the tissue sections with either (ITO) glass slides or manufacturer sample plates by thaw mounting (i.e., the tissue sticks to the glass or steel because it thaws slightly upon touching) and place them in a closed container on dry ice and store at −80°C until use.
5. Before use, bring the tissue sections to room temperature in a desiccator (*see* **Note 3**).
6. When the tissue sections are at room temperature, there is an optional washing step. Washing of the tissue sections is performed to remove the excess of (saline) salt. The presence of a high concentration of salt hampers the matrix crystallization and analyte incorporation. Furthermore, segregation of the salt molecules from the matrix crystals results in an inhomogeneous sample surface where local variations influence the ionization process. The washing step is optional since there are many types of analysis where one is interested in small molecules (e.g., hormones, lipids, drugs), which can be washed out of the tissue section very easily. For macromolecular analysis, good results are obtained by washing twice in ice-cold 70% ethanol. Washing is typically comprised of following steps (a) Place the glass slide containing the tissue section in washing solution number one by immersing very gently and do not stir or shake; (b) Leave the slide for 1 min in the first washing solution and than take it out very gently and remove any big drops from the slide; (c) Place the sample, again as gently as possible, in the second, fresh, washing solution for 1 min; (d) After the second washing step take out the slide, keep it in a horizontal position and gently blow a stream of nitrogen over the surface to assist in drying; and (e) Place the tissue in the desiccator and let it dry completely, else the tissue may detach from the glass slide and start to ripple (*see* **Note 4**).

3.2. Matrix Deposition

1. ESD: Pump matrix solution, 15 mg/mL DHB or 10 mg/mL HCCA (*see* **Note 5**), from a gastight syringe through a stainless steel electrospray capillary maintained at 3.7 kV, for 10 min at a flow rate of 12 μL/h. Needle to sample plate distance is 5.0 mm. No drying or nebulisation gas is used. In the control software the *x*,*y* movement and speed of the sample plate is entered as well as the number of spray cycles. After the last spray cycle, check the matrix coverage using an optical microscope. When the matrix coverage is sufficient, let the tissue sections dry for ~30 min.
2. SprayStation: Matrix solution, ~40 mg/mL DHB (*see* **Note 6**) or 10 mg/mL HCCA, is placed in the pressure resistant container. Before depositing matrix on the tissue, check the

size and the quality of the spray; if needed adjust the pressure on the return spring by adjusting the "flow control knob." Spraying on a clean microscope glass slide should result in a fine mist of micron-sized droplets. In the software, enter the number of spray cycles, the time between cycles (needed to dry the sample), and the number of cycles before a new washing step. After the last spray cycle, check the matrix coverage using an optical microscope and, when sufficient, let the tissue sections dry for ~30 min.

3. Use a TLC sprayer to spray matrix solution, 10 mg/mL HCCA or ~40 mg/mL DHB, at a nitrogen pressure of 0.3–0.4 bar (*see* **Note 6**). Use several spray cycles to achieve homogenous matrix coverage and allow the tissue to dry in-between the spray cycles in a horizontal position. Check matrix coverage using an optical microscope and, when sufficient, let the tissue sections dry for ~30 min.

3.3. Additional Preparations for MALDI-IMS

1. Calibration solution: three standard peptides in applicable mass range are dissolved in the matrix solution and spotted on the target (*see* **Note 7**).
2. For MALDI, coat the matrix covered tissue sections with ~4 nm of gold. (a) Place the sample in the Quorum Technologies SC7640 sputter coater. Make sure a gold target is installed. (b) Press the start sequence button to pump down, flush with argon, further pump down the vacuum chamber and leak in argon until the pressure reaches 0.1 mbar (all done automatically). (c) Enter the density of the metal (19.30 g/cm^3 for gold) and the desired thickness of the sputtered metal layer. (d) Put the discharge voltage on 1 kV and press start. (e) Adjust the plasma current to 25 mA for homogenous coverage.

3.4. ME-SIMS

1. Before conducting SIMS imaging experiments, optimize the setup for image quality using, for example, a copper grid with a 25 μm repeat.
2. Make sure the imaging experiment is conducted in unbunched mode for optimal image quality.
3. Save the acquired data as raw file, in order to be able to post process the data after the measurement is completed.
4. Perform the ME-SIMS experiment in such a way that the analysis is conducted in the static SIMS regime (*see* **Note 8**). In our case, this is accomplished with a primary ion beam current of ~450 pA, a primary pulse length of 30 ns, a spot diameter of 500 nm, and a primary ion energy of 15 kV. At v3 min per experiment, this results in a primary ion dose of 4.9×10^{11} ions/cm^2.
5. For each chemical image, raster the primary beam over a 150 × 150 μm sample area, divided into 256 × 256 square pixels

(larger or smaller areas can also be chosen). To image a significant larger surface, like a whole tissue section, analyze multiple 150 × 150 μm areas by stepping the sample stage in a mosaic pattern. To compensate for small deviations on the sample stage positioning take a 10 μm overlap with the previous acquired sample (the sample stage is moved by 140 μm).

6. After the analysis the acquired spectra can be visualized in the WinCadence software and for selected *m/z* ranges the distributions can be depicted. For large tissue sections, multiple imaging experiments are performed to cover the entire area. Viewing the tissue section in one image, made up out of the multiple experiments, can be accomplished by combining (stitching) the individual images. Two approaches of image stitching are available. First, 150 μm^2 images are stitched together manually in image handling software like Adobe Photoshop. Second, PCA-based methods are used to generate feature-based registration of the imaging data-cubes for the visualization of the imaging data.

3.5. Microprobe MALDI-IMS

1. Determine the tissue section boundaries and take a test spectrum outside the tissue section to determine the acquisition method. Here, the maximum number of data points per spectrum must always be lower than 32,767. Increase the bin size or reduce the mass range to meet this criterion.
2. Select “Manual Acquisition” with one spectrum, define up to 255 laser shots (100 laser shots in our case) and load the 4700 Imaging Tool (freely available from the http://www.maldi-msi.org website).
3. Set the coordinates of the tissue section boundaries and give the raster size, the numbers of pixels on the *XY* scale are calculated using the “Dimensions” button.
4. Give a file name and start the acquisition.
5. Data analysis is performed using BioMap 3.7.4 software, which is freely available from the http://www.maldi-msi.org website. Our experiments were performed on an Applied Biosystems 4700 proteomics analyzer, which creates a, BioMap compatible, *.img file. However, datasets obtained with the Bruker Reflex IV or the Ultraflex II are also compatible with the BioMap software after an additional conversion routine (available free of charge, details are described by Clerens et al. *(24)*). When analyzing the IMS dataset with BioMap, the entire dataset is loaded into the computers’ virtual memory and thus sufficient memory is needed (for medium-sized datasets 2 GB). In BioMap, multiple visualization and data handling routines are available. One important note to make is that the data is automatically shown as interpolated

(smoothed) images. For the real (pixelated) image one has to choose “voxel” instead of “interpolate” in the “display methods” of the image properties window (right mouse click on image).

3.6. Stigmatic MALDI-IMS

1. Before conducting stigmatic MALDI imaging experiments, optimize the setup for image quality using a fine mesh TEM grid (19 µm squares, 25 µm pitch) overlain on peptide-doped matrix crystals. To do so the sample potential and ion optical lens (e.g., immersion lens and transfer lens) voltages are optimized such that a focused image of the grid structure on the phosphor screen is obtained.
2. Set in mosaic imaging parameters the dimensions of the number of pixels to be imaged and the spacing between two pixels. Important factors in setting these parameters are the amount of internal memory needed on the Acqiris card when the acquired data is written to disk after the analysis or the speed of data transfer when writing on the fly.
3. Set the stigmatic image collection parameters depending on the chosen values for the mosaic images. The amount of images acquired is specified in dependence on the number of pixels and number of shots per pixel. Furthermore, the exposure time on the CCD chip of the camera can be set to average over one laser shot or multiple laser shots (per pixel).
4. Two datasets are recorded simultaneously, but data is processed separately and combined afterward. The stigmatic ion images can be stitched together showing either the TIC image of the entire tissue (**Fig. 3a**) section or a high-resolution selected ion distribution. The MALDI mass spectral data is acquired using an Acqiris digitizer resulting in a single *.data file for every laser shot. Acquired data can be analyzed as single shot spectra, as combined spectra per line scan or as combined spectra of the whole measurement. Furthermore, this data can be used to construct low-resolution microprobe selected-ion-count (SIC) images (**Fig. 3b**). Localizations of SICs from the entire microprobe dataset can be visualized using in-house developed software (Data-cube-viewer). Here one can scroll trough the entire mass spectral dataset showing the two-dimensional localization at every m/z value with the intensity data in the z-direction.

3.7. PCA

1. For both the ME-SIMS and MALDI-IMS data, PCA routines *(25)* can be used to find a preselected number of “principal components” (PCs), describing some percentage in variation (the variance) of the original spectral data. In this manner, different correlated molecular components can be

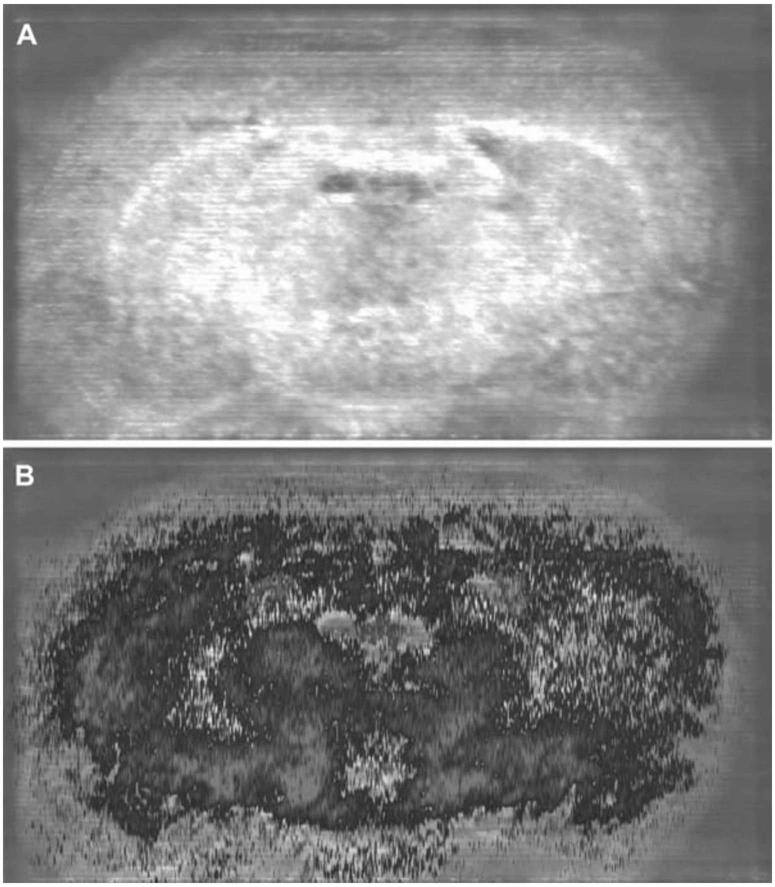


Fig. 3. MALDI-IMS images of a rat brain tissue section with (**a**) stigmatic TIC image and (**b**) same image as in (**a**) with microprobe data of the molecular species at *m/z*= 1,804 (*red*) and *m/z*= 3,478 (*green*) overlaid. (*See Color Plates*)

imaged together, rather than generating images of each individual peak in the dataset.

2. The data is read using MatLab (version 7.0.4, R14, SP2). The maximum size of the dataset is dependent on the available memory. Since the IMS datasets consist of relatively large areas with zero counts, the data is stored in a Harwell-Boeing format (which omits the storage of zero counts). To further reduce the data size, binning is performed in the spectral domain.

4. Notes

1. Recent observations by Svensson et al. *(23)* point toward postmortem changes in the proteome of susceptible peptides and proteins within minutes. To prevent these alterations of the sample, Svensson et al. *(23)* used focused microwave irradiation

to sacrifice the animals. Since focused microwave irradiation is not available in every laboratory, an alternative way to minimize alterations is snap freezing of the dissected tissue and defrosting only just before the sample preparation starts.

2. Be very careful not to make any contact between the Tissue-Tek and parts of the tissue used for sectioning. The Tissue-Tek contains a large amount of polymer substance, which smears over the tissue surface upon cutting. These polymers will dominate the resulting mass spectra.
3. After defrosting, the degenerative processes in the sample continue by the reactivation of multiple endogenous proteases. For this reason the defrosting should be done just before sample preparation starts and tissues should not be left untreated for a sustained period. Furthermore, sample preparation should be done delicately but fast to prevent further degeneration and allow intact molecular species to get incorporated into the MALDI matrix.
4. In any tissue imaging experiment, diffusion of analyte molecules has to be prevented. For this reason, the washing step has to be performed carefully without any steering or shaking of the solution.
5. The ESD needle can get clocked when too high matrix concentrations are used. Adjust the concentration of the matrix solution if needed.
6. Increasing the concentration of DHB (which can be more than 100 mg/mL) results in increased need to wash the nozzle. At high concentrations of DHB, the nozzle becomes partly clogged, resulting in an irregular or "spitting" matrix spray, which in turn results in large droplets and large matrix crystals.
7. Best calibration results are obtained with the calibration components spotted on the same underground as in the analysis (i.e., a tissue surface).
8. In static SIMS, the same area in never sampled twice in order to prevent imaging of induced damage. To do so, only 1% of the sample surface is analyzed, which converts to a primary ion dose of 10^{13} ions/cm^2.

Acknowledgments

We thank Frans Giskes for the help in developing the automated pressure driven spray station. This work was carried out in the context of the Virtual Laboratory for e-Science project (http://

www.vl-e.nl). This project is supported by a BSIK grant from the Dutch Ministry of Education, Culture and Science (OC & W) and is part of the ICT innovation program of the Ministry of Economic Affairs (EZ). The Netherlands Proteomics Centre financially supported this project. This work is part of research program nr. 49 “Mass spectrometric imaging and structural analysis of biomacromolecules” of the “Stichting voor Fundamenteel Onderzoek der Materie (FOM)”, which is financially supported by the “Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO).”

References

1. Pandey, A., and Mann, M. (2000) Proteomics to study genes and genomes. *Nature* 405, 837–46.
2. Hanash, S. (2003) Disease proteomics. *Nature* 422, 226–32.
3. Figeys, D. (2002) Proteomics approaches in drug discovery. *Anal. Chem.* 74, 413a–19a.
4. Stoeckli, M., Chaurand, P., Hallahan, D. E., and Caprioli, R. M. (2001) Imaging mass spectrometry: A new technology for the analysis of protein expression in mammalian tissues. *Nat. Med.* 7, 493–6.
5. Schwartz, S. A., Reyzer, M. L., and Caprioli, R. M. (2003) Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: Practical aspects of sample preparation. *J. Mass Spectrom.* 38, 699–708.
6. Altelaar, A. F. M., Luxembourg, S. L., McDonnell, L. A., Piersma, S. R., and Heeren, R. M. A. (2007) Imaging mass spectrometry at cellular length scales. *Nat. Protoc.* 2, 1185–96.
7. Vickerman, J. C., and Briggs, D. (2001) ToF-SIMS: Surface Analysis by Mass Spectrometry. IM Publications and SurfaceSpectra, Chichester.
8. Wu, K. J., and Odom, R. W. (1996) Matrix-enhanced secondary ion mass spectrometry: A method for molecular analysis of solid surfaces. *Anal. Chem.* 68, 573–882.
9. Altelaar, A. F. M., van Minnen, J., Jimenez, C. R., Heeren, R. M. A., and Piersma, S. R. (2005) Direct molecular imaging of lymnaea stagnalis nervous tissue at subcellular spatial resolution by mass spectrometry. *Anal. Chem.* 77, 735–41.
10. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988) Protein and polymer analyses up to *m/z* 100000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 2, 151–3.
11. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10000 Daltons. *Anal. Chem.* 60, 2299–301.
12. Strick, R., Strissel, P. L., Gavrilov, K., and Levi-Setti, R. (2001) Cation-chromatin binding as shown by ion microscopy is essential for the structural integrity of chromosomes. *J. Cell Biol.* 155, 899–910.
13. Peteranderl, R., and Lechene, C. (2004) Measure of carbon and nitrogen stable isotope ratios in cultured cells. *J. Am. Soc. Mass Spectrom.* 15, 478–85.
14. Chandra, S., Smith, D. R., and Morrison, G. H. (2000) Subcellular imaging by dynamic SIMS ion microscopy. *Anal. Chem.* 72, 104A–14A.
15. Todd, P. J., McMahon, J. M., Short, R. T., and McCandlish, C. A. (1997) Organic SIMS of biologic tissue. *Anal. Chem.* 69, A529–35.
16. Rohner, T. C., Staab, D., and Stoeckli, M. (2005) MALDI mass spectrometric imaging of biological tissue sections. *Mech. Ageing Dev.* 126, 177–85.
17. Altelaar, A. F. M., Taban, I. M., McDonnell, L. A., Verhaert, P. D. E. M., de Lange, R. P. J., Adan, R. A. H., Mooi, W. J., Heeren, R. M. A., and Piersma, S. R. (2007) High-resolution MALDI imaging mass spectrometry allows localization of peptide distributions at cellular length scales in pituitary tissue sections. *Int. J. Mass Spectrom.* 260, 203–11.
19. Skold, K., Svensson, M., Nilsson, A., Zhang, X., Nydahl, K., Caprioli, R. M., Svenningsson, P., and Andren, P. E. (2006) Decreased striatal levels of PEP-19 following MPTP lesion in the mouse. *J. Proteome Res.* 5, 262–9.

20. Chaurand, P., Schwartz, S. A., and Caprioli, R. M. (2004) Profiling and imaging proteins in tissue sections by MS. *Anal. Chem.* 76, 86a–93a.
21. Meistermann, H., Norris, J. L., Aerni, H. -R., Cornett, D. S., Friedlein, A., Erskine, A. R., Augustin, A., De Vera Mudry, M. C., Ruepp, S., Suter, L., Langen, H., Caprioli, R. M., and Ducret, A. (2006) Biomarker discovery by imaging mass spectrometry – Transthyretin is a biomarker for gentamicin-induced nephrotoxicity in rat. *Mol. Cell Proteomics* 5, 1876–86.
22. Luxembourg, S. L., Mize, T. H., McDonnell, L. A., and Heeren, R. M. A. (2004) The Mass Microcope: high-speed, (sub-) micron imaging of peptide and protein distributions. *Anal. Chem.* 76, 5339–44.
23. Svensson, M., Skold, K., Svenningsson, P., and Andren, P. E. (2003) Peptidomics-based discovery of novel neuropeptides. *J. Proteome Res.* 2, 213–9.
24. Clerens, S., Ceuppens, R., and Arckens, L. (2006) CreateTarget and Analyze This! new software assisting imaging mass spectrometry on Bruker Reflex IV and Ultraflex II instruments. *Rapid Commun. Mass Spectrom.* 20, 3061–66.
25. Broersen, A., van Liere, R., and Heeren, R. M. A. (2005) Comparing three PCA-based Methods for the 3D Visualization of Imaging Spectroscopy Data, in *IASTED International Conference on Visualization, Imaging, & Image Processing*, ACTA Press, Benidorm, Spain, pp. 540–45.

Chapter 19

Proteomic Global Profiling for Cancer Biomarker Discovery

Vitor Faca, Hong Wang, and Samir Hanash

Summary

The ultimate goal of cancer molecular diagnostics is the development of simple tests to predict cancer risk, detect cancer early, classify tumors, and monitor response to therapy. Proteomics is well suited for these tasks. However, there are substantial challenges that need to be met to identify the most informative markers using proteomics. Approaches for in-depth quantitative proteomic analysis based on isotopic labeling and protein fractionation are presented in this chapter.

Key words: Biomarkers, Proteomics, LC-MS/MS, Intact proteins, Acrylamide, Isotopic labeling, Serum, Plasma, Multidimensional liquid chromatography.

1. Introduction

The dynamic nature of the proteome of a cell or a tissue, and the fact that many proteome alterations in disease are not predictable from genomic analysis, provides ample justification for studying gene expression in disease directly at the proteomic level *(1)*. The human plasma proteome is one of the most complex proteomes, containing a vast variety of proteins such as "true" plasma proteins (for example proteins in the coagulation and complement systems), proteins that come from leakage of any irrigated tissue, and foreign proteins coming from infectious organisms. These proteins are present in plasma in concentrations ranging from milligrams to picograms or less per milliliter *(2)*. To add to this complexity, plasma proteins often occur in multiple isoforms, cleaved forms, or splice variants *(2, 3)*. Although these characteristics make the plasma proteome one of the most challenging

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_19

to study, it is also expected to be a readily available source of circulating biomarkers that can lead to new disease diagnosis and therapeutics.

Mass spectrometry has evolved enough to detect and identify femtomols of peptides, but the dynamic range of detection, usually between three and five orders of magnitude, is still limiting *(4)*, covering less than half of the range of plasma proteins' abundance. In order to look deeper into the plasma proteome and achieve higher sensitivity, two basic strategies have been used: removal of high abundant proteins, such as albumin and immunoglobulins that interfere with the detection of lower abundance ones, and fractionation of the plasma sample into smaller subsets in order to reduce sample complexity *(3, 5–8)*.

Another challenge is to analyze the plasma proteome quantitatively. Numerous methods have been introduced for quantitative analysis of proteins by mass spectrometry, but some of them cannot be applied to human plasma samples such as in vivo labeling using cell culture media enriched with ^{15}N *(9)* or stable isotopes of amino acids *(10)*. Alternatively, labeling could be performed with various reagents during or after enzymatic digestion *(11, 12)* although samples to be compared have to be processed separately until the labeling step, which may introduce artifactual variations.

In this chapter, we describe the multidimensional protein fractionation platform and the isotope labeling approach we have been applying to profile plasma from several different types of cancer patients for disease-related changes. The approach consists of immunodepletion of six of the most abundant proteins, acrylamide isotope labeling *(13)* followed by an orthogonal two-dimensional intact-protein fractionation procedure. Fractionated proteins are then digested and analyzed by shotgun LC-MS/MS. With this approach, sample complexity was greatly reduced prior to mass spectrometric analysis, permitting the identification and to obtain relative quantitation information for proteins and protein isoforms across seven orders of magnitude in protein concentration.

2. Materials

2.1. Sample

The sample pair (disease and control) to be compared can consist of serum or plasma from two individuals or a pool of two groups of disease and control plasma or serum samples. The protocol described here is designed for a total of 600 μL each of disease and control sample.

2.2. Immunodepletion

1. HU-6 columns (4.6 × 100 mm; Agilent, Wilmington, DE).
2. Buffer A and Buffer B (Agilent – Reagent kit).
3. 0.22-µm syringe filter.

2.3. Acrylamide Isotope Labeling

1. Centricon YM-3 devices (Millipore).
2. Labeling buffer: 8 M urea, 100 mM Tris-HCl pH 8.5, 0.5% OG (octyl-beta-d-glucopyranoside).
3. Dithiotreitol (DTT) (ultrapure, USB).
4. Light acrylamide: acrylamide (>99.5% purity, Fluka).
5. Heavy acrylamide: 1,2,3-$^{13}C_3$-acrylamide (>98% purity, Cambridge Isotope Laboratories, Andover, MA).

2.4. Protein Fractionation

2.4.1. Anion-Exchange Chromatography

1. Mono-Q 10/100 (Amersham Biosciences).
2. Solvent A: 20 mM Tris-HCl in 6% isopropanol, 4 M Urea – pH 8.5.
3. Solvent B: 20 mM Tris-HCl in 6% isopropanol, 4 M Urea, 1 M NaCl – pH 8.5.

2.4.2. Reversed-Phase Chromatography

1. Poros R2 column (4.6 × 50 mm, Applied Biosystems).
2. Solvent A: 95% H_2O, 5% Acetonitrile + 0.1% TFA.
3. Solvent B: 90% Acetonitrile, 10% H2O, + 0.1% TFA at 2.7 mL/min.

2.5. Mass Spectrometry Protein Identification

2.5.1. Protein Digestion

1. Modified trypsin (Promega).
2. Digestion buffer: 0.25 M urea, 50 mM ammonium bicarbonate and 4% acetonitrile (v/v).

2.5.2. Mass Spectrometry Analysis

1. LTQ-FT mass spectrometer (Thermo-Finnigan).
2. nanoACQUITY nanoflow chromatographic system (Waters).
3. Chromatographic column: 25 cm Picofrit 75 µm ID (New Objectives) in house-packed with MagicC18 (100 Å pore size/5 µm particle size).
4. Trap column: nanoACQUITY UPLC™ Symmetry C_{18} 180 µm × 20 mm, (particle size 5 µm).
5. Solvent A: 0.1% Formic Acid in HPLC grade water (v/v).
6. Solvent B: 0.1% Formic acid in chromatography grade acetonitrile (v/v).

2.5.3. Data Analysis

1. All the data management was performed using the Computational Proteomics Analysis System *(14)*.
2. Search Algorithm: X!Tandem *(15)*.
3. Scoring: PeptideProphet *(16)* and ProteinProphet *(17)*.
4. Quantitation: Q3 *(13)*.

3. Methods

The intact protein fractionation is performed based on the Intact Protein Analysis System (IPAS) approach *(3, 5)*, with some modifications (*see* **Note 1**).

3.1. Immunodepletion (see Note 2)

1. Filter plasma or serum samples in a 0.22-µm syringe filter.
2. Equilibrate the HU-6 column with buffer A at 0.5 mL/min for 13 min.
3. Load aliquots of 75 µL of the sample in each depletion run.
4. Collect the flow-through fractions for 10 min and immediately store aliquots at –80°C until use.
5. Elute the bound material and regenerate the column with buffer B at 1 mL/min for 8 min (collect if wanted).
6. Repeat **steps 2–5** for the whole 600-µL aliquot of the disease and control samples to be compared.

3.2. Isotope Labeling (see Note 3)

1. Concentrate separately disease and control immunodepleted serum aliquots in an Amicon YM-3 device until volume <100 µL.
2. Dilute samples up to 1 mL with 8 M urea, 100 mM Tris pH 8.5, 0.5% OG (octyl-beta-d-glucopyranoside) (w/v).
3. Measure protein concentration for disease and control samples using Bradford method (*see* **Note 4**).
4. Reduce protein disulfide bonds in each sample for 2 h at room temperature by adding 0.66 mg dithiotreitol (DTT) per mg of protein.
5. Alkylate the samples with acrylamide isotopes immediately after the reduction by adding 7.1 mg/mg of total protein of light acrylamide to one of the samples (control or disease) and 7.4 mg/mg of total protein of heavy acrylamide to the other sample. The reaction is carried out for 1 h at room temperature.
6. Mix disease and control samples, dilute it to 10 mL with solvent A of anion-exchange chromatography and submit it immediately to the chromatography.

3.3. Protein Fractionation

3.3.1. Anion Exchange Chromatography (First Dimension)

Gradient: Flow rate: 4.0 mL/min
0 min: 0% solv. B
44 min: 35% solv. B
47 min: 50% solv. B
52 min: 100% solv. B
57 min: 100% solv. B

1. Equilibrate the column for at least 10 min with solvent A before injections.
2. Collect fractions in 15 mL tubes at a 1 frac/min rate.
3. Combine collected fractions in 12 pools for second-dimension fractionation (*see* **Note 5**).

3.3.2. Reversed-Phase Chromatography (Second Dimension)

Gradient: Flow rate: 2.7/min
0 min: 5% solv. B
18 min: 50% solv. B
25 min: 80% solv. B
27 min: 95% solv. B
32 min: 95% solv. B

1. Equilibrate the column for at least 10 min with solvent A before injections.
2. Inject the sample and keep the system at the initial solvent concentration of 5% solv. B until absorbance returns to baseline (*see* **Note 6**) before starting the gradient.
3. Collect fractions of 900 μL during the run.
4. Separate aliquots of 400 μL of each fraction tryptic digestion and for mass-spectrometry shotgun analysis (*see* **Note 7**).

3.4. Mass Spectrometry Protein Identification

3.4.1. Protein Digestion (see Note 8)

1. Perform the *in-solution* digestion with individual lyophilized aliquots from the reversed phase.
2. Resuspend individual fractions in 0.25 M urea containing 50 mM ammonium bicarbonate and 4% of acetonitrile and digest overnight with 200 ng of modified trypsin (Promega).
3. Interrupt the digestion by adding of 5 μL of 10% formic acid solution. Individual fractions 4–51 were pooled in a total of 16 pools.

3.4.2. Mass Spectrometry Analysis (see Note 9)

Gradient: Flow rate: 300 nL/min
0 min: 5% solv. B
2 min: 7% solv. B
92 min: 35% solv. B
93 min: 50% solv. B
102 min: 50% solv. B
103 min: 95% solv. B
108 min: 95% solv. B
109 min: 5% solv. B
130 min: 5% solv. B (column equilibration)

1. Inject 5 μL of each pool onto the column.
2. After injection, sample is cleaned up in the trap column for 5 min at 4 μL/min of 5% solv. B.

3. Acquire spectra in a data-dependent mode in *m/z* range of 400–1,800. Select the five most abundant +2 or +3 ions of each MS spectrum for MS/MS analysis. Mass spectrometer parameters are capillary voltage of 2.1 kV, capillary temperature of 200°C, resolution of 100,000, and FT target value of 1,000,000.

3.4.3. Data Analysis

The acquired data is automatically processed by the Computational Proteomics Analysis System – CPAS *(14)* (*see* **Note 10**).

1. For databank searches, light acrylamide alkylation is considered a fixed modification with mass of 71.03657 Da and heavy acrylamide labeled peptides are detected with a variable modification using a delta mass of 3.01006 Da (*see* **Note 11**).
2. Use the minimum criteria for peptide matching of Peptide Prophet score greater than 0.2 *(16)*. Peptides that met these criteria are further grouped to protein sequences using the Protein Prophet algorithm at an error rate of 5% or less (*see* **Note 12**).
3. Acrylamide ratios are obtained by a script called "Q3" that was developed in-house to obtain the relative quantification for each pair of peptides identified by MS/MS that contains cysteine residues (*see* **Note 13**).

4. Notes

1. The original IPAS platform relied on Cy Dye fluorofores and gel-based analysis to obtain quantitative information *(3, 5)*. As with all gel-based methods, the approach was cumbersome and presented low throughput. The gel-based analysis was replaced with shotgun LC-MS/MS analysis of fractionated proteins to provide a comprehensive profiling of plasma samples.
2. This specific column has a maximum load capacity of 80 μL, requiring several injections of the sample. There are a variety of immunodepletion columns now on the market that can be used, including higher capacity columns, which reduce the need of multiple injections. These columns can usually be used for 100–200 cycles of injection/regeneration.
3. The feasibility of acrylamide labeling for protein identification and quantitation has been demonstrated previously using MALDI-TOF for proteins separated by two-dimensional electrophoresis *(18)*. No protein precipitation was observed, indicating that acrylamide-based alkylation is compatible with intact-protein-based approaches. A very high yield of cysteine alkylation is achieved *(13)*. The interruption of reaction after

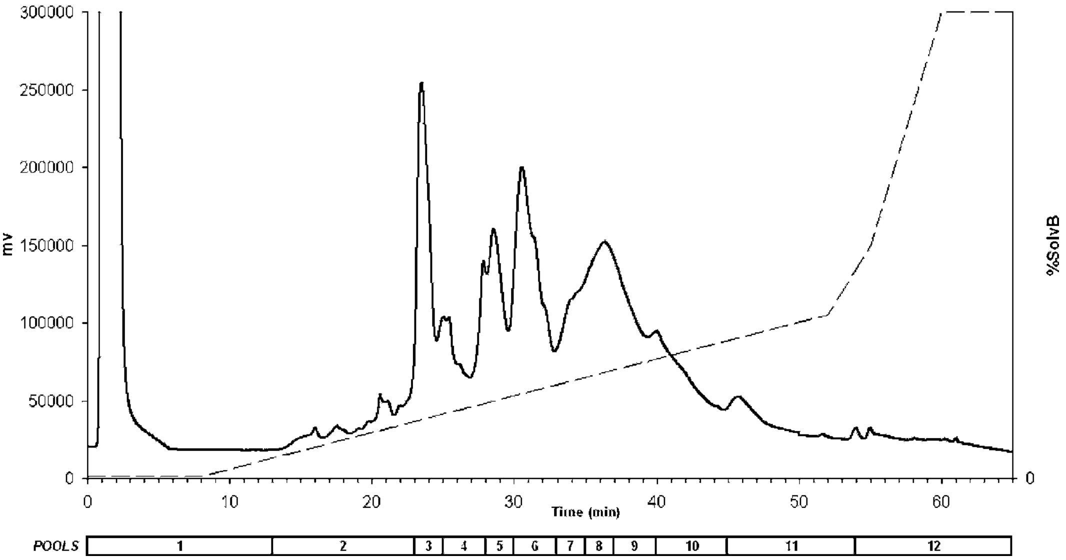


Fig. 1. First-dimension protein fractionation of human serum. The anion-exchange chromatogram corresponds to the injection of approximately 10 mg of immunodepleted human serum proteins. The gradient is represented by the dashed line. Absorbance monitored at 280 nm. The boxes below the chromatogram represent the pooling of 12 fractions that are subsequently fractionated by reversed-phase chromatography.

1 h is very important to avoid side reactions with free amines in the protein.

4. As a result of the immunodepletion step, the total protein concentration is reduced by 85%, from approximately 70 to 10 mg/mL.
5. A representative chromatogram for the first dimension is illustrated in Fig. 1. Pools for subsequent fractionation are formed based on the characteristic features of the chromatogram.
6. Due to the differences in volume or salt concentration for the pools collected from the first dimension (anion-exchange) fractionation step, the injection onto the reversed-phase chromatography was also used as a desalting/concentration step. The separation gradient is started after all the flow-through is washed out of the column.
7. A representative chromatogram of one of the second dimension fractions is illustrated in Fig. 2. All the 864 fractions collected from the 12 reversed-phase chromatographic runs (72 fractions/run) are split into two aliquots (400 and 500 μL). Four hundred microliters of aliquots are used for the LC-MS/MS protein identification step while the 500-μL aliquots are saved for further validation or follow-up experiments.
8. The 400-μL aliquots from the fractionation are lyophilized and resuspended in digestion buffer for in-solution trypsin

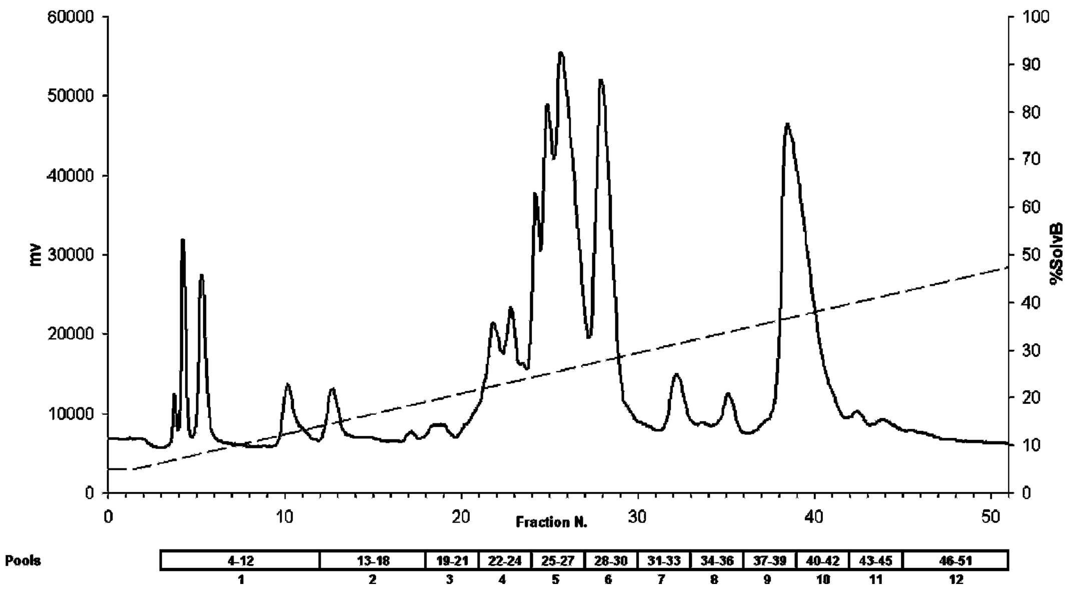


Fig. 2. Second-dimension fractionation of human serum. The reversed-phase chromatogram corresponds to the separation of pool 4 (Fig. 1) from first dimension. The gradient is represented by the dashed line. Absorbance monitored at 280 nm. Boxes represent the pools formed for subsequent LC-MS/MS analysis.

digestion. All the 864 aliquots are digested individually. After digestion, aliquots are combined into 12 pools for each reversed-phase run, performing a total of 144 samples to be analyzed by shotgun LC-MS/MS. The number of samples submitted to the LC-MS/MS analysis (number of pools formed) is chosen based on the time × benefit balance (protein identification).

9. One representative chromatogram for the LC-MS/MS analysis is presented in **Fig. 3**. Wash runs are performed every four samples. A whole experiment takes 18 days to be run in this LC-MS/MS gradient. On average, 100 proteins are identified in each run.

10. CPAS, the Computational Proteomics Analysis System, contains an entire data analysis and management pipeline for Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) proteomics, including experiment annotation, protein database searching and sequence management, and mining LC-MS/MS peptide and protein identifications. CPAS is an open source tool that can be downloaded at http://cpas.fhcrc.org.

11. The efficiency of the acrylamide labeling is virtually 100%, so all the cysteine residues are expected to be fully alkylated. Also, most of the databank search algorithms cannot handle multiple modifications in the same amino acid residue. Databank searches carried out with acrylamide labeling as a variable modification should result in very few unmodified peptides.

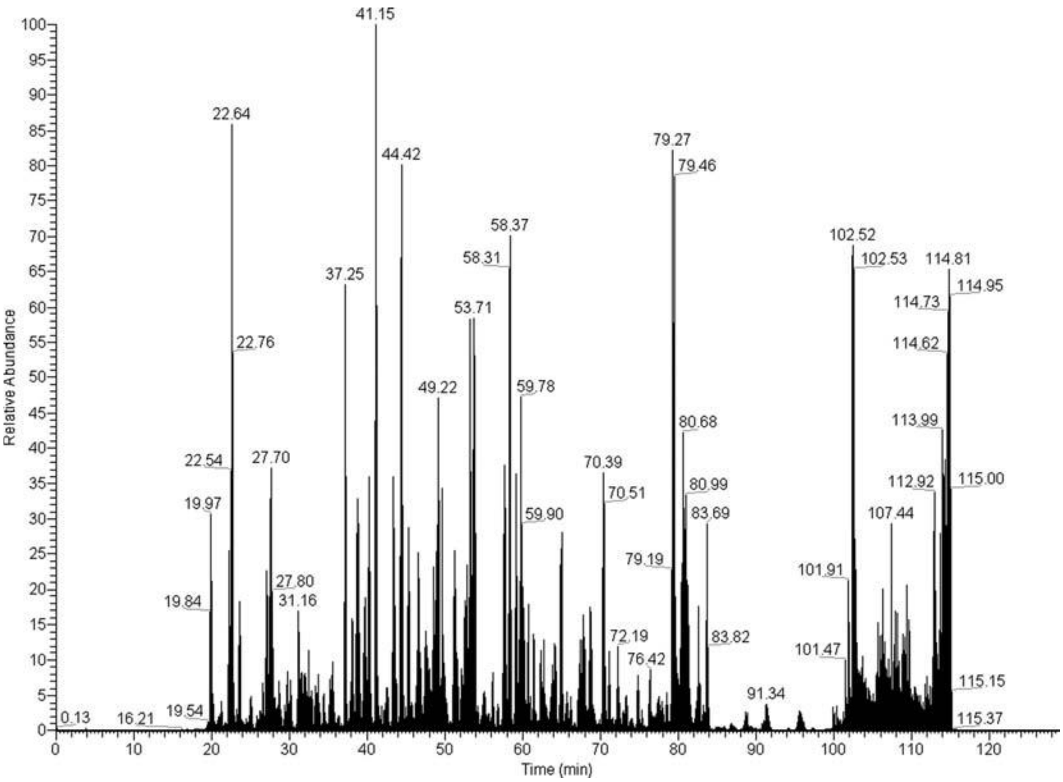


Fig. 3. LC-MS/MS peptide identification. The reconstructed ion current chromatogram corresponds to the injection of approximately 3 μg of digested protein obtained from one pool of the second dimension fractionation. This single 130-min LC-MS/MS run led to the identification of 109 unambiguous proteins with <5% FDR.

12. The criteria for good protein matching can be considered very stringent. Although protein identifications based on single peptide hits are not discarded, these peptides have to achieve a very high Peptide Prophet score to lead to significant protein identification. Around 1,200 proteins are identified in one full experiment, with most of them corresponding to medium-low abundance proteins (**Fig. 4**). In this particular type of experiment where intact proteins are fractionated, different distant fractions containing the same protein can indicate the occurence of isoforms or fragments. The mapping of the peptide identification in the protein sequence can clearly indicate this effect, as illustrated in **Fig. 5**.

13. This quantitation approach relies on the correct identification of the peptide sequence containing a cysteine residue. Briefly, from peptides identified with Peptide Prophet scores greater than 0.75, we obtained the theoretical mass of the monoisotopic light- and heavy-labeled peptides as well as the theoretical mass–charge locations of each isotopic peak. The intensity of peaks in each MS scan was computed and each

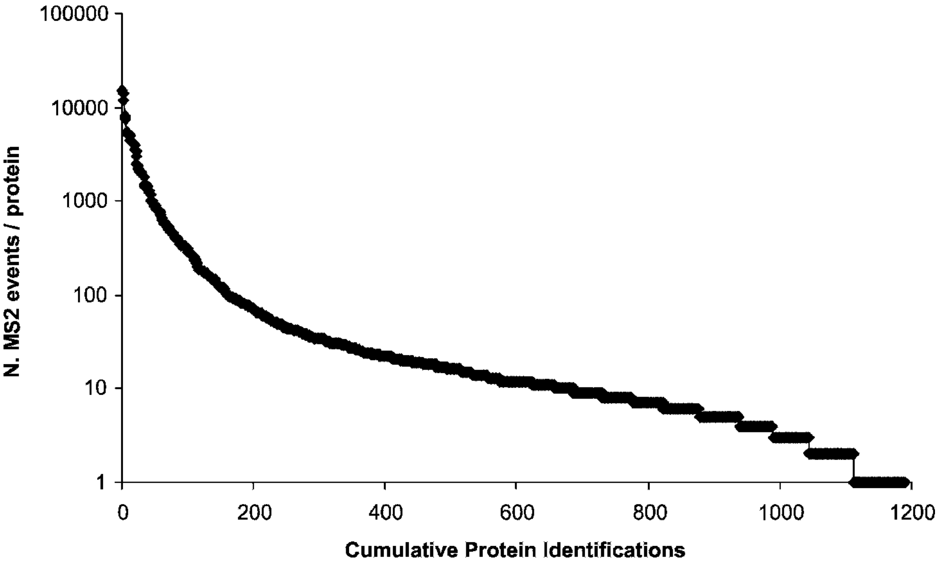


Fig. 4. Number of protein identifications. The plot illustrates that the majority (42%) of the identified protein is low abundant, with <10 MS2 events /proteins. This indicates the efficiency in fractionation to reach low abundant protein levels in serum.

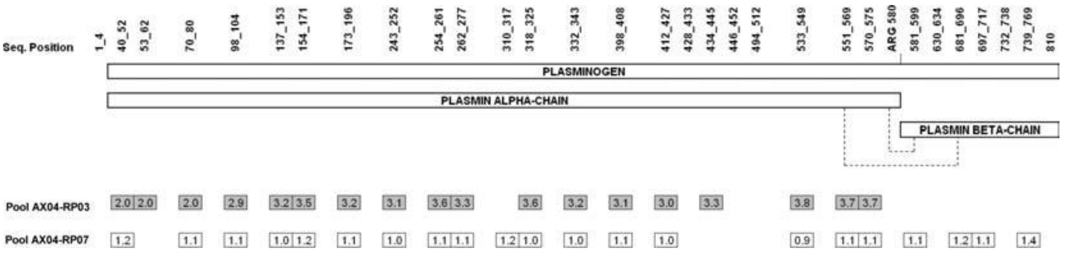

Seq. Position
1_4
40_52
53_62
70_80
98_104
137_153
154_171
173_196
243_252
254_261
262_277
310_317
318_325
332_343
398_408
412_427
428_433
434_445
446_452
494_512
533_549
551_569
570_575
ARG 580
581_599
630_634
681_696
697_717
732_738
739_769
810
PLASMINOGEN
PLASMIN ALPHA-CHAIN
PLASMIN BETA-CHAIN
Pool AX04-RP03
2.0
2.0
2.0
2.9
3.2
3.5
3.2
3.1
3.6
3.3
3.6
3.2
3.1
3.0
3.3
3.8
3.7
3.7
Pool AX04-RP07
1.2
1.1
1.1
1.0
1.2
1.1
1.0
1.1
1.1
1.2
1.0
1.0
1.1
1.0
0.9
1.1
1.1
1.1
1.2
1.1
1.4

Fig. 5. Peptide alignment of plasminogen. In pool AX04-03, the protein coverage corresponds to the alpha-chain of plasmin, while in pool AX04-07, the peptides cover the entire molecule corresponding to the precursor plasminogen. These two protein forms have approximately 4 min difference in retention time. The *numbers* indicated *inside the box* represent the acrylamide isotope ratio (disease/control) for each peptide obtained for these two different isoforms identified for plasminogen. We can observe that the plasmin alpha-chain is upregulated in the disease, while the plasminogen is not.

light peak was paired with the corresponding heavy peak. For the final quantitation, we used the primary scan and, additionally, a certain number of MS scans immediately before and after the MS/MS identification up to a limit of ten scans. For peptides with mass greater than 1,800 Da, the resulting ratio was adjusted slightly by subtracting from the derived quantity of the heavy form, the fraction that can be expected to result from overlapping isotopes of the light form. Proteins containing more than one pair of peptides that yielded quantitation had all the ratios averaged in that particular fraction.

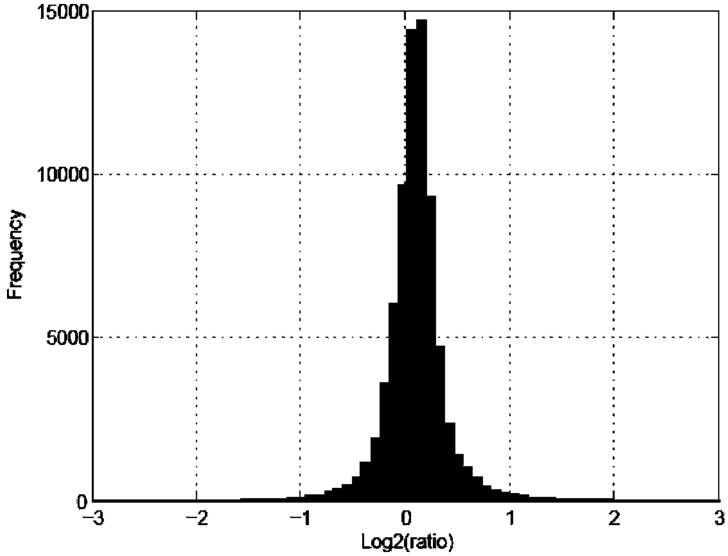


Fig. 6. Representative distribution of quantitative events. Around 80,000 MS events for cysteine containing peptides are acquired from the whole experiment, providing quantitative information for approximately 45% of all proteins identified. The histogram illustrates that for this experiment the majority of the events occurred at less than 1.5-fold ratio.

This way, isoforms eluting in different fractions, as in the particular case of the plasminogen, can be independently quantified. Figure 6 illustrates the distribution of quantitative events obtained in an entire experiment.

References

1. Hanash, S. (2003). Disease proteomics. *Nature* 422, 226–232.
2. Anderson, N. L. & Anderson, N. G. (2002). The human plasma proteome. *Mol Cell Proteomics* 1, 845–867.
3. Misek, D. E., Kuick, R., Wang, H., Deng, B., Zhao, R., Galchev, V., Tra, J., Pisano, M. R., Amunugama, R., Allen, D., Strahler, J., Andrews, P., Omenn, G. S. & Hanash, S. M. (2005). A wide range of protein isoforms in serum and plasma uncovered by a quantitative Intact Protein Analysis System (IPAS). *Proteomics* 5, 3343–3352.
4. Aebersold, R. & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198–207.
5. Wang, H. & Hanash, S. (2005). Intact-protein based sample preparation strategies for proteome analysis in combination with mass spectrometry. *Mass Spectrom Rev* 24, 413–426.
6. Wang, H., Clouthier, S. G., Galchev, V., Misek, D. E., Duffner, U., Min, C.-K., Zhao, R., Tra, J., Omenn, G. S., Ferrara, J. L. & Hanash, S. M. (2005). Intact-protein based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids. *Mol Cell Proteomics* 4, 618–625.
7. Pedersen, S. K., Harry, J. L., Sebastian, L., Baker, J., Traini, M. D., McCarthy, J. T., Manoharan, A., Wilkins, M. R., Gooley, A. A., Righetti, P. G., Packer, N. H., Williams, K. L. & Herbert, B. R. (2003). Unseen proteome: Mining below the tip of the iceberg to find low abundance and membrane proteins. *J Proteome Res* 2, 303–311.
8. Tang, H. Y., Ali-Khan, N., Echan, L. A., Levenkova, N., Rux, J. J. & Speicher, D. W. (2005). A novel four-dimensional strategy combining protein and peptide separation methods enables detection of low-abundance

proteins in human plasma and serum proteomes. *Proteomics* 5, 3329–3342.

9. Oda, Y., Huang, K., Cross, F. R., Cowburn, D. & Chait, B. T. (1999). Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci USA* 96, 6591–6596.
10. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A. & Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376–386.
11. Stewart, II, Thomson, T. & Figeys, D. (2001). 18O labeling: A tool for proteomics. *Rapid Commun Mass Spectrom* 15, 2456–2465.
12. Cagney, G. & Emili, A. (2002). De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat Biotechnol* 20, 163–170.
13. Faca, V., Coram, M., Phanstiel, D., Glukhova, V., Zhang, Q., Fitzgibbon, M., McIntosh, M., Hanash, S. (2006). Quantitative analysis of acrylamide labeled serum proteins by LC-MS/MS. *J Proteome Res* 5, 2009–2018.
14. Rauch, A., Bellew, M., Eng, J., Fitzgibbon, M., Holzman, T., Hussey, P., Igra, M., Maclean, B., Lin, C. W., Detter, A., Fang, R., Faca, V., Gafken, P., Zhang, H., Whitaker, J., States, D., Hanash, S., Paulovich, A. & McIntosh, M. W. (2006). Computational Proteomics Analysis System (CPAS): An extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiments. *J Proteome Res* 5, 112–121.
15. Maclean, B., Eng, J. K., Beavis, R. C. & McIntosh, M. (2006). General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinformatics* 22, 2830–2832.
16. Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74, 5383–5392.
17. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75, 4646–4658.
18. Sechi, S. & Chait, B. T. (1998). Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal Chem* 70, 5150–5158.

Chapter 20

Analysis of Protein Glycosylation and Phosphorylation Using Liquid Phase Separation, Protein Microarray Technology, and Mass Spectrometry

**Jia Zhao, Tasneem H. Patwa, Manoj Pal, Weilian Qiu,
and David M. Lubman**

Summary

Protein glycosylation and phosphorylation are very common posttranslational modifications. The alteration of these modifications in cancer cells is closely related to the onset and progression of cancer and other disease states. In this protocol, strategies for monitoring the changes in protein glycosylation and phosphorylation in serum or tissue cells on a global scale and specifically characterizing these alterations are included. The technique is based on lectin affinity enrichment for glycoproteins, all liquid-phase two-dimensional fractionation, protein microarray, and mass spectrometry technology. Proteins are separated based on pI in the first dimension using chromatofocusing (CF) or liquid isoelectric focusing (IEF) followed by the second-dimension separation using nonporous silica RP-HPLC. Five lectins with different binding specificities to glycan structures are used for screening glycosylation patterns in human serum through a biotin–streptavidin system. Fluorescent phosphodyes and phosphospecific antibodies are employed to detect specific phosphorylated proteins in cell lines or human tissues. The purified proteins of interest are identified by peptide sequencing. Their modifications including glycosylation and phosphorylation could be further characterized by mass-spectrometry-based approaches. These strategies can be used in biological samples for large-scale glycoproteome/phosphoproteome screening as well as for individual protein modification analysis.

Key words: Glycosylation, Phosphorylation, Posttranslational modification, Protein microarrays, Liquid phase separation, Lectin, Mass spectrometry.

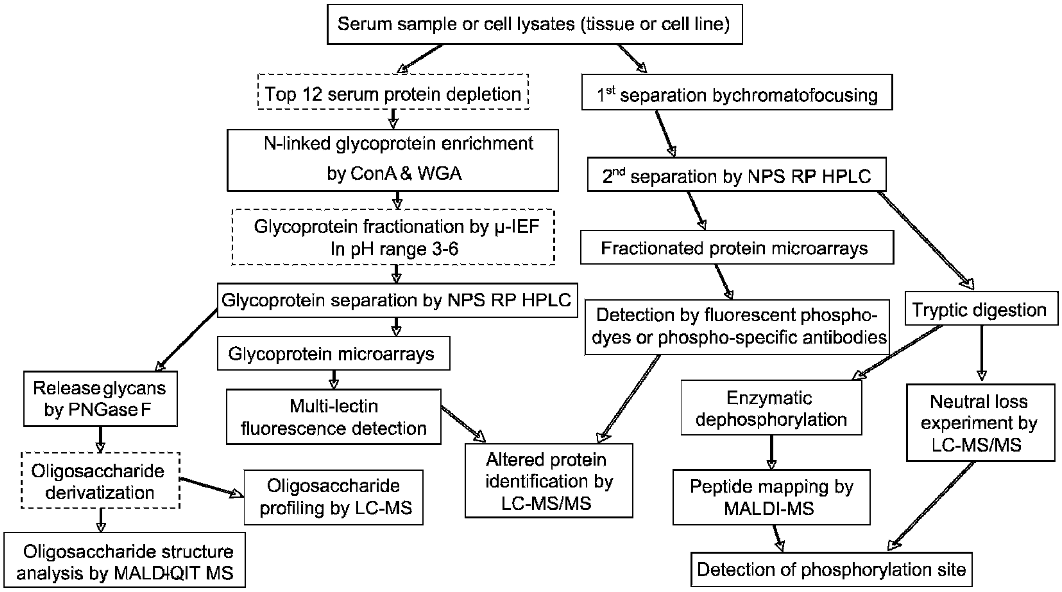
1. Introduction

Protein glycosylation and phosphorylation are very common posttranslational modifications (PTM). The alteration in protein glycosylation on the cell surface and in body fluids has been

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_20

shown to correlate with the progression of cancer and other disease states *(1–4)*. These glycosylation changes in tumor-secreted proteins reflect fundamental changes in the enzymes involved in the glycosylation pathway. The ability to efficiently profile and characterize the variation in glycoproteins could lead to the identification of disease-associated glycan alterations and new diagnostic markers. Phosphorylation and dephosphorylation of proteins regulate a variety of cellular signaling pathways typically associated with major changes in cell function *(5–7)*. It is, therefore, essential to detect the alteration in phosphorylation patterns on a global scale and to be able to identify the critical proteins involved in cell-cycle regulation where such proteins could be important in cancer onset and progression. Glycosylation and phosphorylation have been extensively studied by several technologies such as mass spectrometry and microarrays. Protein microarrays are useful as a high throughput screening method for whole-cell lysates, fractionated proteomes, tissues, and antigen–antibody reactions, whereas mass spectrometry offers the ability for in-depth analysis of the target alterations *(8, 9)*.

The strategies included in this protocol employ an all-liquid phase enrichment and fractionation methodology coupled to protein microarray technology and mass spectrometry. The procedures are briefly outlined in Scheme 1. For serum glycoprotein profiling, glycoproteins are enriched by lectin affinity chromatography. After depositing on the microarray surface, the liquid-phase separated glycoproteins are probed with different lectins with a wide range of binding specificities to different glycan moieties. The glycoprotein–lectin interaction is assessed using a biotin–streptavidin system. Human serum can be screened for different glycan structures using this strategy thereby allowing one to do comparative studies that monitor individual glycosylation changes within a glycoproteome representing different biological states *(10)*. The glycans of target glycoproteins are released by PNGase F and further mass-mapped or fragmented for structure analysis by mass spectrometry. For phosphoprotein analysis, an all-liquid two-dimensional separation method has been explored to map the phosphoprotein expression of a cell lysate to study changes in phosphorylation patterns. Proteins are separated based on pI in the first dimension using chromatofocusing (CF) or liquid isoelectric focusing (IEF) followed by a second-dimension separation based on hydrophobicity using nonporous silica RP-HPLC *(11–14)*. Fluorescent phosphodyes and phosphospecific antibodies are used to detect specific phosphorylated proteins *(15)*. The specific phosphorylation site of the target-phosphorylated protein is detected by enzymatic dephosphorylation or MS2 neutral loss experiment. This method provides


Serum sample or cell lysates (tissue or cell line)
Top 12 serum protein depletion
N-linked glycoprotein enrichment by ConA & WGA
Glycoprotein fractionation by μ-IEF In pH range 3-6
Glycoprotein separation by NPS RP HPLC
Glycoprotein microarrays
Multi-lectin fluorescence detection
Release glycans by PNGase F
Oligosaccharide derivatization
Oligosaccharide profiling by LC-MS
Oligosaccharide structure analysis by MALDI-QIT MS
1st separation bychromatofocusing
2nd separation by NPS RP HPLC
Fractionated protein microarrays
Detection by fluorescent phospho-dyes or phospho-specific antibodies
Altered protein identification by LC-MS/MS
Tryptic digestion
Enzymatic dephosphorylation
Peptide mapping by MALDI-MS
Neutral loss experiment by LC-MS/MS
Detection of phosphorylation site

Scheme 1. Flow chart of the strategies introduced in this protocol for the protein glycosylation and phosphorylation screening. The *dashed line* frame is an optional step in the procedure.

a convenient means of profiling of the phosphoproteome in human cancer cells and is useful to identify potential biomarkers in cancer.

2. Materials

2.1. Sample Preparation

1. Urea, Thiourea, *n*-octyl β-d-glucopyranoside (OG) (Sigma).
2. Tris[2-carboxyethyl] phosphine (TCEP) (Pierce).
3. Glycerol.
4. Protease inhibitor tablet (Roche, Indianapolis, IN).
5. Tissue homogenizer.
6. Centrifuge (refrigerated centrifuge capable of 30,000 × *g*).
7. Desalting column: PD-10 (Sephadex G-25; Amersham Pharmacia Biotech).
8. Lysis buffer: 7.5 M urea, 2.5 M thiourea, 4% *n*-octyl-d-glucopyranoside (OG), 10 mM Tris[2-carboxyethyl]phosphine (TCEP), 10% (v/v) glycerol, 50 mM Tris and 40 μL protease inhibitor solution (10 mg tablet in 1 mL PBS buffer).

2.2. Top-Twelve Protein Depletion (For Serum Sample)

1. ProtromeLab IgY-12 proteome partitioning kit (Beckman Coulter, Fullerton, CA).
2. Spin filters, 0.45 μm, (Millipore, Billerica, MA).
3. 10× Dilution buffer: 100 mM Tris-HCl, pH 7.4, 1.5 M NaCl.
4. 10× Stripping buffer: 1 M glycine-HCl, pH 2.5.
5. 10× Neutralization buffer: 1 M Tris-HCl, pH 8.0.
6. 15-mL, 5-kDa MWCO Amicon filters (Millipore).

2.3. Lectin Affinity Glycoprotein Extraction

1. Agarose bound lectin, Wheat Germ Agglutinin, (WGA) and Agarose bound Concanavalin A (Con A) (Vector Laboratories, Burlingame, CA).
2. Empty spin columns (column capacity: 0.9 mL) (Pierce, Rockford, IL).
3. Lectin affinity binding buffer: 0.15 M NaCl, 20 mM Tris-HCl, 1 mM $MnCl_2$, 1 mM $CaCl_2$, pH 7.4.
4. Elution buffer: 0.3 M *N*-acetyl-glucosamine and 0.3 M Methyl-α-d-mannopyroside in 20 mM Tris-HCl and 0.5 M NaCl, pH 7.0.

2.4. Separation of Proteins Based on pI

2.4.1. Micro-Rotofor Isoelectric Focusing

1. Cathode electrolyte: 0.1 M NaOH.
2. Anode electrolyte: 0.1 M H_3PO_4.
3. Rotofor running buffer: 3% CHAPS, 6 M urea, 2 M thiourea, 10 mM DTT, 2% ampholyte 3–10 (for pH 3–10), 2% ampholyte composed of 0.2% ampholyte 3–10, 0.9% ampholyte 3–5, and 0.9% ampholyte 4–6 (for pH 3–6).
4. Micro-Rotofor system (Bio-Rad).
5. Power supply with constant power capability.
6. 3-mL syringe and 10-mL syringes.
7. Sealing tape.
8. Forceps.
9. Cold trap.
10. Vacuum pump.

2.4.2. Chromatofocusing

1. HPCF-1D column (250 × 2.1 mm) (Beckman Coulter, Fullerton, CA).
2. Polybuffer CF start buffer: 25 mM bis-Tris propane (Sigma), 6 M urea and 0.1% *n*-OG, pH 8.5.
3. CF elution buffer: 7% v/v polybuffer 74 (Amersham Pharmacia, Piscataway, NJ) and 3% polybuffer 96 (v/v), 6 M urea, and 0.1% *n*-OG, pH 4.0.

4. 0.2-µL nylon low extraction membrane filter (500 mL) (Corning Incorporated, Corning, NY).
5. 1 M NaCl solution.
6. Isopropanol (Sigma).
7. HPLC/FPLC system running at ambient temperature and at a flow rate of 0.2mL/min, with fraction collector, UV detector (280 nm).
8. pH electrode/flow cell (Lazar Research Laboratories, Los Angeles, CA).

2.5. Nonporous-Reverse Phase-HPLC Separation

1. Nonporous-reverse phase (NPS-RP)-HPLC column: 33 × 4.6 mm Micra Platinum column, ODSII (recommended) (Eprogen).
2. Column heater.
3. Deionized water was purified using a Millipore RG system (Bedford, MA).
4. Acetonitrile (ACN).
5. Trifluoroacetic acid (TFA).

2.6. Protein Microarrays

2.6.1. Printing

1. Nanoplotter 2.0 (GeSiM).
2. Superamine and Superepoxy glass slides (TeleChem International), Nitrocellulose slides (Whatman, Inc.).
3. Printing buffer: 62.5 mM Tris-HCl (pH 6.8), 1% w/v sodium dodecyl sulfate (SDS), 5% w/v dithiothreitol (DTT), and 1% glycerol. Prepare immediately before reconstituting dried-down sample. Make fresh each time it is needed.
4. Speedvac centrifuge (Thermo, Waltham, MA).

2.6.2. Hybridization of Slides

Glycoprotein Microarrays with Multilectin Detection

1. Blocking buffer: 1% w/v BSA in PBS-T.
2. Primary detection solution: For Concanavalin A (ConA), Aleuria aurentia (AAL), Maackia amurensis (MAL), and Peanut agglutinin (PNA), 5 µg/mL biotinylated lectin solution in PBS-T (0.1% Tween-20) and for Sambuccus Nigra (SNA), 10 µg/mL lectin solution in PBS-T (0.1% Tween-20). All biotinylated lectins purchased from Vector Laboratories (Burlingame, CA).
3. Washing solution: PBS-T (0.1% Tween-20).
4. Secondary detection solution: 1:1,000 solution of 1 mg/mL Streptavidin conjugated to Alexafluor555 (Invitrogen) in 0.5% BSA and 1× PBS-T (0.1% Tween-20).

Phosphoprotein Analysis with Antibodies

1. Blocking buffer: 1% w/v bovine serum albumin (BSA) in PBS-T (phosphate buffered saline and 0.1% v/v Tween-20).

2. Probe buffer: 5 mM magnesium chloride, 0.5 mM DTT, 0.05% v/v TritonX-100, 5% glycerol, and 1% BSA in 1× PBS.
3. Primary antibody solution: 1:500 solution of 1 mg/mL goat polyclonal antiphophoserine and goat polyclonal antiphosphothreonine (Zymed laboratories) and mouse monoclonal antiphosphotyrosine (Upstate signaling) in probe buffer.
4. Secondary antibody solution: 1:1,000 solution of 1 mg/mL donkey antigoat conjugated to Cy5 in probe buffer and 1:1,000 solution of 1 mg/mL donkey antimouse conjugated to Cy5 in probe buffer.

Phosphoprotein Analysis with ProQ Diamond

1. Blocking buffer: 1% BSA in PBS-T (0.1% Tween-20).
2. ProQ Diamond dye (Invitrogen).
3. ProQ Diamond destaining solution (Invitrogen).
4. 1× PBS.

2.6.3 .Slide Scanning

1. Axon 4000A scanner (Molecular Devices, Sunnyvale, CA).

2.7. Protein Identification by Peptide Sequencing

1. SpeedVac concentrator (Thermo).
2. Ammonium bicarbonate (Sigma-Aldrich).
3. TPCK modified sequencing grade porcine trypsin (Promega, Madison, WI).
4. HPLC grade water.
5. Paradigm HPLC pump (Michrom Bio Resources, Inc., Auburn, CA).
6. Nanoreverse phase column (0.075 mm × 150 mm, C18).
7. Peptide nanotrap (0.15 × 50 mm, C18).
8. A 75-μm metal spray tip (Michrom).
9. Finnigan LTQ mass spectrometer (Thermo electron).
10. SEQUEST algorithm incorporated in Bioworks software (Thermo).

2.8. Protein Glycosylation Analysis by Mass Spectrometry

2.8.1. Protein Deglycosylation Using PNGase F

1. PNGase F (QA-Bio, Palm Desert, CA).
2. 5× Reaction Buffer 7.5 for PNGase F: 250 mM sodium phosphate, pH 7.5.
3. Denaturation Solution: 2% SDS, 1 M β-mercaptoethanol.
4. Triton X-100: 15% solution.

2.8.2. Glycan Purification

1. Glycoclean R cartridges (Prozyme).
2. 3–5 mL Syringes.
3. Acetic acid and methanol.

4. Black graphitized carbon cartridge (150 mg, 0.5 mL column volume) (Alltech, Deerfield, IL).
5. Trifluoroacetic acid (TFA).

2.8.3. Glycan Profiling by LC-MS

1. TSKgel Amide-80 resin (5 µm, 80 Å, Tosoh Biosciences, Montgomeryville, PA).
2. Capillary pump (Ultra-Plus II MD, Micro-Tech Scientific, Vista, CA).
3. ESI-TOF (LCT premier, Micromass/Waters, Milford, MA).
4. Solvent: ACN, formic acid, deionized water.

2.8.4. Permethylation of Oligosaccharides

1. Anhydrous dimethyl sulfoxide (DMSO).
2. Sodium hydroxide powder (NaOH).
3. Chloroform ($CHCl_3$).
4. Methyl iodide (CH_3I).

2.8.5. Glycan Structure Analysis

1. MALDI-QIT (Shimadzu Biotech, Manchester, UK).
2. 2′,4′,6′-Trihydroxyacetophenone monohydrate (THAP) (Sigma-Aldrich).
3. 2,5-dihydroxybenzoic acid (DHB).

2.9. Protein Phosphorylation Analysis Using Mass Spectrometry

2.9.1. Enzymatic Dephosphorylation

1. Calf alkaline phosphatase.
2. Zip Tips (Millipore).
3. MALDI-TOF MS (Micromass, Inc., TOFSpec2E).

3. Methods

3.1. Sample Preparation

3.1.1. Cell Lysate of Cell Lines and Tissue Sample

Cell lysis is a process that extracts proteins from cells. In order to keep extracted proteins intact in solution, the proteases are inhibited by protease inhibitors and the pH of the lysis buffer is controlled close to physiological pH. Maintaining low ionic strength is a key issue for the subsequent isoelectric focusing or chromatofocusing experiment. Nonionic detergents, reducing agents, and chaotropic agents are used to enhance the solubility of focused proteins.

1. Tissue or cell line samples were quickly thawed and immediately lysed with 2 mL of lysis buffer.
2. Tissue samples could be homogenized on ice using a tissue grinder, mechanical homogenizer, or mini-bead beater according to manufacturer’s instruction.

3. The sample with lysis buffer should be vortexed frequently over a period of 1 h at room temperature and then centrifuged at approximately 23,000 × *g* for 70 min at 4°C.
4. After the supernatant was collected, buffer exchange was conducted using a PD-10 G-25 column. The PD 10 column was equilibrated with the equilibration buffer of the subsequent experiment and then used to exchange the cell lysate from the lysis buffer to the above buffer according to the manufacturer's protocol.
5. Protein concentration in each sample was quantified by the Bradford assay *(16)*. The buffer exchanged protein mixtures were stored in a −80°C freezer until further use.

3.1.2. Preparation of Serum Sample

1. 40 cm^3 of blood provided by each patient were permitted to sit at room temperature for a minimum of 30 min (and a maximum of 60 min) to allow the clot to form in the red top tubes.
2. The samples were centrifuged at 1,300 × *g* at 4°C for 20 min. The serum was removed, transferred to a polypropylene capped tube, and frozen. The frozen samples were stored at −70°C until assayed.

3.2. Removing High Abundance Proteins Using IgY Antibody Column (For Serum Sample)

The serum proteome is dominated by a few highly abundant proteins which constitute about 90% of the total protein content of the serum. These proteins severely interfere with the quantification and identification of proteins of lower abundance. Most of the high abundance proteins in serum are glycosylated proteins *(12)* and the presence of these high abundance glycoproteins masks the detection of glycoproteins of lower abundance. Although albumin is not a glycoprotein, it binds to many other glycoproteins. Since many important marker proteins are detected in low concentration in biological samples, removing the high abundance proteins may be a critical strategy for serum biomarker discovery. Human serum was depleted using the ProtromeLab IgY-12 proteome partitioning kit. This antibody column removes albumin, IgG, α1-antitrypsin, IgA, IgM, transferrin, haptoglobin, α_1-acid glycoprotein, α_2-macroglobin, HDL (apolipoproteins A-I & A-II), and fibrinogen using an affinity column based on avian antibody (IgY)–antigen interactions.

1. Depletion was performed according to the standard operating protocol of this kit. Before setting up the column onto the LC, the HPLC system is flushed with dilution buffer for 20 min at a flow rate of 2 mL/min to make sure that the tube is filled with dilution buffer. Then the antibody column is equilibrated with dilution buffer for 30 min at a flow rate of 2 mL/min to ensure a flat baseline (*see* **Note 1**).

2. 125-μL serum is diluted in 500 μL dilution buffer and briefly centrifuged using a 0.45 μm spin filter for 1 min at 9,200 × *g* to remove sample particulates (*see* **Note 2**).
3. 625-μL diluted sample is injected to the antibody column. The column is washed with dilution buffer at 0.5 mL/min for 30 min. The unbound fraction, which is the top-twelve protein-depleted fraction, flows out at ~8 min as shown in **Fig. 1** and the fraction is collected. The final volume of depleted fraction in dilution buffer is approximately 15 mL. The sample is concentrated and buffer exchanged with equilibration buffer of next step (lectin binding buffer for glycoprotein microarray experiment) using a 5 kDa MW cutoff Amicon filter.
4. The bound fraction, which is the top-twelve protein fraction, is eluted with stripping buffer for 18 min at 2 mL/min and then the neutralization buffer is applied immediately for 10 min at a flow rate of 2 mL/min to regenerate the column. Finally, the column is re-equilibrated with dilution buffer for 10 min at 2 mL/min.
5. 1/10 volume of 10× neutralization buffer is added to the top-twelve protein fraction.
6. Protein assays are performed by the Bradford assay method. The result indicates that around 5–10% of serum proteins are recovered in the depleted fraction.

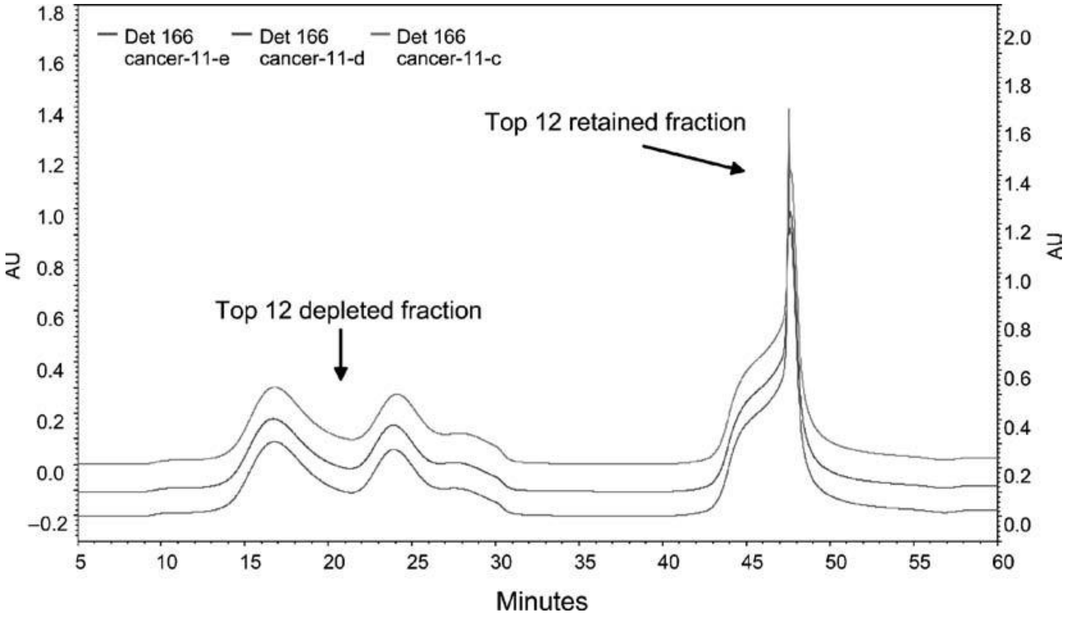


Fig. 1. UV Chromatogram of three125-μL serum depletion from three patients by IgY antibody column to remove 12 high abundant proteins. During the binding process, the fraction flowing through was collected as top-twelve depleted serum fraction and the top-twelve retained protein fraction was collected during elution. The absorption was set at 280 nm.

3.3. Lectin Affinity Glycoprotein Extraction

In proteomics studies, fractionation of samples can help to reduce their complexity and to enrich specific classes of proteins for subsequent downstream analyses. One method for fractionation and enrichment of glycoproteins is lectin affinity chromatography. Lectins are proteins that are able to specifically and reversibly bind carbohydrates. Lectin affinity chromatography has been used for enrichment of different classes of glycoproteins by many groups *(17, 18)*. In this protocol, ConA and WGA lectins are used for extraction of most N-linked glycoproteins. ConA recognizes α-linked mannose including high mannose-type and mannose-core structures. WGA lectin can interact with some glycoproteins via sialic acid residues and it also binds oligosaccharides containing terminal *N*-acetylglucosamine. Almost all N-linked glycoproteins could be extracted since these two lectins can interact with high mannose, complex, and hybrid type glycans.

1. 350-μL agarose-bound WGA and 250 μL agarose-bound ConA were packed into disposable screw end-cap spin column with filters at both ends (*see* **Note 3**).
2. For serum sample, protease inhibitor stock solution is prepared by dissolving one complete EDTA-free Protease inhibitor cocktail tablet in 1 mL H_2O. The stock solution is added to binding buffer and elution buffer at a ratio of (v/v) 1:50. Nondepleted serum sample (~50 μL) is diluted with binding buffer to make the final volume of 500 μL. Depleted serum sample (250–600 μL) is buffer-exchanged into 500 μL binding buffer using Amicon 5000 MWCO filter.
3. The column is first equilibrated with 500-μL binding buffer by centrifuging the spin column at 500 rpm for 2 min three times.
4. Diluted sample is loaded onto the column and incubated for 15 min. The column is centrifuged for 2 min at 500 rpm to remove the nonbinding fraction (*see* **Note 4**).
5. The column is washed with 600-μL binding buffer twice to wash off the nonspecific binding. The captured glycoproteins are released with 250-μL elution buffer and the eluted fraction is collected by centrifugation at 500 rpm for 2 min. This step is repeated twice and the eluted fractions are pooled. Proteins bound with WGA are eluted by 0.3 M *N*-acetyl-glucosamine and proteins bound with ConA are eluted by 0.3 M Methyl-α-d-mannopyroside. The protein assay suggests that around 5–10% of the protein content is extracted from nondepleted serum by the lectin affinity columns.

3.4. First Dimension: Fractionation of Proteins Based on pI

For a complex sample like the nondepleted human serum, a one-dimensional separation might not be sufficient to fractionate glycoproteins, especially due to the many high abundance proteins that include glycoproteins. In addition, glycoproteins can exist in multiple glycoforms. Multiple dimension separation is necessary for nondepleted serum glycoprotein mixtures. Here glycoproteins extracted from nondepleted serum are separated based on their pI in the first dimension. The liquid-phase IEF separation is performed using a commercial apparatus called the Rotofor (Bio-Rad). The protocol describes a standard experiment for the micro-Rotofor including materials and step-by-step separation procedures. The Rotofor separates proteins in solution based on their isoelectric point (pI). A protein is positively charged when the pH of the environment is lower than its pI and negatively charged when the pH of the environment is higher than its pI. Upon application of an electric potential, positively charged proteins move toward the anode and negatively charged proteins move toward the cathode until they are neutral. Thus, the proteins are focused at their respective pIs. The typical pH gradient range is 3–10. For the separation of glycoproteins, a Rotofor sample buffer for formation of pH gradient 3–6 is used for a better resolution in the acidic range since most of the glycoproteins have pI lower than 6. Glycoproteins from human serum can be fractionated into ten pI fractions in this pH range using the microscale Rotofor. Figure 2 shows the pH distribution of the ten wells collected from the IEF separation of glycoproteins enriched from nondepleted human serum. The pH of most fractions fall in the range of 3–6 except fraction 1, 9, and 10. The pH of the fractions in the middle increases smoothly from 3 to 6 with a pH difference down to 0.2.

3.4.1. Fractionation with Microliquid-Phase IEF Cell

1. Ion exchange membrane is equilibrated in appropriate electrolyte solution overnight before experiment (anode membrane in 0.1 M H_3PO_4 and cathode membrane in 0.1 M NaOH) (*see* Note 5).

2. The pH gradient (3–10) is formed by the 2% Bio-lyte 3/10 ampholyte. The acidic pH gradient (3–6) is formed by 2% Bio-lyte ampholyte composed of 10% 3/10 ampholyte, 45% 3/5 ampholyte, and 45% 4/6 ampholyte in the buffer. Chaotropic agent, detergent, and reducing agent are added to the buffer for denaturing and solubilizing the proteins. For separation of lectin-enriched glycoproteins, a 5 k MW cutoff Amicon filter is used to remove the lectin elution buffer from the sample and Rotofor sample buffer is added to the concentrated sample to make the final volume of 2 mL.

3. The focusing chamber is assembled according to the manufacturer's manual. Slowly load the sample through the centermost loading port of the focusing chamber. Make sure that all compartments are filled and no bubbles remain. Seal the loading port with tape (*see* Note 6).

4. Place the focusing assembly into the focusing station. Add 0.1 M H_3PO_4 through the vent hole of anode (red) assembly and 0.1 M NaOH through the cathode (black) assembly. Close the

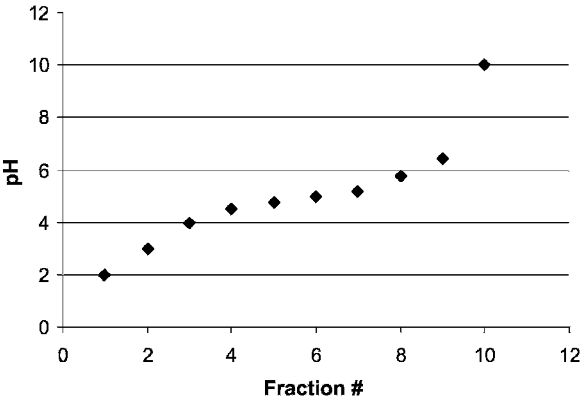


Fig. 2. The pH of the ten fractions collected from micro-Rotofor run was measured. The glycoproteins enriched from 750 μL nondepleted human serum by ConA affinity column were separated by the pH gradient 3–6 formed by mixing of the Biolyte ampholyte 3/10, 3/5, and 4/6.

cooling block and switch to cooling setting I (10°C). Equilibrate the system to the setting temperature for 15 min.

5. Set power at a constant of 1 W. The run is complete when the voltage stabilizes. It usually takes 2.5–3 h to complete the run.
6. Turn off power; remove the seal from loading ports. Harvest the fractions by applying vacuum through the vacuum chamber. Ten fractions are collected in the harvest tray.
7. The proteins are separated to ten fractions. The proteins with pI higher than 6 will be focused in the last fraction and the one with pI lower than 3 will be in the first fraction. The pH of collected fractions can be tested with a pH electrode.

3.4.2. Protein Fractionation by Chromatofocusing

CF is a column-based chromatographic method that uses a weak anion exchange column to generate a descending linear pH gradient. The exchanger is titrated using ampholytes resulting in a pH separation where fractionation can be performed down to <0.1 pH units. A second dimension such as NPS-RP-HPLC can then be used to analyze each pI fraction. This two-dimensional column technology is available in the Beckman ProteomeLab protein fractionation system kit. In the CF process, accurate pH is measured online by a postdetector pH electrode/cell with low dead volume. Each pH range can be automatically collected using a fraction collector. Up to 30 mg proteins could be separated using this technique. The result is a two-dimensional UV map of protein expression, with high reproducibility and excellent fractionation of complex samples such as a cell lines or tissue lysates.

1. CF separation is performed on a HPCF-1D column (250 × 2.1 mm) (Beckman Coulter). The pH of the start buffer is adjusted to pH 8.5 and the pH of elution buffer is adjusted to pH 4.0 with saturated iminodiacetic acid solution and concentrated ammonium hydroxide. Prepared buffer is filtered using a 0.2 µm low extractable membrane filter to remove particles. Prior to sample loading, the column is equilibrated with the start buffer for ~2 h at a flow rate of 0.2 mL/min.
2. After equilibration, 5 mg proteins in start buffer are loaded onto the CF column (*see* **Note 7**).
3. Elution is achieved by applying a pH 4.0 elution buffer at a flow rate of 0.2 mL/min. A linear pH gradient is generated at the outlet of the column in ~1 h. As a result, proteins elute off the column sequentially according to their pI. Separation is monitored at 280 nm using a UV detector. Fractions are collected every 0.2 pH unit change from pH 8.5 to pH 4.0.
4. The column is washed with 1 M NaCl for 1 h followed by H_2O for 1 h at flow rate of 0.2 mL/min.
5. The column could be washed with isopropanol for 1 h at 0.2 mL/min for removal of the residual proteins and storage of the column (*see* **Note 8**).

3.5. Second Dimension: Fractionation of Proteins Based on Hydrophobicity

Separation in reverse phase (RP)-HPLC is based on hydrophobicity. Particle size, pressure, recovery, stability, sample capacity, and efficiency need to be balanced when selecting the right column for a separation. This protocol employs 1.5-µm ODS (C18) nonporous silica (NPS) particles. NPS columns provide a faster high-resolution separation with improved protein recovery than porous columns since they eliminate issues of macromolecular pore adsorption and stagnant mobile phase in the intraparticle void volume associated with porous stationary phases *(19, 20)*. As a result, higher column efficiency is obtained with much improved resolution. NPS packing materials are generally more stable than their porous counterparts at elevated temperatures. In addition, small (1.5-µm) silica particles compensate for the decreased surface area, hence the sample capacity associated with nonporous materials. The high efficiency of NPS 1.5-µm packing allows for reproducible, high-resolution, high-speed protein separations. The enriched glycoproteins or the collection from the pI fractionation (CF or Rotofor) is separated by NPS-RP-HPLC in the second dimension. The LC separation chromatogram of lectin-extracted glycoproteins from depleted serum is shown in **Fig. 3a** and the two-dimensional UV map of micro-Rotofor fractionated glycoproteins (pH 3–6) from nondepleted serum is shown in **Fig 3b**. It is observed that the depletion step effectively simplifies the glycoprotein content in serum (**Fig. 3a**) and most of the glycoproteins in serum are focused in ~pH 5 (**Fig. 3b**).

For phosphorylation analysis, the proteins from whole cell lysate are separated by CF as the first dimension and by NPS-RP-HPLC as the second dimension.

1. Sample in aqueous buffer is loaded onto NPS-RP-HPLC. High separation efficiency is achieved by using an ODSII (4.6 × 33 mm) column (*see* **Note 9**).
2. To collect purified proteins from NPS-RP-HPLC, the reversed-phase separation was performed at 0.5 mL/min and monitored at 214 nm using a Beckman 166 Model UV detector. Proteins eluting from the column are collected. The reversed phase column is heated to 60°C by a column heater.

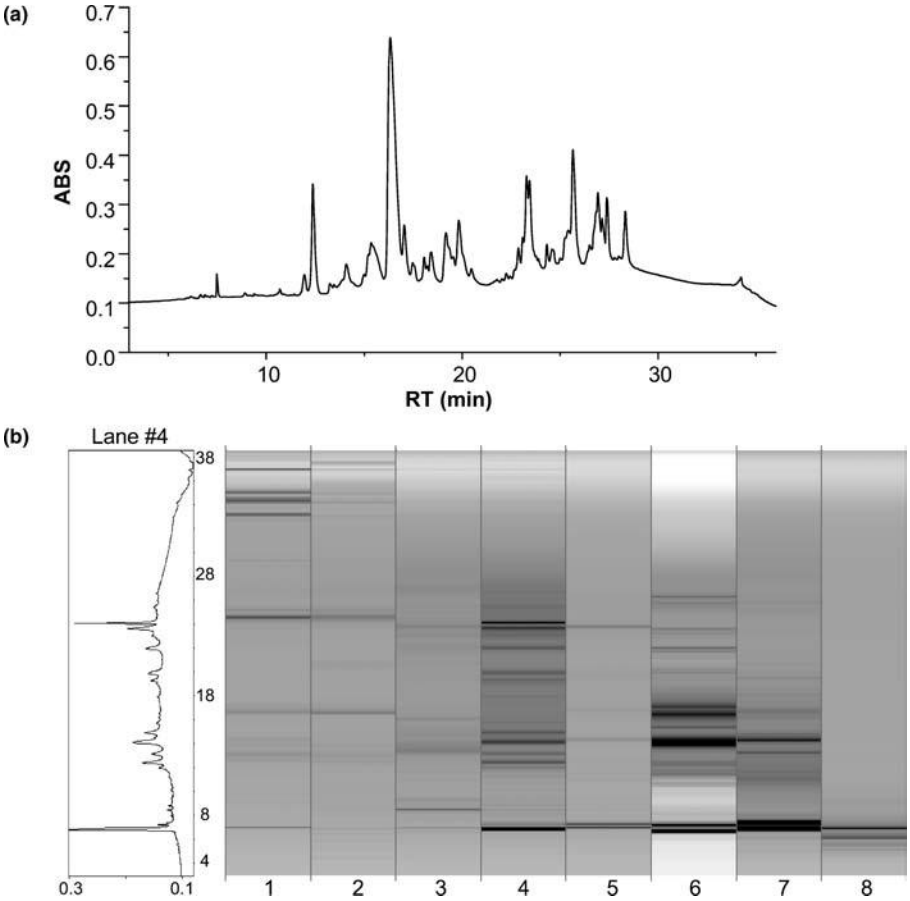


Fig. 3. **a** NPS-RP-HPLC separation of ~30 μg glycoproteins extracted from depleted serum by ConA-WGA lectin column. The UV wavelength is 214 μm. Flow rate is 0.5 mL/min. **b** Two-dimensional UV map generated from the NPS-RP-HPLC separation of glycoproteins extracted from nondepleted human serum. Rotofor fractions 2–9 (pH 3–6) are shown. The *x*-axis is the fraction number corresponding to Fig. 2 and the *y*-axis is the retention time. The chromatogram shown in *left* is the lane #4 (fraction #5).

3. Both mobile phase A (water) and B (ACN) contain 0.1% v/v TFA. The gradient profile used is as follows: 5–15% B in 1 min, 15–25% B in 2 min, 25–30% B in 3 min, 30–41% B in 15 min, 41–47% B in 4 min, 47–67% B in 5 min and 67–100% B in 2 min.

3.6. Protein Microarrays

Protein microarrays are a convenient method to simultaneously scan multiple proteins for a specific property. When separated proteins are exchanged into an appropriate printing buffer, the proteins can be deposited on a covalently modified glass slide or nitrocellulose slide where they bind either covalently or by strong noncovalent interactions. The resulting protein microarray can then be probed with a variety of agents to determine reactivity of proteins. To determine phosphorylated proteins, the slides can be incubated with ionic phosphodyes or antiphosphoprotein antibodies. In contrast, the same slides can be probed with multilectins through a biotin–streptavidin detection scheme to elucidate the glycosylation status of the same proteins. Five lectins are used for detection of glycoproteins with different structures. AAL recognizes fucose linked (α-1,6) to *N*-acetylglucosamine or to fucose linked (α-1,3) to *N*-acetyllactosamine. ConA recognizes α-linked mannose including high mannose-type and mannose-core structures. Both MAL and SNA recognize sialic acid on the terminal branches, while SNA binds preferentially to sialic acid attached to terminal galactose in an (α-2,6) and to a lesser degree, an (α-2,3) linkage *(21)*. MAL could detect glycans containing NeuAc-Gal-GlcNac with sialic acid at the 3 position of galactose *(22)*. In contrast, PNA binds desialylated exposed galactosyl (β-1,3) *N*-acetylgalactosamine. In fact, sialic acid in close proximity to the PNA receptor sequence will inhibit its binding. Use of the combination of these five lectins should be highly successful in covering a majority of N-glycan types reported and differentiating them according to their specific structures. When such methods are used in a sandwich assay format, high sensitivity with low limits of detection can be obtained. Studies with standard proteins established the limits of detection to be in the 2.5–5-fmol range with high lectin selectivity *(10)*.

Figure 4 shows a glycoprotein microarray section highlighting differences in glycosylation in H factor 1 and alpha acid glycoprotein as determined by processing with five different lectins. When processing human samples, protein concentrations may vary between individuals. In order to eliminate concentration effects all data is normalized with respect to peak heights of the data being analyzed. The bar graph under array images show the extent to which the abundance or specific glycan structures change with respect to disease. For example, it is evident that mannosylation is more abundant in pancreatitis compared to normal serum (as seen in response with ConA) while fucosylation

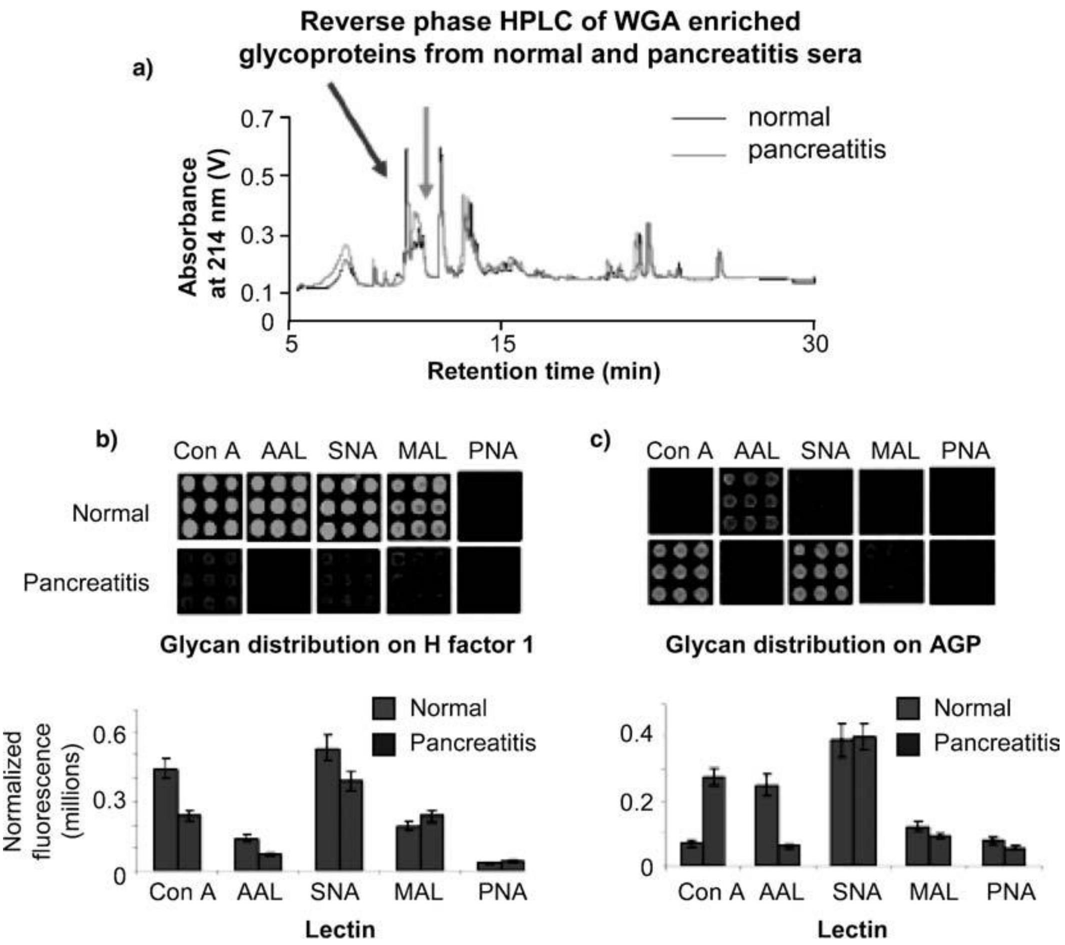


Fig. 4. Identifying differences in glycosylation from sera of different biological states. **a** Reversed-phase chromatogram of enriched glycopoteins from normal and pancreatitis sera with differences highlighted. *Bar graphs* showing integrated fluorescence values of spots shown in array images after background subtraction and normalization based on UV peak area for peak shown with **b** red arrow and **c** orange arrow in the reversed-phase chromatograms (reproduced from ref. 10 with permission from ACS).

is more abundant in normal compared to pancreatitis serum (as seen in response with AAL).

Figure 5 shows sections of a protein microarray probed with ProQ Diamond phosphoprotein gel stain. The Sum52PE cell line was studied and the effect of an inhibitor on protein phosphorylation events in the FGFR2 pathway was studied *(15)*. Both Sum52PE cell lines are separated by CF followed by RP-HPLC. It can be seen that important proteins such as Eps15, STAT3, and MAPK7 interacting proteins were dephosphorylated when the cell line was treated with the inhibitor.

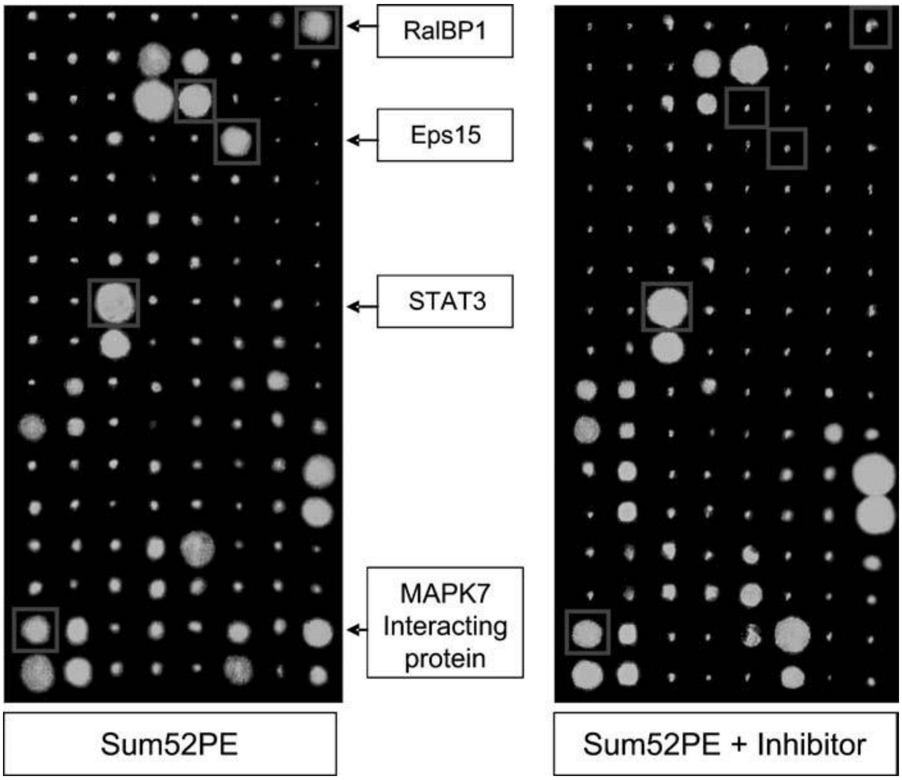


Fig. 5. Phosphoprotein microarray images of two cell lines highlighting phosphorylation differences between the two. On the *left* is the Sum52PE cell line while on the *right* is the Sum52PE treated with inhibitors of a key signaling pathway. Notice how some proteins are dephosphorylated in this cell line (Proteins of interest are highlighted in *blue boxes* and their IDs are indicated by *arrows*.) (Reproduced from ref. *15* with permission from ACS).

3.6.1. Printing

1. Fractionated proteins in 96-well plates are dried down completely using a speedvac centrifuge.
2. The dried fractions are stored at –80°C until needed.
3. Prior to printing, all fractions are resuspended in 15-μL printing buffer and left on a gel shaker at 4°C, overnight.
4. The printing program is set to print five droplets of 500 pL each per spot. All spots are printed 0.6 mm apart and the spot diameter is about 400 μm.
5. At the end of printing all slides are stored in a dessicator cabinet at 4°C.

3.6.2. Hybridization of Slides

Hybridization of Slides: Glycoprotein Analysis with Multilectins

1. Printed slides are blocked overnight by placing each slide in a 50 mL tube containing 40 mL of blocking buffer.
2. In the morning the slide is incubated in primary incubation solution. Each slide is placed in a plastic pouch (made in-house using Kpak sealPAK pouches and a heat sealer) containing 4 mL of primary incubation solution. The slides are incubated for 2 h on a rotator at 4°C.

3. After primary incubation, the slides are washed five times for 5 min each in washing solution. Washing is done by placing all slides on a slide rack in a container with probe buffer such that all slides are completely submerged followed by shaking it on a gel shaker.
4. The slides are subsequently incubated in secondary incubation solution using the same procedure as primary incubation. However, the incubation time is only 1 h.
5. After secondary incubation, the slides are again washed five times for 5 min each in washing solution.
6. The slides are then spin-dried using a high-speed centrifuge.
7. It is recommended that all slides be scanned immediately after processing for best results. However, if stored, they should be kept at 4°C in a dessicator cabinet.

Hybridization of Slides: Phosphoprotein Analysis with Antibodies

1. Printed slides are blocked overnight by placing each slide in a 50-mL tube containing 40 mL of blocking buffer.
2. The tubes are attached to a Clay Adam's brand nutator and left with shaking at 4°C.
3. In the morning the slide is incubated in primary incubation solution. Each slide is placed in a plastic pouch (made in-house using Kpak sealPAK pouches and a heat sealer) containing 4 mL of primary incubation solution. The slides are incubated for 2 h on a rotator at 4°C.
4. After primary incubation, the slides are washed five times for 5 min each in probe buffer. Washing is done by placing all slides on a slide rack in a container with probe buffer such that all slides are completely submerged followed by shaking it on a gel shaker.
5. The slides are subsequently incubated in secondary incubation solution using the same procedure as primary incubation. However, the incubation time is only 1 h.
6. After secondary incubation, the slides are again washed five times for 5 min each in probe buffer.
7. Once all antibodies are washed off, the slides are spin-dried using a high-speed centrifuge.
8. It is recommended that all slides be scanned immediately after processing for best results. However, if stored, they should be kept at 4°C in a dessicator cabinet.

Hybridization of Slides: Phosphoprotein Analysis with ProQ Diamond

1. Printed slides are blocked overnight by placing each slide in a 50-mL tube containing 40 mL of blocking buffer (*see* **Note 10**).

2. The slides are then incubated for 1 h in ProQ Diamond phosphoprotein gel stain with shaking at room temperature. A tube filled with ProQ dye can be used for two slides.
3. After incubation, the slides are destained by washing with destaining solution three times for 10 min each.
4. Once destaining is complete the slides are rinsed with PBS to wash off any residual solutions after which the slides are spun dry using a high-speed centrifuge.
5. It is recommended that all slides be scanned immediately after processing for best results. However, if stored, they should be kept at 4°C in a dessicator cabinet.

3.6.3. Slide Scanning

1. Dried, processed slides are scanned on an Axon 4000A scanner.
2. Cy3, Alexa 555 and ProQ treated slides are scanned in the green channel; while Cy5 and Alexa 647 are scanned in the red channel.
3. The photomultiplier tube (PMT) gain should be adjusted until best signal without saturation is seen. Typically, this value ranges from 400 to 700 on our instrument.

3.7. Protein Identification by Peptide Sequencing

The purified protein with alterations as detected by protein microarray was collected and subjected to trypsin digestion. The tryptic-digested peptides were further separated by microscale or nanoscale liquid chromatography and analyzed using mass spectrometry. Liquid chromatography-tandem mass spectrometry is a rapid, sensitive method for protein identification. This method allows identification of proteins at low fmol levels and enables identification of numerous proteins in protein mixtures. High-quality MS/MS spectra that contain peptide sequence information could be generated by tandem mass spectrometer. b and y ions are predominantly generated by collision induced dissociation (CID) where peptides were fragmented by collision with nonreactive gases. The commercial available SEQUEST computer program is used to match the sequence information in the spectra to a database of known protein sequences. High-throughput protein identification is enabled by the constantly expanding protein sequence information in various databases.

3.7.1. Tryptic Digestion

1. Fractions obtained from NPS-RP-HPLC (~1–2 μg) are concentrated down to ~20 μL using a SpeedVac concentrator operating at 45°C.
2. 20 μL of 100–300 mM ammonium bicarbonate was then mixed with each concentrated sample to obtain a pH value of ~7.8.

3. 0.5 μL of TPCK modified sequencing grade porcine trypsin was added and vortexed prior to a 12–16 h incubation at 37°C on an agitator.

3.7.2. Peptide Separation by Capillary RP-HPLC and Sequencing by Tandem Mass Spectrometry

1. Before separation, peptide samples were dried down and redissolved in 10 μL HPLC grade water.
2. The peptide mixtures were separated using a reverse phase column attached to a Paradigm HPLC pump. For nano-LC-ESI-MS/MS experiments, a nanotrap platform was set up prior to the electrospray source. It includes a peptide nanotrap (0.15 × 50 mm) and a separation column (0.075 mm × 150 mm, C18). The peptide sample was injected and first desalted on the trap column with 5% solvent B (0.3% formic acid in 98% ACN) at 50 μL/min for 5 min
3. The peptides were eluted using a 45 min gradient from 5 to 95% B at a flow rate of 0.25 μL/min, where solvent A was 0.3% formic acid in 2% ACN.
4. A Finnigan LTQ mass spectrometer equipped with a nano-ESI source was used to acquire spectra. A 75-μm metal spray tip was used and the spray voltage was set at 3 kV. The capillary temperature was 200°C, spray voltage was 4.2 kV, and capillary voltage was 30 V.
5. The instrument was operated in data-dependent mode with dynamic exclusion enabled. The MS/MS spectra on the five most abundant peptide ions in full MS scan were obtained. The high abundant solvent or contamination peaks were placed into an exclusion list to avoid being selected for fragmentation. The normalized collision energy was set at 35% for MS/MS.

3.7.3. Database Searching

All MS/MS spectra were searched against the human protein database from SwissProt using SEQUEST algorithm incorporated in Bioworks software and the Swiss-Prot human protein database. The sequence-database search tool was set to expect the following variable modifications: Methione oxidation (+18 Da), carboxymethylated cysteines if the alkylation was preformed (+57 Da). Trypsin was used as a specific protease with two missed cleavages allowed. Searches were repeated without specification of a protease to search for nonspecific cleavage products. Positive protein identification was accepted for a peptide with X_{corr} of greater than or equal to 3.0 for triply-, 2.5 for doubly- and 1.9 for singly charged ions.

3.8. Protein Glycosylation Analysis by Mass Spectrometry

The glycoproteins of interest were collected from LC separation and further characterized using mass spectrometry. The molecular weights and the composition of the enzyme-released glycans could be obtained by mass mapping; a further stage of fragmentation is needed to obtain the sequence or linkage information. The experimentally observed masses were matched with the theoretical value by adding the residue masses of the monosaccharides, the reducing terminal derivative, the mass of any additional groups, and the mass of the adduct. Analyzing the glycosyl moieties on proteins usually involves four steps (a) glycan release and purification; (b) separating (profiling) the glycans in the pool; (c) derivatization of the glycans; (d) glycan structure analysis using mass spectrometry.

3.8.1. Protein Deglycosylation Using Pngase F

Intact N-linked glycans can be released from glycoproteins or glycopeptides enzymatically with peptide *N*-glycosidase F. This enzyme is an amidase that cleaves the linkage between the core GlcNAc and the Asn residue of all asparagine-linked complex, hybrid or high mannose oligosaccharides except alpha(1–3) core fucosylated glycoproteins which is commonly found on plants and can be cleaved by endoglycosidase A. The optimum working pH for this enzyme is 6–10 with the optimum ~7.5. Steric hindrance slows or inhibits the action of PNGase F in some cases. Denaturation of the glycoprotein by heating with SDS and β-mercaptoethanol (β-ME) greatly increases the rate of deglycosylation.

1. Add 10 μL 250 mM sodium phosphate pH 7.5 and 2.5 μL 20× denaturation solution. Heat the solution at 100°C for 5 min.
2. Cool down the solution and add 2.5 μL of Triton X-100.
3. Add 2.0 μL (0.01 U) of PNGase F to the reaction. Incubate overnight at 37°C.

3.8.2. Purification of Enzyme-Released Glycans

N-glycanse reaction may contain salt, detergents, proteins (including the enzyme), and released glycans in the reducing form. The glycan sample needs to be purified for further analysis. A reverse phase cartridge could be used to remove the proteins/peptides and detergent since the oligosaccharides were more hydrophilic and will not retain on the cartridge. The glycans could be desalted using a graphitized carbon column as the oligosaccharides are more hydrophobic than salt and will bind this material. The columns or cartridges packed with hydrophilic material could also be used for cleaning the glycan samples.

1. Wash the reverse phase cartridge (e.g., Glycoclean R cartridges from Prozyme) with 5 mL methanol followed by 5 mL 5% acetic acid.
2. The sample should be in aqueous solution with organic solvent less than 5%. Use syringe to pass the sample through the cartridge and collect the eluent. Washed the cartridge with 1 mL 2.5% acetic acid three times and pool the eluent.
3. Wash the graphitized carbon column with ACN and water.

4. Pass the eluent from the RP cartridge through the column and wash the column with 1 mL deionized water three times.
5. The glycan sample was eluted with 2 mL 25%ACN and 2 mL 40% ACN in 0.05% TFA and the eluent was pooled.

3.8.3. Profiling the Enzyme-Released Oligosaccharides

Separation of oligosaccharides is challenging because of the complexity of structures. A method commonly used for liquid separation of complex carbohydrate mixtures is HPLC followed by online ESI-MS detection *(23–25)*. Fragile groups, such as sialic acid, which are unstable in MALDI due to in-source decay, are retained using this soft ionization technique *(26)*. For liquid-based separation of glycans, analytical scale normal phase chromatography following fluorescent labeling *(27, 28)* and graphitized carbon chromatography are generally used *(29, 30)*. Normal phase separations are based upon hydrophilic interactions of the oligosaccharide hydroxyl groups with the stationary phase. Samples applied in high concentrations of organic solvent adsorb to the column surface and are eluted with an aqueous gradient so that glycans are resolved on the basis of hydrophilicity. Typically, larger oligosaccharides are more hydrophilic than smaller ones, and require higher concentrations of aqueous solvent to elute. Isocratic structures can be separated using the high resolution of this technique. As shown in **Fig. 6**, N-linked carbohydrates were separated by μ normal phase (0.2 × 120 mm) LC and detected online by ESI-TOF MS. All the proposed glycan structures are shown on the corresponding doubly charged peaks. Elution positions of the glycans could be expressed as glucose units (GU) by comparison with the elution times of the glucose ladder. The contribution of individual monosaccharides to the overall GU value of a given glycan can be calculated, and these values can be used to predict the structure of an unknown sugar from its GU value.

1. The capillary normal phase column (200 μm × 120 mm) was packed in-house using TSKgel Amide-80 resin (5 μm, 80 Å, Tosoh). A capillary pump was used for separation. The capillary column was directly mounted to a microinjector with a 500-nL internal sample loop. The flow from the solvent delivery pump was split precolumn to generate a flow rate of ~3 μL/min.
2. Before injection, purified complex N-glycan mixtures were dried down and redissolved in 80% ACN. The glycans were separated in a 45 min linear gradient which is 100% B to 40% B in 45 min. Solvent A is 50 mM formic acid with pH 4.5 adjusted by ammonium hydroxide. Solvent B is 20% solvent A in ACN.

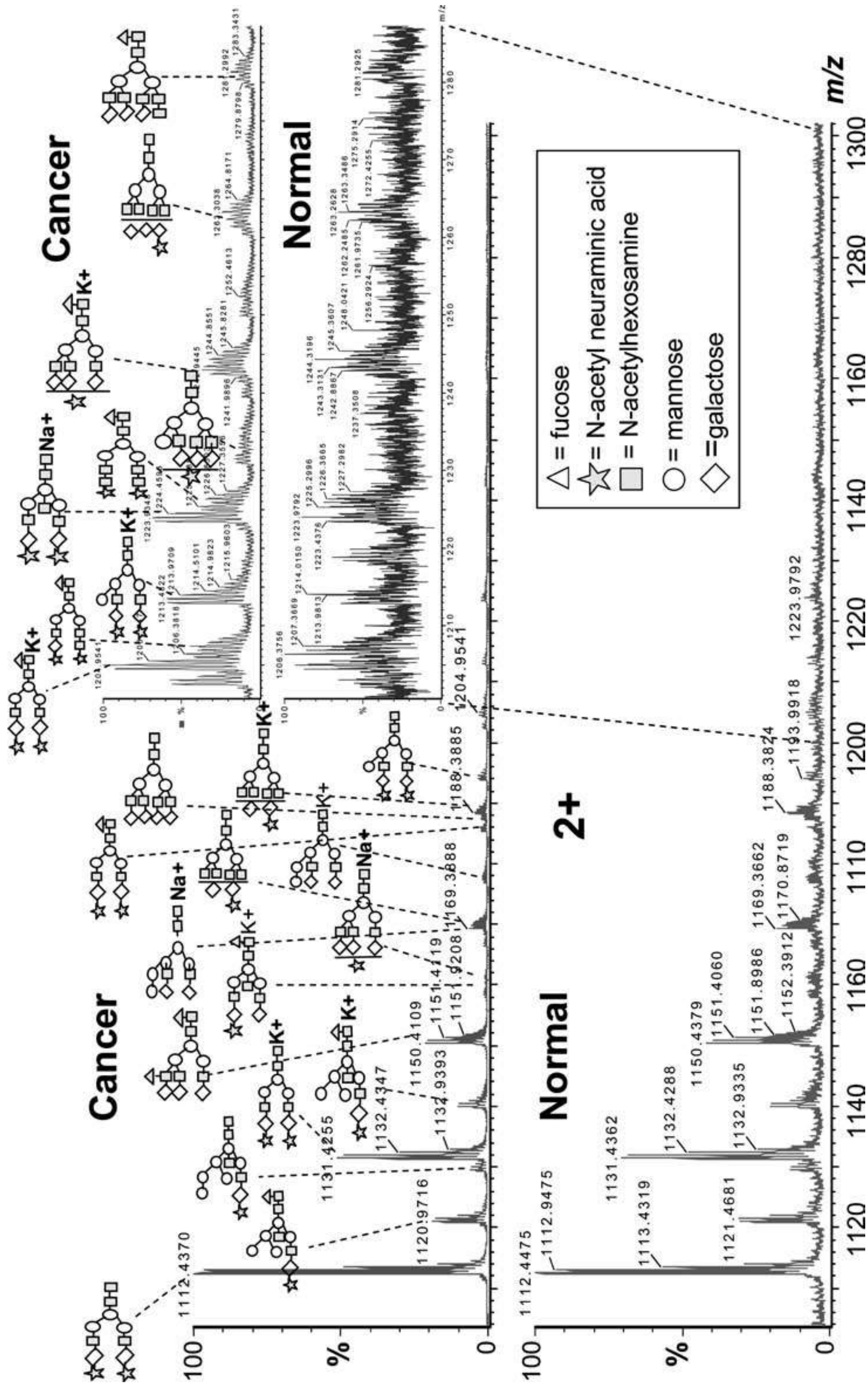


Fig. 6. N-linked carbohydrates were separated by μ normal phase LC and detected online by ESI-TOF MS. The combined mass spectra of retention time 28–29 min are shown. All of the peaks presented here are doubly charged. The spectrum for cancer sample is shown above the normal sample. The proposed glycan structures are shown on the corresponding peaks. *Triangle*, fucose; *star*, *N*-acetyl neuraminic acid (sialic acid); *square*, HexNAc; *circle*, mannose; *diamond*, galactose.

3.8.4. Permethylation of Oligosaccharides

Commonly used derivatization reagents include 2-aminobenzamide, *p*-aminobenzoic acid ethyl ester 2-aminopyridine, etc. This type of derivatization facilitates chromatographic purification and enhances the formation of reducing-end fragment ions in MS experiments. Permethylation is the most popular full derivatization method. Most of the functional groups are converted into their methylated derivatives as permethylation stabilizes the sialic acid residues in acidic oligosaccharides, yielding more predictable ion products when subjected to MS/MS experiments. In addition, methylated sugars seem to ionize more efficiently. Figure 7 shows the MS/MS spectrum of a permethylated mono-sialylated biantennary structure where detailed structure information is yielded by CID fragmentation of permethylated glycans.

1. For permethylation, the glycan sample (~1 μg) was dried and add 2 μL (this volume should be scaled up if more sample are analyzed) of freshly vortexed DMSO/NaOH suspension (0.2 g NaOH in 500 μL anhydrous DMSO) were added. The DMSO/NaOH reagent may be stored in desiccators at room temperature indefinitely but must be re-votexed before each use.

2. After incubation for 30 min, one and half volume of iodomethane was added and the mixture was allowed to sit for another 30 min. The same volumes of DMSO/NaOH and CH_3I were added again and were left to stand as above for 60 min.

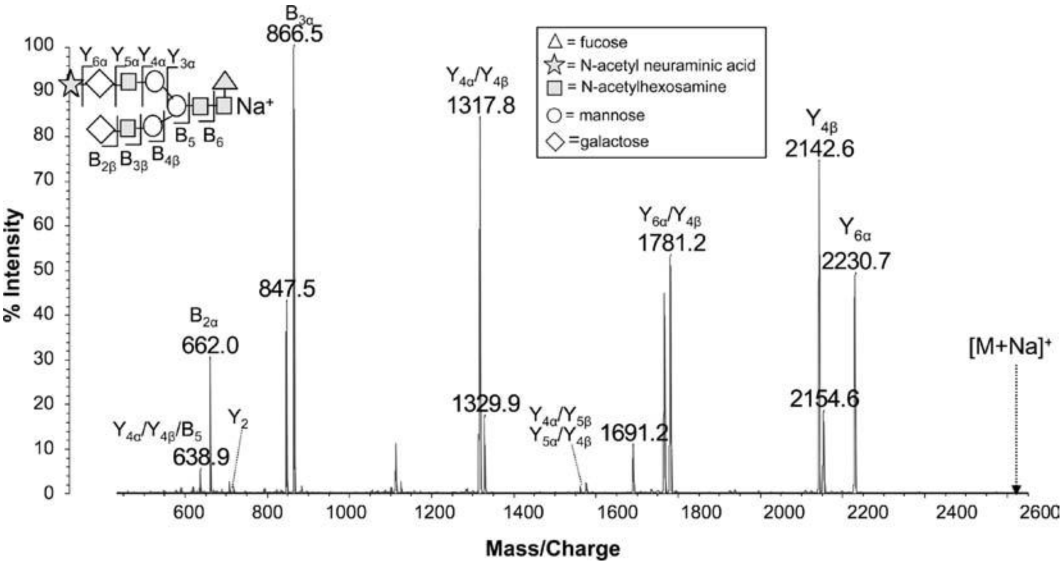


Fig. 7. Fragmentation spectra of permethylated sialylated oligosaccharide $[M + Na]^+$ obtained from MALDI-QIT MS in positive mode. The structure is mono-sialylated bi-antennary structure with fucose: $(Hex)_2$ $(HexNAc)_2(NeuAc)_1(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ at *m*/*z* 2,604.95.

3. Five volumes of $CHCl_3$ were added and the sample was then washed with two volumes of water. The water wash was repeated until the aqueous layer was no longer basic. The organic phase was dried in a centrifugal evaporator and the permethylated glycans were redissolved with 50% methonal for MALDI analysis.

3.8.5. Glycan Structure Analysis Using Tandem Mass Spectrometry

Fragmentation of glycans occurs by cleavage between the sugar rings (glycosidic cleavage) and cleavage across the rings. Glycosidic cleavage gives information on composition, sequence, and sometimes branching. Cross-ring cleavages provide linkage information. Glycosidic cleavages are the most commonly observed in the collision-induced dissociation. Cross-ring cleavage ions are usually minor peaks in the spectra of sugars ionized as sodium or potassium adducts by low-energy collisions. Abundant cross-ring cleavage ions can be produced in instruments capable of producing high-energy collisions. Glycan structural isomers could be revealed by their specific fragments. A sialic acid is easily lost in positive-ion MS/MS spectra. In the negative-ion mode, however, sialic acid is relatively stable *(31, 32)*. MALDI has proved to be the most effective method to ionize N-link carbohydrates since it does not require the carbohydrates to be derivatized *(33)*. Ion trap instruments have the capability to perform multiple successive stages of fragmentation, which allows probing the details of carbohydrate structure. The interface of MALDI to the quadrupole iontrap-ToF provides a means of performing multiple stages of CID with high mass accuracy and resolution *(34–37)*. This unique instrument is used to study the derivatized or underivatized glycans from the target protein where the carbohydrate could be studied with fragmentation up to MS4 *(38, 39)*. The MS/MS spectrum of a premethylated glycan with one sialic acid and a high mannose glycan were shown in **Figs. 7** and **8**. It was observed that $[M + Na]^+$ ion were generated in positive mode and the b and y ions from glycosidic cleavage dominate the spectrum.

1. MS and MS^2 spectra of glycan samples were acquired on a Axima QIT MALDI quadrupole ion trap-TOF (MALDI-QIT). 25 mg/mL 2,5-dihydroxybenzoic acid (DHB) was prepared in 50% ACN with 0.1% TFA.

2. For positive mode analysis, the oligosaccharide solution was applied to the MALDI probe followed by 0.01 M NaCl (0.5 μL). After the sample was dried in air, 0.5 μL matrix was added. For negative mode, matrix solution was prepared by dissolving 5 mg/mL 2′,4′,6′-Trihydroxyacetophenone monohydrate (THAP) and 10 mM ammonium citrate in 50% ACN.

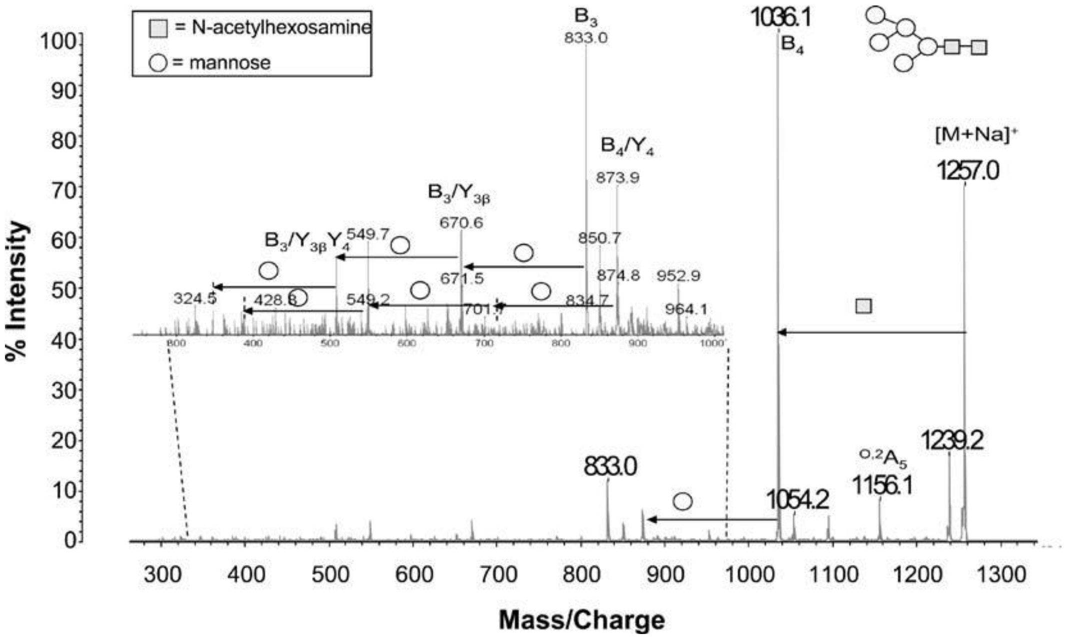


Fig. 8. Fragmentation spectra of high mannose oligosaccharide $[M + Na]^+$ obtained from MALDI-QIT MS in positive mode. The composition of this structure is $(Hex)_5$ $(HexNAc)_2$ + $(Man)_3(GlcNAc)_2$ at *m/z* 1,257.0 (reproduced from ref. *10* with permission from ACS).

3. Acquisition and data processing were controlled by Launchpad software (Karatos). A pulsed N_2 laser light (337 nm) with a pulse rate of 5 Hz was used for ionization. Each profile results from two laser shots. Argon was used as the collision gas for CID and helium was used for cooling the trapped ions.
4. TOF was externally calibrated using 500 fmol/µL of bradykinin fragment 1–7 (757.40 *m/z*), angiotensin II (1,046.54 *m/z*), P14R (1,533.86 *m/z*), and ACTH (2,465.20 *m/z*) (Sigma-Aldrich) for both positive and negative mode.

20.3.9. Protein Phosphorylation Analysis Using Mass Spectrometry

3.9.1. Enzymatic Dephosphorylation

Calf intestine alkaline phosphatase has been widely utilized to dephosphorylate both proteins and nucleic acids. Alkaline phosphatase has a broad pH profile with an optimum between pH 8.0 and 8.5. Calf intestine alkaline phosphatase effectively dephosphorylates proteins containing phosphoserine and phosphothreonine residues, which together account for >97% of the protein-bound phosphate in eukaryotic cells. This enzyme can be utilized for in vitro dephosphorylation of proteins under conditions that do not denature the substrate proteins. In this approach, tryptic digested mixtures from intact and dephosphorylated protein are mass mapped by MALDI-TOF for detection of the shifted mass due to the dephosphorylation. In Fig. 9a–d, the MALDI spectra

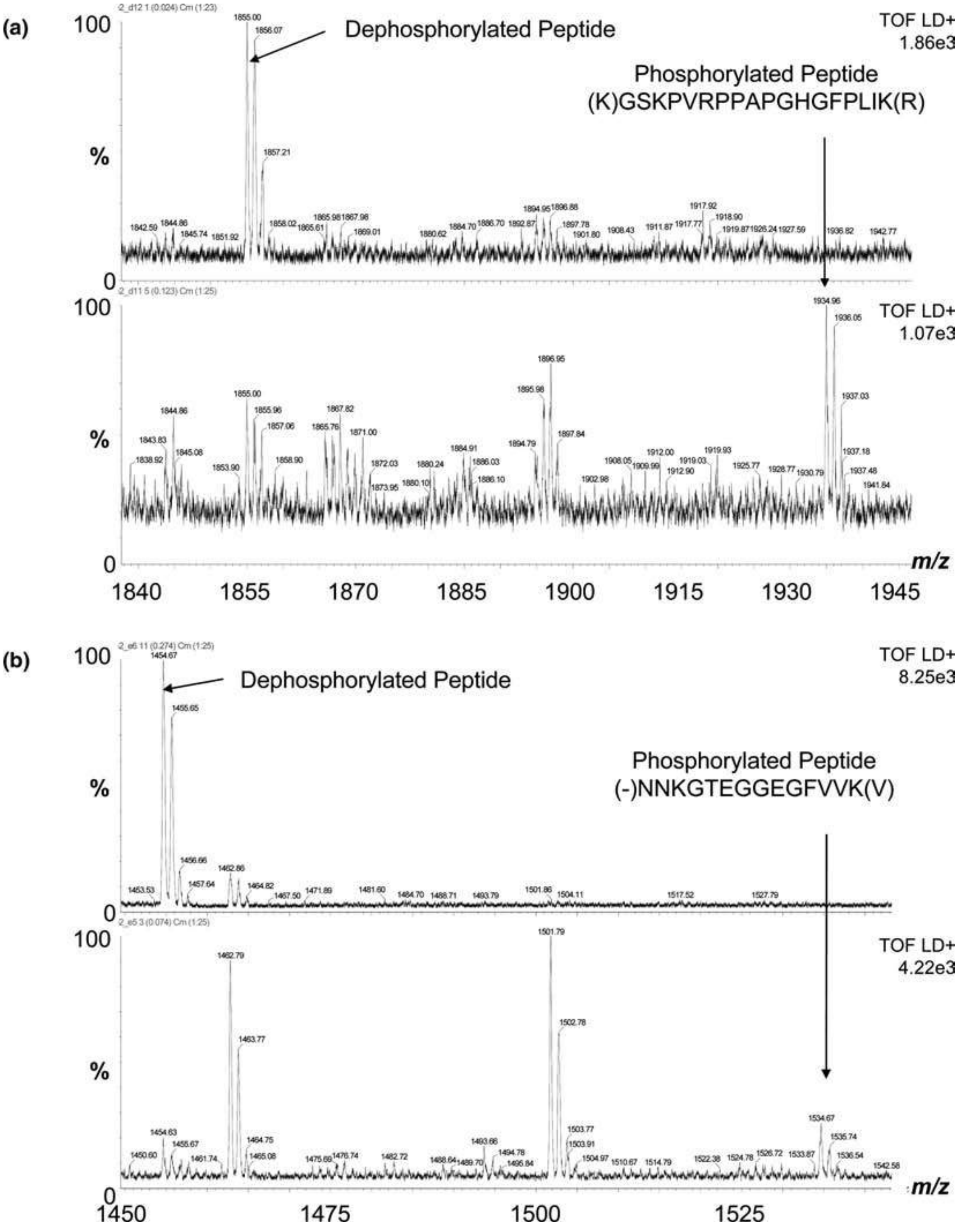


Fig. 9. **a** MALDI spectrum of Rab13 interacting protein (MIRab13) (MICAL-like protein 1) obtained before (*bottom*) and after (*top*) dephosphorylation by CAP. **b** MALDI spectrum of heterogeneous nuclear ribonucleoprotein H (hnRNP H) obtained before and after dephosphorylation (reproduced from ref. *15* with permission from ACS).

of Rab13 interacting protein and heterogeneous nuclear ribonucleoprotein H before and after dephosphorylation by CAP were shown. The observed mass shifts clearly indicate the phosphorylation on this peptide.

1. After completion of the proteolytic cleavage, the samples were divided into two equal parts. The enzymatic dephosphorylation step was performed by treating one part with five units calf alkaline phosphatase reconstituted in 25 mM NH_4HCO_3 buffer (pH 8.0).
2. The mixture was incubated at 37°C for 2 h, and 2.5% TFA was added to stop the enzymatic reaction. The other part was treated as a control.
3. In preparation for MALDI-MS, the samples were first aspirated using Zip Tips, and then 1 μL of the eluent was mixed with an equal volume of CHCA matrix solution prepared in 60% ACN/0.1% TFA and spotted on a MALDI plate.
4. Once the spot dried, 1 μL of 9:1 THAP/DAC matrix solution prepared in 60% ACN/0.1% TFA was applied on top. The spot was allowed to dry slowly afterward.
5. MALDI-TOF MS (Micromass) was used to generate peptide masses of analyzed protein. To verify and locate the phosphorylation sites on proteins, the MALDI-MS spectrum of the phosphorylated (control) and the dephosphorylated samples were compared.

3.9.2. Detection of Phosphorylation Site by Neutral Loss Experiment

The low abundance of phosphopeptide coupled with the higher acidity increases the complexity of their analysis by mass spectrometry in positive ESI mode. In tandem mass spectrometry, phosphopeptide precursor ions typically exhibit a prominent neutral loss of phosphate group (98 Da) during fragmentation. To overcome inadequate peptide fragmentation and diagnostic sequence ion information typical in phosphopeptide analysis, a data-dependent experiment could be carried out on a linear ion trap mass spectrometer (LTQ). It selectively triggers MS3 scans on only the MS/MS fragment ions where a specific neutral loss ion (–98 Da) was detected. The increased ion capacity and improved trapping efficiency of the linear ion trap significantly enhances the quality and abundance of ions for the MS3 data that is essential for identification of phosphorylated proteins.

4. Notes

1. Do not freeze this column, store it in 4°C with 1× dilution buffer containing 0.02% (w/v) sodium azide.

2. The Proteome Lab IgY-12 columns for different sample capacity (20 μL, 50 μL, 250 μL) are also available from Beckman Coulter.
3. This spin column is for up to ~50 μL nondepleted serum sample; larger scale column could be packed if higher amount of sample is needed.
4. The flow through fraction could be reloaded to the lectin column. Higher recovery could be achieved while nonspecific binding also increases.
5. The membrane should be stored in electrolyte or distilled water between runs to avoid cracks which cause electrolyte to leak into the focusing chamber.
6. Adding sample slowly until all of the compartments are filled. Aspirating the sample and reload it if bubble exists. Air bubbles will disrupt the electric field and lead to poor separation.
7. Make sure the sample is properly desalted. High ionic strength will interfere with the protein binding. After injecting sample, the pressure may rise rapidly, wait until the pressure decreases and is stable for another injection.
8. Between washing with isopropanol and 1 M NaCl, the column must be equilibrated with water for at least 1 h. NaCl will precipitate in isopropanol which causes the pressure to increase immediately.
9. For lectin-enriched fraction, the sample may need to be concentrated with 10 k MW cutoff centrifuge filter (Millipore) and rediluted with deionized water since the high concentration of sugar in the lectin elution buffer may cause the column pressure to increase rapidly and shorten the column lifetime.
10. It is possible to probe glass slides with both ProQ Diamond dye as well as antibodies. In this case, the slides should be probed with ProQ dye first followed by incubation with the antibodies. Combining both experiments on one slide reduces the number of slides that need to be used as well as provides valuable information about the type of phosphorylation present on the protein.

Acknowledgment

This work was supported in part by the National Cancer Institute under grant R01CA106402 (D.M.L.), the National Institute of

Health under grant R01GM49500 (D.M.L.), and a Michigan Economic Development Grant MEDC03–622 (D.M.S.). Support was also generously provided by Eprogen, Inc. and Beckman-Coulter. We thank Bio-Rad for the gift of the micro-Rotofor device.

References

1. Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA. (2001) Glycosylation and the immune system. *Science* 291: 2370–2376.
2. Block TM, Comunale MA, Lowman M, et al. (2005) Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. *Proc Natl Acad Sci USA* 102: 779–784.
3. Zhao J, Simeone DM, Heidt D, Anderson MA, Lubman DM. (2006) Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: Application to pancreatic cancer serum. *J Proteome Res* 5: 1792–1802.
4. Kobata A, Amano J. (2005) Altered glycosylation of proteins produced by malignant cells, and application for the diagnosis and immunotherapy of tumours. *Immunol Cell Biol* 83: 429–439.
5. Robinson DR, Wu YM, Lin SF. (2000) The protein tyrosine kinase family of the human genome. *Oncogene* 19: 5548–5557.
6. Gschwind A, Fischer OM, Ullrich A. (2004) The discovery of receptor tyrosine kinases: Targets for cancer therapy. *Nat Rev Cancer* 4: 361–370.
7. Cohen P. (1982) The role of protein phosphorylation in neural and hormonal control of cellular activity. *Nature* 296: 613–620.
8. Orchekowski R, Hamelinck D, Li L, et al. (2005) Antibody microarray profiling reveals individual and combined serum proteins associated with pancreatic cancer. *Cancer Res* 65: 11193–11202.
9. Yan F, Sreekumar A, Laxman B, Chinnaiyan AM, Lubman DM, Barder TJ. (2003) Protein microarrays using liquid phase fractionation of cell lysates. *Proteomics* 3: 1228–1235.
10. Patwa TH, Zhao J, Anderson MA, Simeone DM, Lubman DM. (2006) Screening of glycosylation patterns in serum using natural glycoprotein microarrays and multi-lectin fluorescence detection. *Anal Chem* 78: 6411–6421.
11. Zhu K, Zhao J, Lubman DM, Miller FR, Barder TJ. (2005) Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins. *Anal Chem* 77: 2745–2755.
12. Zhao J, Chang AC, Li C, et al. (2007) Comparative proteomic analysis of Barrett's metaplasia and esophageal adenocarcinomas using 2-D liquid mass mapping. *Mol Cell Proteomics* 6: 987–999.
13. Zhao J, Zhu K, Lubman DM, et al. (2006) Proteomic analysis of estrogen response of premalignant human breast cells using a 2-D liquid separation/mass mapping technique. *Proteomics* 6: 3847–3861.
14. Yoo C, Zhao J, Pal M, et al. (2006) Automated integration of monolith-based protein separation with on-plate digestion for mass spectrometric analysis of esophageal adenocarcinoma human epithelial samples. *Electrophoresis* 27: 3643–3651.
15. Pal M, Moffa A, Sreekumar A, et al. (2006) Differential phosphoprotein mapping in cancer cells using protein microarrays produced from 2-D liquid fractionation. *Anal Chem* 78: 702–710.
16. Bradford MM. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254.
17. Rademacher TW, Parekh RB, Dwek RA. (1988) Glycobiology. *Annu Rev Biochem* 57: 785–838.
18. Yang Z, Hancock WS, Chew TR, Bonilla L. (2005) A study of glycoproteins in human serum and plasma reference standards (HUPO) using multilectin affinity chromatography coupled with RPLC-MS/MS. *Proteomics* 5: 3353–3366.
19. Banks JF, Gulcicek EE. (1997) Rapid peptide mapping by reversed-phase liquid chromatography on nonporous silica with on-line electrospray time-of-flight mass spectrometry. *Anal Chem* 69: 3973–3978.
20. Hanson M, Unger KK, Mant CT, Hodges RS. (1996) Optimization strategies in ultrafast reversed-phase chromatography of proteins. *Trans. Anal. Chem.* 15: 102–110.

21. Shibuya N, Goldstein IJ, Broekaert WF, Nsimba-Lubaki M, Peeters B, Peumans WJ. (1987) Fractionation of sialylated oligosaccharides, glycopeptides, and glycoproteins on immobilized elderberry (*Sambucus nigra L.*) bark lectin. *Arch Biochem Biophys* 254: 1–8.
22. Wang WC, Cummings RD. (1988) The immobilized leukoagglutinin from the seeds of Maackia amurensis binds with high affinity to complex-type Asn-linked oligosaccharides containing terminal sialic acid-linked alpha-2,3 to penultimate galactose residues. *J Biol Chem* 263: 4576–4585.
23. Thomsson KA, Karlsson H, Hansson GC. (2000) Sequencing of sulfated oligosaccharides from mucins by liquid chromatography and electrospray ionization tandem mass spectrometry. *Anal Chem* 72: 4543–4549.
24. Kawasaki N, Itoh S, Ohta M, Hayakawa T. (2003) Microanalysis of N-linked oligosaccharides in a glycoprotein by capillary liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. *Anal Biochem* 316: 15–22.
25. Wuhrer M, Deelder AM, Hokke CH. (2005) Protein glycosylation analysis by liquid chromatography-mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 825: 124–133.
26. Harvey DJ, Martin RL, Jackson KA, Sutton CW. (2004) Fragmentation of N-linked glycans with a matrix-assisted laser desorption/ionization ion trap time-of-flight mass spectrometer. *Rapid Commun Mass Spectrom* 18: 2997–3007.
27. Naka R, Kamoda S, Ishizuka A, Kinoshita M, Kakehi K. (2006) Analysis of total N-glycans in cell membrane fractions of cancer cells using a combination of serotonin affinity chromatography and normal phase chromatography. *J Proteome Res* 5: 88–97.
28. Merry AH, Neville DC, Royle L, et al. (2002) Recovery of intact 2-aminobenzamide-labeled O-glycans released from glycoproteins by hydrazinolysis. *Anal Biochem* 304: 91–99.
29. Schulz BL, Packer NH, Karlsson NG. (2002) Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis. *Anal Chem* 74: 6088–6097.
30. Hashii N, Kawasaki N, Itoh S, Hyuga M, Kawanishi T, Hayakawa T. (2005) Glycomic/glycoproteomic analysis by liquid chromatography/mass spectrometry: Analysis of glycan structural alteration in cells. *Proteomics* 5: 4665–4672.
31. Deguchi K, Takegawa Y, Ito H, et al. (2006) Structural assignment of isomeric 2-aminopyridine-derivatized monosialylated biantennary N-linked oligosaccharides using negative-ion multistage tandem mass spectral matching. *Rapid Commun Mass Spectrom* 20: 1480–1481.
32. Deguchi K, Ito H, Takegawa Y, Shinji N, Nakagawa H, Nishimura S. (2006) Complementary structural information of positive- and negative-ion MSn spectra of glycopeptides with neutral and sialylated N-glycans. *Rapid Commun Mass Spectrom* 20: 741–746.
33. Harvey DJ. (1999) Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. *Mass Spectrom Rev* 18: 349–450.
34. Fountain ST, Lee H, Lubman DM. (1994) Ion fragmentation activated by matrix-assisted laser desorption/ionization in an ion-trap/reflectron time-of-flight device. *Rapid Commun Mass Spectrom* 8: 407–416.
35. Ding L, Kawatoh E, Koichi T, Smith AJ, Kumashiro S. (1999) High efficiency MALDI-QIT-ToF mass spectrometer. *Proc Int Soc Optical Eng* 3777: 144.
36. Chien BM, Michael SM, Lubman DM. (1993) Enhancement of resolution in matrix-assisted laser-desorption using an ion-trap stroage reflectron time-of-flight mass spectrometer. *Rapid Commun Mass Spectrom* 7: 837–843.
37. Doroshenko VM, Cotter RJ. (1998) A quadrupole ion trap/time-of-flight mass spectrometer with a parabolic reflectron. *J Mass Spectrom* 33: 305–318.
38. Demelbauer UM, Zehl M, Plematl A, Allmaier G, Rizzi A. (2004) Determination of glycopeptide structures by multistage mass spectrometry with low-energy collision-induced dissociation: Comparison of electrospray ionization quadrupole ion trap and matrix-assisted laser desorption/ionization quadrupole ion trap reflectron time-of-flight approaches. *Rapid Commun Mass Spectrom* 18: 1575–1582.
39. Ojima N, Masuda K, Tanaka K, Nishimura O. (2005) Analysis of neutral oligosaccharides for structural characterization by matrix-assisted laser desorption/ionization quadrupole ion trap time-of-flight mass spectrometry. *J Mass Spectrom* 40: 380–388.

Chapter 21

Transthyretin Mass Determination for Detection of Transthyretin Familial Amyloid

John F. O'Brien and H. Robert Bergen III

Summary

The concentration range of plasma proteins exceeds the dynamic range of any single analytical method. It has been estimated that the concentration range of serum proteins exceeds ten orders of magnitude *(1)*. Because of this, prior immunoselection of even abundant proteins facilitates the relative nonquantitative observations required to show structural abnormality in primary or in posttranslational structure. Determination of atypical proteins by mass measurement has been reported for genetic defects in glycosylation *(2, 3)* and for monitoring for transthyretin (TTR) defects *(4)*. Here we describe a rapid method of purification and electrospray introduction of TTR into a mass spectrometer to detect mass changes due to amino acid substitutions. The method currently forms the basis for a clinical assay to ascertain TTR mutations resulting in amyloidosis.

Key words: Amyloidosis, Transthyretin, Immunopurified, Electrospray mass spectrometry.

1. Introduction

Transthyretin (TTR) is a plasma protein, which circulates as a homotetramer and has unique binding sites for thyroid hormone and retinol-binding protein *(5–7)*. TTR-related familial amyloid polyneuropathy (FAP) is a protein folding disorder, which results in tissue deposition of structurally abnormal TTR monomers *(8)*. A major cause of misfolded TTR monomers is the presence of amino acid substitutions as well as spontaneous misfolding of wild-type TTR (senile systemic amyloidosis) *(9)*. Deposition of the unstable monomer occurs in muscle or nervous tissues forming amyloid precipitates or plaques. The familial form of amyloid involving TTR is dominant and thus normally occurs in individuals

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_21

who have both normal and abnormal copies of the gene *(10, 11)* (*see* **Note 1**).

Though the primary symptoms of FAP are neuropathic, often patients suffer from cardiomyopathy and muscle weakness. Treatment can involve liver transplantation since TTR is synthesized in the liver. More current therapeutic interventions with thyroxine analogues that, like thyroid hormone, bind to and stabilize the tetrameric form are currently in clinical trials *(12–15)* (*see* **Note 2**).

The detection of abnormal TTR monomers in serum forms the basis for diagnosis. This can be done by mass analysis of the monomer, which has a mass of 13,761 Da. Alternatively, direct mutation analysis of the *TTR* gene can be done. Sequencing the entire *TTR* gene is currently favored over identifying specific mutations since the gene comprises only four exons. We have developed a workflow that first determines if the mass of TTR differs from normal by an amino acid substitution, and when such a mass shift is detected, sequencing of the gene is performed *(4)*. Since ~88% of the currently known mutations differ in mass by 10 Da, this is a useful method for the determination of the presence of known or previously unidentified mutations. Additionally, several variants are isobaric and the exact mutation must be determined by subsequent DNA sequence analysis. The mass spectrometry assay of TTR is the subject of the method described herein.

2. Materials

2.1. Equipment

1. Eighteen- or 21-gauge beveled needle, 1 in.
2. Fifteen-milliliter screw cap polypropylene tube.
3. Barnstead/Thermolyne Labquake Shaker/Rotisserie.
4. Vortex-Genie 2 Mixer with 96-microwell plate insert.
5. Beckman/Coulter Allegra 21 Centrifuge with S2096 Microplate rotor.
6. PE Sciex API 150 LC/MS with TurboIonSpray ion source or equivalent.
7. Perkin Elmer Series 200 Autosampler or equivalent.
8. Perkin Elmer Series 200 Micropump *(2)* or equivalent.
9. Shimadzu System Controller, SCL-10Avp.
10. Shimadzu Liquid Chromatography LC-10ADvp pump *(2)* or equivalent.
11. Two Position Micro Electric Actuator and Control Module, Valco Instruments Co. Inc.
12. PEEK tubing, 1/16 in., 0.007-in. I.D. maximum for all connections.

13. Valco Cheminert six-port valve, 1/16-in. fittings, 0.006-in. ports, PN C2–0006E.
14. Upchurch Scientific, Inc. Microfilter, M-560.
15. Phenomenex Security Guard Kit KJO-4282.
16. Phenomenex Widepore C_4 (Butyl), 4-mm H, 2-mm I.D. Cartridge AJO-4329.
17. Isolute Array Base Plate, P/N 120-1000-P1; Argonaut Technologies, Inc.
18. Isolute Array 10-μm Fritted Reservoirs, P/N 120-1040-R; Argonaut Technologies, Inc.
19. Microliter Analytical Supplies, 96-well PP 350-μL plate PN 07–1100.
20. Microliter Analytical Supplies, 96 Position Round Well Pierceable Cover, PP, PN 07-0011N.
21. Affinity resin (refer to TTR affinity resin preparation procedure).

2.2. Immunoaffinity Purification: Reagents

1. Rabbit Anti-Human Prealbumin (TTR; cat. # A0002), Dako North America, Inc., Carpinteria, CA.
2. Poros AL (Aldehyde) Self Pack Media (cat. # PN 2-3115-00), Applied Biosystems, Foster City, CA.
3. Pierce Slide-A-Lyzer 3.5 kDa, 3 mL (cat. # PN 66330), Pierce, Rockford, IL.
4. 100 mM phosphate buffer, pH 7.4: In a 1-L beaker, add 2.6 g of $NaH_2PO_4.H_2O$ and 11.5 g of Na_2HPO_4. The pH will be about 7.4. Adjust the pH to 7.4 with 1N NaOH or 1 M HCl. Stable for 6 months at 4°C.
5. *Phosphate buffer saline (PBS)*. 10 mM phosphate, 154 mM NaCl, pH 7.4. In a 1-L volumetric flask, dissolve 9.0 g of NaCl in approximately 800 mL of RO water. Add 100 mL of 100 mM phosphate, pH 7.4 (reagent 4). Adjust the pH to 7.4 with 1N NaOH or 1 M HCl and dilute to 1 L. Stable for 6 months at 4°C.
6. *High-salt buffer*. 1.5 M Na_2SO_4, 100 mM phosphate, pH 7.4. In a 50-mL polypropylene screw cap tube, dissolve 10.6 g of Na_2SO_4 in 50 mL of 100 mM phosphate buffer (reagent 4). This solution is stable for 6 months at room temperature (*see* **Note 3**).
7. *Sodium cyanoborohydride ($NaCNBH_3$)*. 100 mg/mL. Transfer approximately 75–100 mg of $NaCNBH_3$ to a 1.5 or 2 mL tared screw cap or microcentrifuge tube in a chemical hood. Reweigh the tube and determine the mass of $NaCNBH_3$. Add RO water to make a 100 mg/mL solution. N.B. Cyanoborohydride is toxic! Skin contact and inhalation should be avoided.

8. *Capping buffer*. 0.2 M Tris, pH 7.2. In a 1-L beaker, dissolve 24.2 g of Tris base in 800 mL of RO water. Adjust the pH to 7.2 with concentrated HCl and dilute to 1 L. Stable for 3 months when stored at 4°C.
9. *Sodium azide (NaN_3)*. 20% solution. Prepare a 20% wt/vol solution by measuring 100–200 mg of NaN_3 into a tared 1.5-mL eppendorf tube. Calculate the appropriate amount of water to make a 20% solution (e.g., if the total amount of NaN_3 measured is 150 mg, 0.15/0.2 = 0.75 mL of water is added).

2.3. Immunoaffinity Purification: Antibody–Bead Coupling

Day 1. Solutions 4 and 5 are necessary for the dialysis step of the procedure. It is not necessary to make up other reagents until the next day, after overnight dialysis.

Day 2. Solutions 6 and 7 are necessary for coupling antibody to the beads. It is not necessary to make up other reagents until the next day if the bead binding solution is going to be agitated over night.

Day 3. Solutions 8 and 9 need only be made up prior to the final steps of the antibody binding process.

1. Dialysis is required to remove the azide preservative that Dako includes with the anti-TTR antibody. Transfer 2 mL of the rabbit anti-human TTR (~8.2-mg protein) (reagent 1) to a 3-mL Slide-A-Lyzer (3.5 kDa) unit (reagent 3) with a 1-in. 18- or 21-guage beveled needle by carefully following the manufacturers' enclosed instructions or go to their web site for current protocols (http://www.piercenet.com). Do not allow the needle to contact the membrane. This is accomplished by inserting the needle such that the needle opening is facing sideways to the slide. With the needle facing up draw excess air from the Slide-A-Lyzer. Invert the Slide-A-Lyzer and syringe such that the syringe is facing down. Inject the antibody solution into the Slide-A-Lyzer. Use the same syringe to remove as much air as possible from the Slide. Never use a Slide-A-Lyzer numbered hole more than once.
2. Dialyze against 1 L of PBS (reagent 5) for 2 h at 4°C. Discard the PBS and dialyze for 12–24 h (overnight) against 1 L of fresh PBS (reagent 5) at 4°C temperature. Using a syringe, recover the dialyzed antibody solution. *Record the total volume recovered as V_i*. Usually 1.8–2 mL is recovered. Place recovered antibody solution in a 15-mL polypropylene screw cap tube.
3. If some of the solution is lost, adjust the amount of Poros to give 1-mg protein/50-mg Poros. Calculate the amount of resin to use by using the formula:
4. $(V_i \times 410\ mg)/2\ mL$ = mg of Poros AL to use.

5. Add the calculated amount of Poros AL resin (reagent 2) to the tube containing the antibody solution. Avoid inhalation of resin.
6. Calculate the amount of 100 mg/mL $NaCNBH_3$ (reagent 7) to add to give a final concentration of 5 mg/mL after the high-salt buffer has been added, e.g.,

$$\left(\frac{(3.75V_i)\cdot\ 5\,\text{mg/mL}}{100\,\text{mg/mL}}\right) = \text{Volume of 100 mg/mL NaCNBH3 to add.}$$

   The 3.75 V_i value comes from the initial volume (V_i) plus the volume of high-salt buffer you will add below (V_t), which will be calculated as $V_t = 2.75 \times V_i$. Thus $V_i + 2.75\,V_i = 3.75\,V_i$.
7. Add the volume of $NaCNBH_3$ (reagent 7) calculated in **step 4** to the antibody/resin solution.
8. Calculate the total amount of high-salt buffer (reagent 6) to add.
9. $V_t = 2.75\,V_i$, where V_t is the total volume of high-salt buffer to add and V_i is the total volume of antibody recovered from the Slide-A-Lyzer.
10. For example, if 2 mL of antibody solution is recovered, the total volume of high-salt buffer to add is 2 mL × 2.75 = 5.5 mL of high-salt buffer.
11. The final concentration desired is 1.1 M Na_2SO_4 after addition of high-salt buffer.
12. Add 1 mL of high-salt buffer (reagent 6) in 5-min intervals until V_t of high-salt buffer has been added. The solution should be rocked between additions and then allow the solution to rock/rotate for 90 min or overnight.
13. Centrifuge the solution at 2,000 × *g* for 2 min and using a syringe, aspirate the bottom layer of buffer taking care not to remove the resin that will be floating on top. Discard the aspirate.
14. Suspend the resin in 10 mL of PBS (reagent 5). Suspend gently and centrifuge at 2,000 × *g* for 2 min. The resin should now pelletize. Remove the supernatant.
15. Suspend the media in 5 mL of capping buffer (reagent 8) and add 0.25 mL of 100 mg/mL $NaCNBH_3$ (reagent 7). Rock/rotate gently for 1 h.
16. Centrifuge the solution at 2,000 × g for 2 min and remove the supernatant. Wash the resin twice with 10 mL of PBS (reagent 5) and remove the supernatant with centrifugation as before.
17. Add 10 mL PBS (reagent 5) and add 10 μL of 20% NaN_3 (reagent 9) to bring the final NaN_3 concentration to 0.02% NaN_3. Store at 4°C.

2.4. Immunoaffinity Purification: Reagents

1. *Binding/wash buffer*. PBS, 10 mM phosphate, 154 mM NaCl pH 7.4. In a 1-L beaker, dissolve 9.0 g of NaCl in approximately 900 mL of RO water. Add 0.23 g of $NaH_2PO_4{\cdot}H_2O$ and 1.15 g of Na_2HPO_4. The pH will be about 7.4. Adjust the pH to 7.4 with 1N NaOH or 1 M HCl. Adjust the final volume to 1 L. Stable for 6 months at 4°C.
2. *Elution buffer*. 100 mM glycine, 2% acetic acid, pH 2.5. In a 1-L beaker, dissolve 7.5 g of glycine in 800 mL of RO water. Slowly add 20 mL of glacial acetic acid and mix. The pH will be about 4.0. Adjust the pH to 2.5 with about 5 mL of concentrated HCl. Bring the final volume to 1 L with RO water. Stable for 6 months at 4°C.

2.5. Reverse Phase Chromatography

1. *Aqueous mobile phase (gradient pump A)*. 1% Acetic acid/methanol/acetonitrile (98/1/1). In a 500-mL reagent bottle, add 5 mL of acetic acid, 5 mL of methanol, and 5 mL of acetonitrile to 490 mL of RO water. Mix well. Make fresh weekly.
2. *Organic mobile phase (gradient pump B)*. 0.5% Acetic acid–0.02% trifluroacetic acid/methanol/acetonitrile (5/48/48). In a 500 mL reagent bottle, add 2.5 mL of acetic acid, 200 μL of TFA, 240 mL of methanol, and 240 mL of acetonitrile to 25 mL of RO water. Mix well. Make fresh weekly.
3. *Loading phase (loading pump)*. Water/methanol/trifluoracetic acid (90/10/0.1). In a 500-mL reagent bottle, add 445 mL of RO water, 50 mL of methanol, and 500 μL of trifluoroacetic acid. Stable for 6 months at 4°C.

TTR is purified from plasma using affinity purification by incubating plasma with an immunoaffinity resin prepared earlier. After a 1-h incubation of plasma with anti-TTR resin in a 96-well plate (with agitation), the resin is washed free of nonbound plasma components with PBS in a 96-well filter plate. Purified TTR is eluted off the anti-TTR resin with 100 mM glycine of pH 2.5. The purified TTR is reduced with Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) to simplify the mass spectra by removing Cys10 adducted species. After a 15-min incubation, the sample is injected and trapped on a 2 × 4 mm C_4 trap at 50 μL/min, washed free of glycine for 3 min, and eluted into the mass spectrometer at 50 μL/min with an organic gradient. After deconvolution, normal patients present with a single peak corresponding to wild-type TTR, which serves as a reference. FAP patients are typically heterozygous and are detected by the presence of two peaks (i.e., wild-type TTR and mutant TTR) differing from each other in mass.

2.6. Specimens

Serum and plasma are both acceptable sample matrices. Both are stable for 7 days at 4°C or indefinitely at −80°C.

2.7. Additional Reagents

1. *Myoglobin Standard Concentrated, 10 pmol/µL.* Prepare a concentrated stock solution of myoglobin (10 pmole/µL) by dissolving 1.7 mg of horse heart myoglobin in 10 mL of RO water. Aliquot the myoglobin solution into 50-µL aliquots and store it in frozen state.
2. *TCEP 100 mM.* To a 1-mL vial add 28.6 mg of TCEP. Add 1 mL of RO water for a final concentration of 100 mM. Stable for 2 weeks in frozen state. Measured amount of TCEP can be plus/minus about 0.5 mg as the amount of reducing agent (TCEP) is in excess.

3. Methods

3.1. Sample Processing

1. Whole blood in a 3-mL ACD or EDTA tube should arrive in the lab refrigerated at 4°C. Aliquot about 1.5 mL of well-mixed blood to a tube marked Molecular Genetics tube (*see* Note 4) and store this aliquot in the freezer for future genetic analysis if necessary. The remaining blood is centrifuged and the plasma aliquoted for MS testing.
2. If the MS test is abnormal the Molecular Genetics test will automatically be ordered.

3.2. Valves and MS Configuration

The valving configuration is shown in Fig. 1. A loading pump (water/methanol/trifluoroacetic acid 90/10/0.1) is plumbed into the autosampler and then a six-port two-position valve at 50 µL/min. The injected sample is trapped on a six-port valve

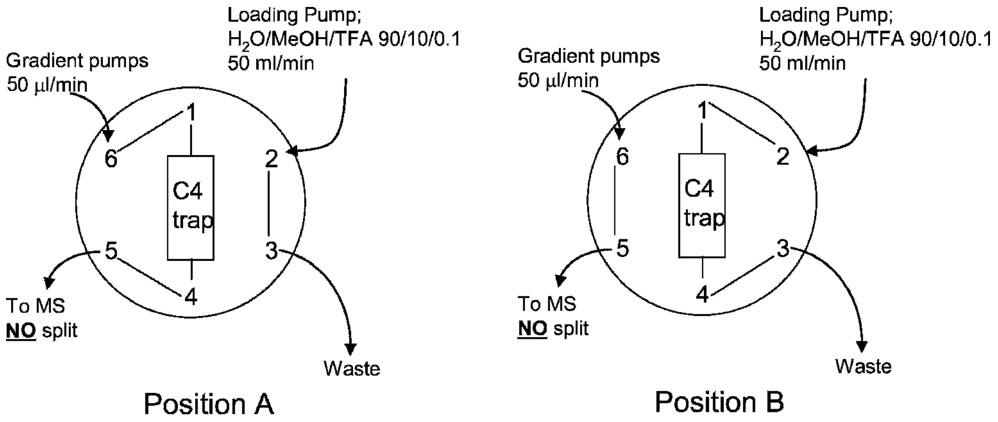

Loading Pump;
H_2O/MeOH/TFA 90/10/0.1
50 ml/min
Gradient pumps
50 µl/min
1
2
3
4
5
6
C4 trap
To MS
NO split
Waste
Position A
Loading Pump;
H_2O/MeOH/TFA 90/10/0.1
50 ml/min
Gradient pumps
50 µl/min
1
2
3
4
5
6
C4 trap
To MS
NO split
Waste
Position B

Fig. 1. A six-port valve is plumbed with the C4 trap across ports 1 and 4. The loading pump is plumbed with the autosampler and samples are trapped and desalted in position "B." After 3 min of washing, the valve switches to position "A" and an organic gradient elutes the desalted TTR into the mass spectrometer.

with the C_4 trap in valve position "B" (**Fig. 1**) and desalted (3 min) prior to the valve switching to position "A". After switching to position "A", an organic gradient (0–3 min 0% B; 3–6 min 0–75% B; 6–12 min 75% B, 12–14 min 75–95% B; 14–16 min 95–0% B) at 50 µL/min elutes the desalted TTR into the MS system without splitting.

The mass spectrometer electrospray source is set to 5,500 V with a desolvation temperature of 120°C. Orifice potential was 51 V. A scan rate of 1.7 s/scan over the mass range 800–2,500 Da was utilized. Scans containing TTR are summed and transformed onto a relative mass scale.

3.3. Mass Spectrometer Quality Control

A myoglobin standard is utilized to ascertain LC–MS system status.

1. Obtain one vial of Myoglobin Standard Concentrate (reagent 1) containing 50 µL of 10 pmole/µL myoglobin. Add 950-µL RO Water and mix for an operating concentration of 0.5 pmole/µL. Inject 45 µL of the 0.5 pmol/µL myoglobin standard. Make at least four injections of myoglobin standard (two from well 95 and two from well 96) or until chromatography is clean and good data for myoglobin are obtained. Analyze the summed spectra of the eluting myoglobin.
2. Verify that the myoglobin multiply charged spectra looks comparable to a reference myoglobin spectra. A proper spectrum indicates that the LC–MS system is working properly.
3. Wild-type TTR and TTR isolated from a control volunteer with the polymorphic variant (Gly6Ser +30 Da) are analyzed with the patient samples. Absolute mass of the wild-type peak (current working mean is 13,763 ± 2 Da; Swiss-Prot mass for TRR is 13,761 Da) and peak width at half height (10.5 Da, s.d. = 1.3 Da) are useful quality control parameters since mass shifts such as that seen in the homozygous V122I (+14) variant or small mass shifts without complete resolution from wild-type TTR can be detected using these parameters *(3)*.

3.4. Patient Samples

1. For each sample, add 100 µL of PBS (reagent 2) and 40 µL of anti-TTR resin suspension to a well in a 96-well polypropylene sample plate (*see* **Note 5**).
2. Transfer 20 µL of plasma to be analyzed to the prepared wells of the sample plate.
3. Cover the sample plate with parafilm or plate cover and shake on vortex-genie 2 for 1 h with sufficient motion (setting on mixer of 2–3) to maintain a suspension of the resin in the sample.
4. Fit enough filter reservoirs to the filter base plate assembly (filter plate) to accommodate the patient samples (not wells 95 and 96) and place the assembly on the wash plate.

5. After 1-h incubation, transfer each sample from the sample plate to a separate 10-µm fritted reservoir on the filter plate and centrifuge the filter plate fitted with the wash plate with an appropriate counter balance (2,000 rpm, 2 min, and acceleration factor 5) (*see* **Note 6**).
6. Discard the filtrate in the wash plate. Blot the tips of the fritted reservoirs gently on paper towel to remove any adhered liquid and replace on the wash plate.
7. Wash the resin with 2 × 200 µL of PBS (reagent 2) and discard the filtrate after each wash (2,000 rpm, 2 min, and acceleration factor 5).
8. Blot the tips of the fritted reservoirs gently on a paper towel after each wash cycle and discard the wash plate after the last wash.
9. Prepare the autosampler plate by adding 10 µL of 100 mM TCEP to each of the sample wells (not wells 95 and 96) and fit the filter plate to the autosampler plate (*see* **Note 7**).
10. Add 55 µL of 100 mM glycine, pH 2.5 (reagent 3) to each fritted reservoir and resuspend the resin by repeated aspiration with a pipette.
11. Wait for 5–10 min and then centrifuge the filter plate and the autosampler plate (2,000 rpm, 2 min, and acceleration factor 5).
12. Cover the autosampler plate with round well pierceable covers and incubate for 15 min at 55°C.
13. The sample is now ready for analysis. Inject 45 µL via autosampler.
14. Refer to **Subheading 3.2** for instrument parameters.
15. After analysis is complete, sum the scans in the center of the eluted TTR and transform multiply charged spectra with BioSpec Reconstruct centering the envelope over the $(M + 8H^+)^{+8}$ charge state (1,721 Da) and a peak width of 800 Da. An output start and stop mass of 13,000 and 14,500 Da are used, respectively, with a 0.1-Da step size over five iterations.
16. Appearance of peaks within ± 265 Da from wtTTR at 30% of the peak at 13,761 Da in the deconvoluted spectra is marked as positive for TTR mutation.

3.5. Reporting Results/Interpreting Results

Figure 2 shows the results from seven plasma samples – six positive and one negative for the presence of a TTR variant. The exact variant is determined by genetic analysis and is shown for each of the positive samples *(4)* (*see* **Note 8**).

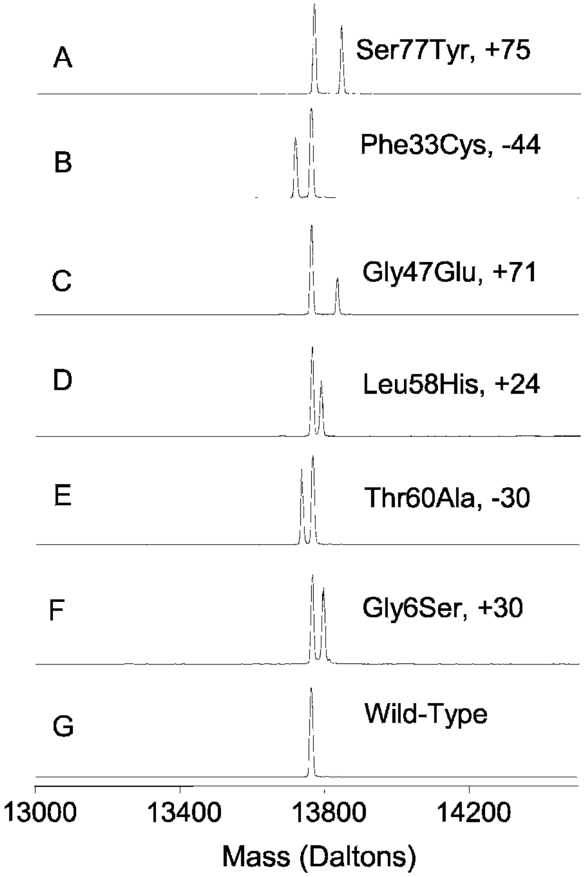


Fig. 2. Transformed spectra of seven plasma samples. Six of these samples are heterozygous for a TTR variant and are easily identified by an additional peak adjacent to the wild-type peak at 13,731 Da. The wild-type sample (G) is homozygous for wild-type TTR (reprinted with permission, ©American Association of Clinical Chemistry, from ref. 25).

4. Notes

1. *Expression levels.* Both *TTR* genes are expressed and both a wtTTR and mutant TTR are usually evident upon analysis in affected patients. Additionally, a low variant to wild-type ratio exists in most samples *(16, 17)*. The mass at 13,761 has always been more pronounced than mutant TTR in all analyses to date. One exception being the Gly6Ser +30 polymorphism in which both the 13,761 and 13,791 Da peaks are comparable in intensity. This lends weight to the determination that this is a benign polymorphism. DNA sequencing is suggested nonetheless.
2. *Clinical significance.* Familial amyloidotic polyneuropathy (FAP) is an inherited systemic disease in which protein deposits ultimately lead to organ failure(s) and death *(9)*. One of the major clinical manifestations is some type of peripheral neuropathy. FAP is associated with genetic variants of TTR in which

single amino acid substitutions have been found in the primary structure *(9)*. TTR is a plasma transport protein that circulates as a tetramer with a monomeric mass of 13,761 Da containing 127 amino acids *(18)*. A recent count indicates 90 genetic variants, the vast majority being amyloidogenic *(19–21)*.

The capability to quickly and easily identify the presence of a TTR genetic variant by an online analysis should not be overlooked as a diagnostic tool for FAP. Early diagnosis of the disease can contribute to the patient's survival chances. Currently, the major therapeutic is liver transplantation. There is evidence that livers transplanted shortly after the disease onset give better outcomes *(22–24)*.

3. Na_2SO_4 may crystallize out of solution in which case the solution may be heated or agitated to redissolve the Na_2SO_4.
4. If clinical diagnosis is strongly suspected or there is a positive family history, molecular sequencing is recommended to rule out the possibility of any false-negative result.
5. You must vortex the anti-TTR resin intermittently while pipetting to maintain a homogenous suspension, e.g., between every two additions.
6. It is important to use an acceleration factor 5 to prevent resin passing through the filter. Utilizing higher acceleration factors has been shown to pass resin through the filter.
7. *Removal of Cys10 adducts.* TTR has one cysteine residue at position 10. This cysteine residue exists without adduction but is predominately found adducted in a disulfide linkage with cysteine, glutathione, cys–gly, or S-sulfonated (i.e., Cys–S–SO_3H). These adducts complicate the analysis and thus the analysis is simplified by a simple reduction step prior to analysis, thus the TCEP reduction. An additional benefit is the increased signal/noise ratio resulting from collapsing the signal from the adducted ions into free TTR. Persistence of a significant +79-Da peak may suggest underreduction of the S-sulfonated species and may require doubling of the amount of TCEP and extending the time for reduction. If the +79-Da peak is not reduced following rigorous reduction, the presence of a variant is indicated.
8. Results are reported as negative, positive, or indeterminate. (a) *Negative.* Normal patient samples will be those that are determined to contain a single peak at 13,761 Da. This finding does not exclude the diagnosis of FAP or the presence of TTR mutations. Resolution of TTR variants from wtTTR requires that the mutant TTR differs in mass from wtTTR by at least 10 Da when analyzed by a unit resolution MS system. Mutations are known to exist that confer mass differences of less than 10 Da and one (Leu to Ile) is isobaric. Additionally, several examples exist of double mutations or homozygous mutations that give

a single peak or small mass differences that this assay will not detect (e.g., a cis −28 and +30 with an overall mass difference of +2 would not be detected). This fact should be included as a disclaimer in all reporting. (b) *Positive*. The presence of a second peak at or near 13,761 Da is indicative of a TTR variant. The largest mass shift for a single mutation is 129 Da for a Gly to Trp or Trp to Gly. Larger shifts would accompany double mutations. Looking at masses 265 Da from 13,761 Da would take into account the largest possible shifts including a double mutation. The signal from mutant TTR will most likely have a smaller signal than the wtTTR. DNA sequence analyses of the three exons associated with TTR are recommended for abnormal reports to determine the exact mutation. Possible mutations can be suggested by utilizing mutation mass shift tables (e.g., http://prospector.ucsf.edu/prospector/4.0.8/html/misc/mutation.htm). There are also examples of patients who are homozygous for the same mutation. One recent example is a V122I (+14 mass shift) mutation that is present in both alleles giving rise to a single peak at 13,775 Da, which is shifted up 14 Da. Technicians should read masses carefully to avoid mistaking the homozygous variant for wild-type TTR. (c) *Indeterminate*. If the peak shape is broader than normal, it is possible that a mutation exists with a small mass that is not completely resolved. DNA sequence analysis of the three TTR exons is suggested.

The diagnosis of FAP should be considered in all patients with symptoms beginning in the third to seventh decade of life. These symptoms include peripheral neuropathy, autonomic neuropathy, and varying degrees of systemic amyloid involvement. Other clinical manifestations of hereditary amyloidosis depend on which organ systems are involved.

References

1. Anderson, N. L. and Anderson, N. G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics. 1*, 845–867
2. Lacey, J. M., Bergen, H. R., Magera, M. J., Naylor, S. and O'Brien, J. F. (2001) Rapid determination of transferrin isoforms by immunoaffinity liquid chromatography and electrospray mass spectrometry. *Clin Chem. 47*, 513–518
3. Bergen, H. R., Lacey, J. M., O'Brien, J. F. and Naylor, S. (2001) Online single-step analysis of blood proteins: the transferrin story. *Anal Biochem. 296*, 122–129
4. Bergen, H. R., III, Zeldenrust, S. R., Butz, M. L., Snow, D. S., Dyck, P. J., Klein, C. J., O'Brien, J. F., Thibodeau, S. N. and Muddiman, D. C. (2004) Identification of transthyretin variants by sequential proteomic and genomic analysis. *Clin Chem. 50*, 1544–1552
5. Blake, C. C., Geisow, M. J., Oatley, S. J., Rerat, B. and Rerat, C. (1978) Structure of prealbumin: secondary, tertiary and quaternary interactions determined by Fourier refinement at 1.8 A. *J Mol Biol. 121*, 339–356
6. Monaco, H. L., Rizzi, M. and Coda, A. (1995) Structure of a complex of two plasma proteins: transthyretin and retinol-binding protein. *Science. 268*, 1039–1041
7. Nilsson, S. F., Rask, L. and Peterson, P. A. (1975) Studies on thyroid hormone-binding proteins. II. Binding of thyroid hormones,

retinol-binding protein, and fluorescent probes to prealbumin and effects of thyroxine on prealbumin subunit self association. *J Biol Chem. 250*, 8554–8563

8. Sipe, J. D. (1994) Amyloidosis. *Crit Rev Clin Lab Sci. 31*, 325–354
9. Benson, M. (1995) Amyloidosis In *The Metabolic and Molecular Bases of Inherited Disease* (D., ed) pp. Valle, 4159–4191, McGraw-Hill, New York
10. Benson, M. D. and Dwulet, F. E. (1985) Identification of carriers of a variant plasma prealbumin (transthyretin) associated with familial amyloidotic polyneuropathy type I. *J Clin Invest. 75*, 71–75
11. Dwulet, F. E. and Benson, M. D. (1984) Primary structure of an amyloid prealbumin and its plasma precursor in a heredofamilial polyneuropathy of Swedish origin. *Proc Natl Acad Sci U S A. 81*, 694–698
12. Sekijima, Y., Dendle, M. A. and Kelly, J. W. (2006) Orally administered diflunisal stabilizes transthyretin against dissociation required for amyloidogenesis. *Amyloid. 13*, 236–249
13. Tojo, K., Sekijima, Y., Kelly, J. W. and Ikeda, S. (2006) Diflunisal stabilizes familial amyloid polyneuropathy-associated transthyretin variant tetramers in serum against dissociation required for amyloidogenesis. *Neurosci Res. 56*, 441–449
14. Miller, S. R., Sekijima, Y. and Kelly, J. W. (2004) Native state stabilization by NSAIDs inhibits transthyretin amyloidogenesis from the most common familial disease variants. *Lab Invest. 84*, 545–552
15. Klabunde, T., Petrassi, H. M., Oza, V. B., Raman, P., Kelly, J. W. and Sacchettini, J. C. (2000) Rational design of potent human transthyretin amyloid disease inhibitors. *Nat Struct Biol. 7*, 312–321
16. Kishikawa, M., Nakanishi, T., Miyazaki, A., Shimizu, A., Nakazato, M., Kangawa, K. and Matsuo, H. (1996) Simple detection of abnormal serum transthyretin from patients with familial amyloidotic polyneuropathy by high-performance liquid chromatography/ electrospray ionization mass spectrometry using material precipitated with specific antiserum. *J Mass Spectrom. 31*, 112–114
17. Theberge, R., Connors, L., Skinner, M., Skare, J. and Costello, C. E. (1999) Characterization of transthyretin mutants from serum using immunoprecipitation, HPLC/ electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem. 71*, 452–459
18. Kanda, Y., Goodman, D. S., Canfield, R. E. and Morgan, F. J. (1974) The amino acid sequence of human plasma prealbumin. *J Biol Chem. 249*, 6796–6805
19. Saraiva, M. J. (2001) Transthyretin mutations in hyperthyroxinemia and amyloid diseases. *Hum Mutat. 17*, 493–503
20. Theberge, R., Connors, L. H., Skinner, M. and Costello, C. E. (2000) Detection of transthyretin variants using immunoprecipitation and matrix-assisted laser desorption/ ionization bioreactive probes: a clinical application of mass spectrometry. *J Am Soc Mass Spectrom. 11*, 172–175
21. Connors, L. H., Lim, A., Prokaeva, T., Roskens, V. A. and Costello, C. E. (2003) Tabulation of human transthyretin (TTR) variants, 2003. *Amyloid Int J Clin Exp Invest. 10*, 160–184
22. Ando, Y., Tashima, K., Tanaka, Y., Nakazato, M., Ericzon, B. G., Duraj, F. F., Sakashita, N., Kimura, E., Ando, E., Yonehara, T.et-al. and (1994) Treatment of a Japanese patient with familial amyloidotic polyneuropathy with orthotopic liver transplantation. *Intern Med. 33*, 730–732
23. Bergethon, P. R., Sabin, T. D., Lewis, D., Simms, R. W., Cohen, A. S. and Skinner, M. (1996) Improvement in the polyneuropathy associated with familial amyloid polyneuropathy after liver transplantation. *Neurology. 47*, 944–951
24. Suhr, O. B., Holmgren, G., Steen, L., Wikstrom, L., Norden, G., Friman, S., Duraj, F. F., Groth, C. G. and Ericzon, B. G. (1995) Liver transplantation in familial amyloidotic polyneuropathy. Follow-up of the first 20 Swedish patients. *Transplantation. 60*, 933–938

Chapter 22

Characterization of Microorganisms by MALDI Mass Spectrometry

Catherine E. Petersen, Nancy B. Valentine, and Karen L. Wahl

Summary

Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) for characterization and analysis of microorganisms, specifically bacteria, is described here as a rapid screening tool. The objective of this technique is not comprehensive protein analysis of a microorganism but rather a rapid screening of the organism and the accessible protein pattern for characterization and distinction. This method is based on the ionization of the readily accessible and easily ionizable portion of the protein profile of an organism that is often characteristic of different bacterial species. The utility of this screening approach is yet to reach its full potential but could be applied to food safety, disease outbreak monitoring in hospitals, culture stock integrity and verification, microbial forensics, or homeland security applications.

Key words: Bacteria, Spores, Microorganisms, Mass spectrometry, MALDI, Proteins, Fingerprinting.

1. Introduction

There is an ever-increasing need for the consistent and rapid identification of intact microorganisms. Mass spectrometry is a powerful analytical tool that can also be used for screening microorganisms rapidly. Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) has been used to identify microorganisms based on expressed protein profiles. MALDI-TOF MS provides rapid analysis time (< 1 min/sample analysis), has low sample volume requirements (<1 µL of fluid), and yields very specific and unbiased analysis based on the molecular weights of true components of the sample. The high sensitivity and high tolerance toward contaminants have made

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_22

MALDI-TOF MS a viable technique for the analysis of complex biological samples.

MALDI-TOF MS has established itself as an analytical technique having the ability to identify bacteria to the species level in pure cultures and simple mixtures of bacteria *(1)*. It can be used as a rapid screening tool in distinguishing between pathogenic and nonpathogenic species, which potentially makes it a powerful tool in countering terrorism. MALDI-TOF MS has also been successfully extended to the identification of sporulated varieties of Gram-positive strains of *Bacillus* with modification to the sample preparation techniques *(2)*.

The screening of bacterial samples by mass spectrometry for quick identification can be accomplished with direct analysis of a subset of the protein profile by MALDI-TOF MS analysis. While MALDI-TOF MS is known to have a wide dynamic range of analysis it is also a competitive ionization method and therefore may not yield a profile representative of all components present in the sample. The chemical complexity in most vegetative bacterial samples along with the many orders of magnitude difference in concentration often results in only a minor number of cell components observed by direct MALDI-TOF MS analysis of the entire cell. However, this only presents a significant challenge if the desired result is the complete profiling of the cell contents. A reproducible pattern of putative proteins is readily observed from simple sample preparation and analysis of bacterial cultures *(3, 4)*. The sample preparation and analysis method described here was developed as a rapid screening tool for identification of microorganisms without individual protein identification *(5, 6)*. While relatively straightforward, there are a few steps that are helpful to follow in order to successfully obtain data from microorganisms directly by MALDI-TOF MS. The purpose of this chapter is to provide some guidance on one possible approach to analysis of microorganisms by MALDI-TOF MS. This is not the only method available but is similar to many other published methods and approaches *(7, 8)* and commercial protocols (e.g., Micromass, Bruker) as well. Two review articles *(7, 8)* provide a good overview of this research field as well as a book devoted to this subject of microorganisms analysis by mass spectrometry that provides additional information and approaches to the analysis of these complex biological samples *(9)*.

Several steps are important for successful MALDI-TOF MS analysis of microorganisms (specifically bacteria). The first step is obtaining relevant microorganism samples for analysis. While only small volumes (μL) of sample are required for MALDI-TOF MS analysis, currently at least 10^3 cells are necessary. Analyses of approximately 10^6 cells per microliter are used more routinely for this direct MALDI-MS analysis. The relative concentrations of matrix and analyte are critical for successful analysis and will be discussed in more detail later. Removal of the growth media from the microorganisms is also important for successful mass spectrometric analysis. A proven and effective way to clean the bacterial samples harvested from liquid growth media is to pellet the cells and wash with water or appropriate volatile solution, such as 2% ammonium chloride, repeatedly. Once adequate bacterial samples are prepared/obtained, the second step is to spot the microbial sample (*see* **Note 1**) onto a MALDI sample plate along with a MALDI matrix compound, which is used to aid in desorption and ionization of intact protein and nonvolatile components within the bacterial cells. There are numerous matrix recipes and spotting procedures in the literature for successful MALDI-MS analysis that are for the most part applicable to analysis of components of microorganisms. The research community currently performing bacterial analysis does not consistently use the same sample matrix and spotting procedures. Provided here is the method we have optimized in our laboratory and have successfully applied to analysis and characterization of vegetative and sporulated bacteria. We have also analyzed fungi with the use of double-stick tape for sampling directly from a fungal colony and applying to the MALDI sample plate *(10)*. After the sample is effectively spotted onto a commercial MALDI sample plate, the sample is analyzed with straightforward instrument parameters. Note that the instrument needs to be well calibrated prior to data collection as with any analysis. The final step is data processing and analysis. PNNL has developed algorithms for justified comparison of sample data with a collected database of MALDI-TOF MS spectra for comparative identification. There are other commercial products available that are designed for use with specific vendor instrumentation and other published approaches to data analysis *(9)*.

Mass spectra can be obtained from unknown bacterial samples and compared with reference spectra to provide information for a correct identification with a computed degree of confidence. Well-defined sample preparation procedures and calibration routines must be adhered to during the collection of replicates for reference spectra. Protein profiles, or fingerprints, obtained from microorganisms by means of MALDI-TOF MS can vary in connection with the choice of solvent system for the matrices *(11)*. Sufficient replication is also necessary to capture minor variations in growth or handling procedures (*see* **Notes 2** and **3**). Thereafter, automated statistics-based data analysis algorithms can be used to compile these replicates into a reference spectrum for inclusion into a database *(6)*. This process has also been successfully extended for the identification of microorganisms grown under different growth conditions making it a viable tool in terms of attribution in forensics *(12)*.

Matrix-assisted laser desorption ionization as its name suggests depends on the cocrystallization of the analyte material and matrix molecules *(13, 14)*. A variety of matrices are compatible for use in instruments equipped with a nitrogen laser (337 nm). The ionization process is well suited for mass spectrometric analysis of large biomolecules. The analyte substance is imbedded in the crystallized matrix and irradiated by sufficient laser power to assist in ionization of intact molecules but does not result in fragmentation. Thus, MALDI is often referred to as a *soft* ionization technique. MALDI matrices have conjugated ring structures and typically differ only in their attached moieties. Structures for three of the most commonly used matrices for vegetative and sporulated bacteria are shown in **Fig. 1**.

Each matrix has a unique initial velocity when exposed to a pulsed laser beam under vacuum and as such can provide advantages when interrogating different mass ranges of interest. For example, alpha-cyano-4-hydroxy cinnamic acid (ACHC), often referred to as a hot matrix, has a high initial velocity, and is a preferred matrix for smaller molecules. Sinapinic acid (SA) or 3,5-dimethyoxy-4-hydroxy cinnamic acid and 3-methoxy-4-hydroxy cinnamic acid, commonly referred to as ferulic acid (FA), are more suited for use with most intact vegetative bacterial cells as well as bacterial spore preparations. Typical mass ranges for the ions observed by direct MALDI-MS analysis of the bacterial cells with the protocols described here are 2–20 kDa. A schematic of the preparation of the MALDI sample spot and a scanning electron microscopy image of a portion of the sample spot is provided in **Fig. 2**.

Once the matrix has been selected, an appropriate solvent system must also be determined. Matrices will be dissolved, usually to the point of saturation, in some ratio of H_2O and organic material. The solvent system is a crucial component of sample preparation and has a major impact both on crystal structure and the degree to which the analytes are incorporated into them.

For the analysis of most readily accessible and ionizable components from intact bacterial cells or spores it is desirable to have the solvent system at a low pH, which is accomplished

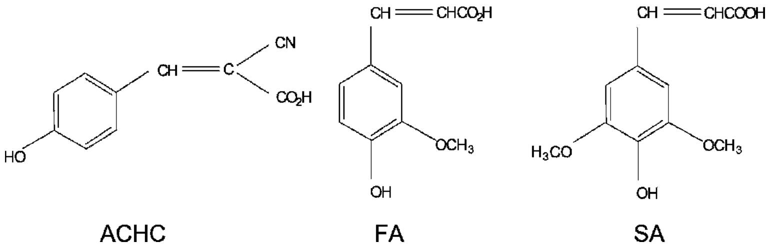


Fig. 1. Structures of common MALDI matrices.

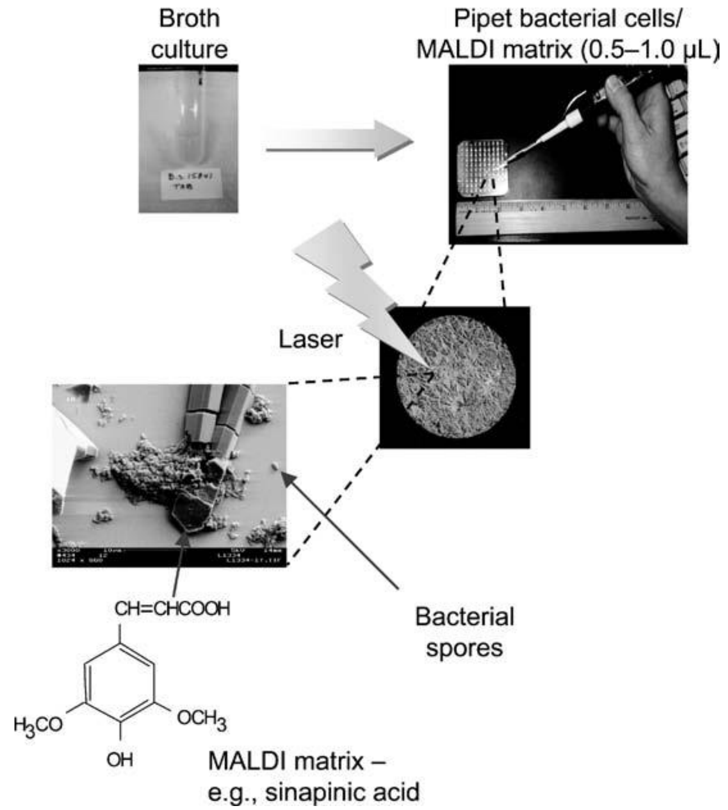

Broth culture
Pipet bacterial cells/ MALDI matrix (0.5–1.0 µL)
Laser
CH=CHCOOH
H3CO
OCH3
OH
Bacterial spores
MALDI matrix – e.g., sinapinic acid

Fig. 2. Schematic of MALDI sample preparation.

through the addition of an acid such as trifluoroacetic acid (TFA) or formic acid (*see* **Note 49**). Small acid-soluble proteins (SASPs) are key components of bacterial cells that are readily accessible and ionizable *(15)*.

2. Materials

2.1. Reagents

1. Alpha cyano-4-hydroxycinnamic acid (ACHC), and sinapinic acid (SA) (Bruker Daltonics, Bremen, Germany).
2. Ferulic acid (Aldrich, Milwaukee, WI).
3. Protein mix 1 and 3 (Bruker Daltonics, Bremen, Germany).
4. Horse heart cytochrome *c* and angiotensin I (Sigma, St. Louis, MO).
5. TFA (Aldrich, Milwaukee, WI). Note: Trifluoroacetic acid is corrosive and causes severe burns. Suitable protection, including gloves, laboratory coat, and eye and face protection should be donned when working with concentrated solutions.

6. Vegetative bacteria (American Type Culture Collection, (ATCC®) Manassas, VA).
7. Bacto Luria Bertani (LB), Broth Miller (Difco), and Bacto tryptic soy broth (TSB) without dextrose (Difco), and Bacto nutrient broth (Difco) (Becton Dickinson, Sparks, MD).
8. Lab Lemco Medium: Lab Lemco 23 g and 1.0 L of E-Pure® distilled deionized water.
9. *Nutrient sporulation medium (16).* 3.0-g tryptone, 3.0-g yeast extract, 2.0-g Bacto agar, 23-g LL agar, and 1.0-µL 1% $MnCl_2 \cdot 4H_2O$ in 1.0 L of E-Pure® distilled deionized water.
10. Acetonitrile and ammonium chloride (J.T. Baker, Phillipsburg, NJ).
11. Teflon coated 96 × 2 stainless steel MALDI plate (Applied Biosystems, Foster City, CA).
12. Stainless steel insert plate (Bruker Daltonics, Bremen, Germany).

2.2. Matrix Recipes

1. ACHC: 2-mg ACHC, 140-µL acetonitrile, and 60-µL 0.1% TFA.
2. SA: 2-mg sinapinic acid, 140-µL acetonitrile, and 60-µL 0.1% TFA.
3. FA (recipe #1): 2-mg ferulic acid, 140-µL acetonitrile, and 60-µL 0.1% TFA.
4. FA (recipe #2) *(17)*: 7-mg ferulic acid, 123-µL 65% formic acid, 155-µL acetonitrile, and 192-µL H_2O (*see* **Note 4**).

2.3. Instrumen-tation

Two different commercial MALDI-TOF MS instruments were used during the development of microorganism analysis in our laboratory due to the instrumentation available at different stages of the research. Any commercial MALDI mass spectrometer should be capable of obtaining data as discussed here.

1. An Applied Biosystems Voyager-DE RP MALDI-TOF MS operated in linear, delayed extraction, positive mode was used primarily for data collection of the vegetative cells.
2. A Bruker Autoflex II MALDI TOF/TOF MS instrument operated in reflector, positive mode was used for data collection of the spores.

2.4. Software

Instrument control and data collection is performed using the instrument control software as provided with the commercial instrument. The majority of the data collected in our laboratory for microorganism analysis has been done with manual data collection. However, automated data collection routines can be used.

1. *Applied Biosystems Grams.* Standard instrument control software supplied with the Voyager DE RP MALDI-TOF MS instrument.

2. *Bruker Daltonics Flex Control and Analysis Programs.* Standard instrument control software supplied with the Bruker Autoflex TOF/TOF MS instrument.
3. *Algoworks.* Algorithms developed at PNNL were used for data analysis. This includes the use of a novel patented peak detection algorithm incorporated with algorithms to account for the variability in the MALDI spectra due primarily to the variability in the microbial samples. Briefly these algorithms statistically compute a weighted average of sample replicate analyses and provide a means for statistically comparing an unknown sample to a reference library of MALDI spectra for different microorganisms *(6, 18, 19)*.

3. Methods

Several different MALDI matrices have been used for analysis of the microorganisms and the recipes are provided. As previously discussed, there is a difference in the mass spectra of a given sample obtained with the use of different MALDI matrices.

3.1. Vegetative Bacterial Sample Preparation

Bacterial samples can be cultured in a variety of culture media and culture conditions. Typical culture conditions for each organism of interest were used initially to get some reference data of several different types of organisms. *Bacillus subtilis (Bs)*, *Bacillus thuringiensis (Bt)*, *Bacillus atrophaeus*, *Bacillus cereus (Bc)*, *Serratia marcescens*, and the "unknown control," *Bacillus sphaericus*, were cultured in TSB and incubated overnight in a shaker incubator at 30°C, 120 rpm. *Pseudomonas stutzeri*, *Pantoea agglomerans*, and *Pseudomonas putida* were cultured in TSB and incubated at 37°C, 130 rpm. *Escherichia coli* (*E. coli*) was cultured in Luria Berani (LB) broth and incubated at 37°C, 130 rpm. The cells were centrifuged at 14,000 rpm for 2 min, decanted, and washed twice with 1 μL of 2% ammonium chloride to remove the majority of the culture medium and cell debris. In order to establish the effect of biological variability on the MALDI mass spectra, it is important to culture triplicates of each sample at a minimum to capture the biological variability.

3.2. Bacterial Spore Preparation

Spore cultures can be prepared in many different medium types. For research purposes to develop mass spectrometric methods several sporulation recipes were chosen to ensure sufficient sample for development purposes. *Bacillus subtilis* spores can be prepared in media such as nutrient sporulating media (NSM) or Lab Lemco (LL).

1. The agar plates are inoculated with 200 μL of starter culture grown in TSB at 30°C overnight (~14 h).
2. The plates are inverted and incubated in a 37°C incubator for 5–7 days.
3. The spores are harvested from the medium using sterile E-Pure® water and a sterile loop.
4. To produce clean spores the sample can be pelleted and washed by centrifugation 5–7 times in 20–30 μL of sterile E-Pure® water. The centrifugation is performed as follows: 1,000 rpm for 1–2 min, 5,000 rpm for 1–2 min, 11,000 rpm for 6–8 min. The sample is then decanted and a clean aliquot of water is added for the next wash. A clean spore preparation is approximately 95–98% phase bright spores as determined by phase contrast microscope evaluation. Spore counts can be determined by plate counts and are typically in the 10^8 CFU/mL range or higher for a clean spore preparation.

3.3. Estimation of Bacterial Spore Mass Used for MALDI-MS Analysis

Fergenson et al. *(20)* estimate the mass of one *Bacillus anthracis* spore being approximately 1×10^{-12} g or 1 pg. Based on this estimation and the use of approximately 0.5-μL sample deposited onto the MALDI sample plate, MALDI-MS analysis uses at most ~50 μg of bacterial spores per analysis. The majority of this sample remains after MALDI-MS analysis and can be retrieved from the MALDI sample plate for further analysis or archiving.

3.4. MALDI Vegetative Cell Sample Preparation

1. Ferulic acid (recipe #1), with two internal standards, cytochrome *c* and angiotensin I (at 5 and 2.5 mg/μL, respectively), was prepared.
2. A layering method was used for the bacterial analysis in which 1 μL of the bacterial sample was applied to the sample plate and allowed to air-dry. The concentration of cells was ~10^5 cells/μL when deposited onto the MALDI sample plate. Then 1 μL of the ferulic acid matrix solution was applied to the bacterial sample spot and allowed to air-dry (*see* **Note 5**).
3. Each spectrum was obtained averaging 128 laser shots. Each spectrum was internally calibrated (*see* **Note 6**) with the monomer ion of cytochrome *c* (*m/z* 12,361) and the monomer of angiotensin I (*m/z* 1,297). Regular stainless steel plates manufactured by Applied Biosystems were used.
4. The data files were then transferred to the data analyst for automated peak extraction and analysis using Algoworks algorithms.
5. Library reference fingerprints were created using ten replicate MALDI-TOF MS spectra/organism collected from replicate cultures on each of 3 days, for a total of 60 spectra/bacterium.

3.5. Matrix-to-Analyte Ratio Consi-derations

The ratio of matrix to cell concentration is important for successful MALDI mass spectral analysis. When possible, concentration of the sample with either cell counts or absorbance measurement is useful. However, if an unknown sample with no concentration information is to be analyzed, it might be helpful to prepare several different ratios of analyte to matrix. Shown in **Fig. 3** is a representation of the type of spectra obtained when the sample is too concentrated relative to the matrix added, too dilute relative to the matrix used, and appropriate ratio of analyte to matrix. In this case the use of an internal standard is helpful to determine whether an unknown sample is too dilute or too concentrated when no useful MALDI-MS data is collected *(21)* (*see* **Note 6**). The internal standard (cytochrome *c*) ions are not observed when the sample is too concentrated for the amount of matrix used. However, when the sample concentration is too dilute, the internal standard ions are observed.

3.6. MALDI Spore Sample Preparation

Without some pretreatment, bacterial spore samples do not yield as much MALDI mass spectral data compared with vegetative cells. One method for obtaining additional protein data from bacterial spore samples is to subject them to a simple pretreatment procedure.

1. A Petri dish, 10 cm in diameter, was used to accommodate an Applied Biosystems Teflon-coated 96 × 2 stainless steel MALDI plate (*see* **Note 7**).
2. Several layers of paper towels were cut into the shape of a circle slightly smaller than the Petri dish.
3. The layers of paper towels were placed into the bottom of the dish and wetted with distilled water, and then the Petri dish was covered.

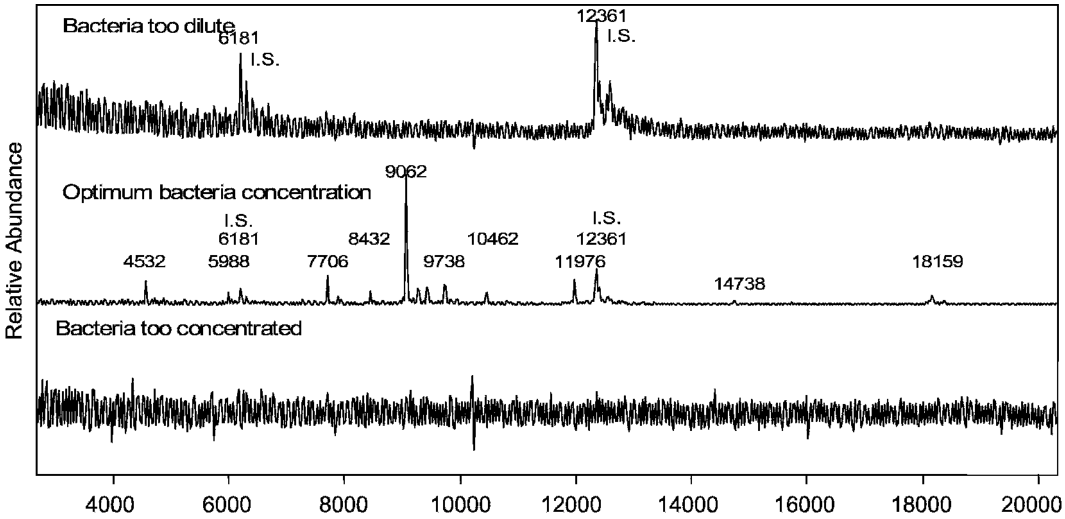


Fig. 3. Use of internal standard to determine appropriate relative ratio of sample to matrix.

4. The Petri dish was then placed into a previously heated 37°C oven to create a moisture chamber (*see* **Note 8**).
5. A 2% solution of TFA was prepared. Using a mechanical pipettor a 1-μL aliquot of 2% TFA was placed onto the MALDI spots in an *every other spot arrangement*.
6. Spore samples were vortexed slightly to resuspend and homogenize.
7. A 0.5-μL aliquot of the spores was pipetted onto each of the droplets of 2% TFA.
8. The prepared MALDI plate with the spots still wet was then placed into the warmed Petri dish with moist towels and covered.
9. The covered Petri dish was then placed back into the 37°C oven for 30 min.
10. After 30 min the Petri dish and MALDI plate were removed from the oven and allowed to dry.
11. Once the spots were dried 0.5-μL aliquots of an appropriate matrix were applied on top of the spore spots.
12. The spots were then analyzed after the instrument was properly calibrated (quadratic fit) using peptide 1 mix and protein 1 mix commercially available from Bruker. One of three different Bruker reflectron methods with parameters optimized for the three different matrices (ACHC, FA recipe #2, or SA) was loaded and used for data collection. Five 100-shot spectra were added together using the sum buffer to generate one saved spectrum.

3.7. MALDI Mass Spectral Data

Representative samples of the types of mass spectral data that can be obtained from direct analysis of vegetative bacterial cells and spores are presented here. **Figures 4** and **5** demonstrate

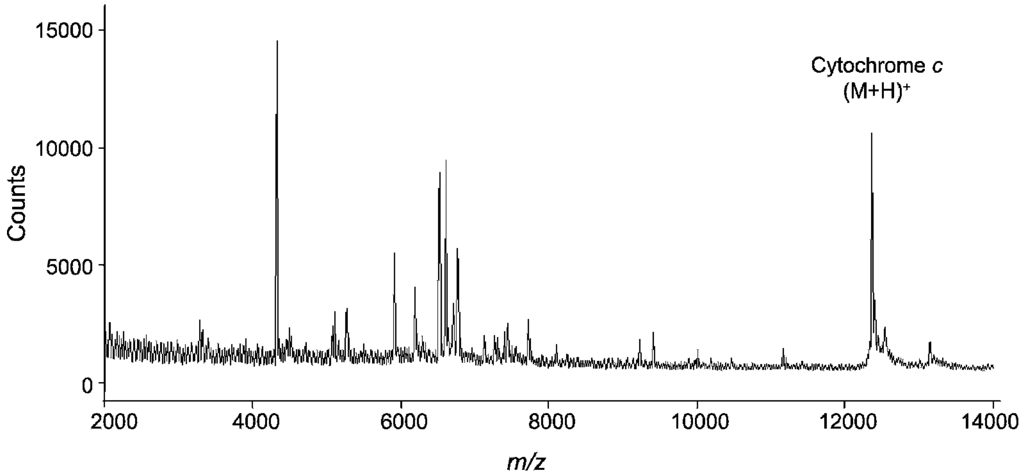


Fig. 4. Representative MALDI-MS spectrum of vegetative *Bacillus subtilis* 15841 cultured in tryptic soy broth with internal standard added prior to analysis.

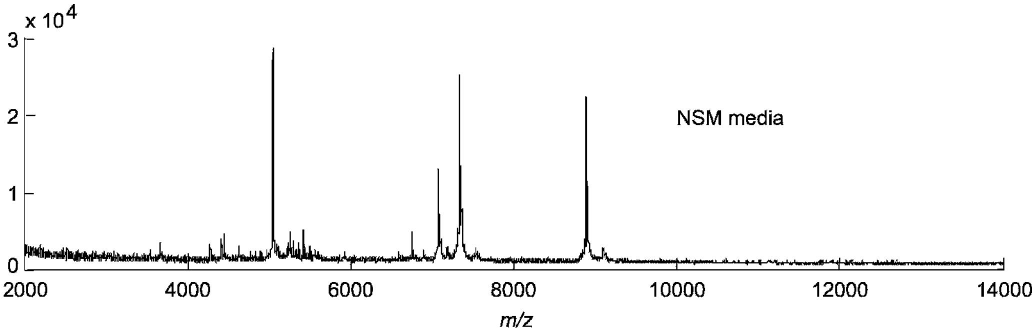


Fig. 5. Representative MALDI-MS of sporulated *Bacillus subtilis* 49760 in nutrient sporulating media agar plates without internal standard.

the species-specific differential biosignatures that emerge when vegetative cells are grown in different media. Contrasted are *Bacillus subtilis* grown in TSB and NSM agar. The effect of adding an internal standard to the sample spot prior to analysis is underscored as well.

4. Notes

1. If automation is to be employed in data collection, the best level of homogeneity obtainable across the MALDI spot is desired. When spotting MALDI plates it is best to allow wicking action to *pull* the aliquot of sample or matrix from the pipettor rather than touching the surface with the pipette tip.
2. Toward the goal of improving reproducibility when manually collecting data, it is important to capture as much variability across the MALDI spot as possible. Average as many laser shots from as many different locations across the spot as possible to achieve the desired level of reproducibility. The bacterial samples are more heterogeneous than protein solutions/standards, and therefore the MALDI sample spot containing bacterial cells is often more heterogeneous than for other samples.
3. Microorganisms are inherently heterogeneous variable samples compared with protein solutions and standards, and therefore some variability in MALDI spectra of intact bacteria and other microorganisms is expected. We have found the variability from the sample (biology) to be much more significant than the variability in the analytical method *(5)*.
4. To minimize the formylation of sample components, the formic acid content in prepared matrices should be at or below 17%.

5. While it is true that MALDI, as an ionization technique, is known for its tolerance toward impurities, there are concentration levels, beyond which it is nearly impossible to obtain useful signal. If you suspect that detergents or extraneous salts are interfering with data collection a simple spot washing step after cocrystalization of the sample and matrix may be all that is necessary. Simply apply a microliter of Milli-Q water to the spot and allow it to stand for several seconds. The water can then be removed with a pipette tip or with a corner of a Kimwipe. This step may need to be repeated for challenging samples/spots. As an alternative, pipette tips with packed beds such as C18 material may be used to prefractionate proteins and peptides from such unwanted components (e.g., ZipTips from Millipore).
6. The choice of internal standard vs. external standard for mass axis calibration is always a good question. The internal standard is useful for most accurate mass determinations but also leads to potential issues with competitive ionization in MALDI and suppression of some sample peaks. However, as discussed previously when having difficulty in getting useful signal from a true unknown sample, the use of internal standard can sometimes help determine if a negative MALDI-MS result is due to the sample being too concentrated (and no ions are observed) or too dilute (only ions from the internal standard are observed). It might be useful in the early stages of method development to spot samples both with and without internal standards to determine the variability in the *m/z* values detected as well as the appropriate matrix to analyte ratio for optimal MALDI-MS analysis.
7. Successful transfer of this method to the commercial Bruker sample plates was not completed in our laboratory due to project priorities.
8. To enhance reproducibility, within a sporulated sample set, always prewarm the moisture chamber to the desired temperature necessary, prior to insertion of the spotted MALDI plate.

References

1. Wahl, K., S. Wunschel, K. Jarman, N. Valentine, C. Petersen, M. Kingsley, K. Zartolas, and A. Saenz, (2002) Analysis of Microbial Mixtures by MALDI Time-of-Flight Mass Spectrometry. *Anal. Chem.* 74, 6191–6199
2. Hathout, Y., P.A. Demirev, Y.-P. Ho, J.L. Bundy, V. Rhyzov, L. Sapp, J. Stutler, J. Jackman, and C. Fenselau, (1999) Identification of *Bacillus* Spores by Matrix-Assisted Laser Desorption Ionization-Mass Spectrometry. *Appl. Environ. Microbiol.* 65, 4313–4319
3. Arnold, R.J. and J.P. Reilly, (1999) Observation of *Escherichia coli* Ribosomal Proteins and Their Posttranslational Modifications by Mass Spectrometry. *Anal Biochem.* 269, 105–112
4. Warscheid, B. and C. Fenselau, (2003) Characterization of *Bacillus* Spore Species and Their Mixtures Using Postsource Decay with a Curved-Field Reflectron. *Anal Chem.* 75, 5618–5627
5. Saenz, A.J., C.E. Petersen, N.B. Valentine, S.L. Gantt, K.H. Jarman, M.T. Kingsley, and

K.L. Wahl, (1999) Reproducibility of Matrix-Assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry for Replicate Bacterial Culture Analysis. *Rapid Commun. Mass Spectrom.* 13, 1580–1585

6. Jarman, K.H., S.T. Cebula, A.J. Saenz, C.E. Petersen, N.B. Valentine, M.T. Kingsley, and K.L. Wahl, (2000) An Algorithm for Automated Bacterial Identification Using Matrix-assisted Laser Desorption/Ionization Mass Spectrometry. *Anal. Chem.* 72, 1217–1223
7. Lay, J., (2001) MALDI-TOF Mass Spectro-metry of Bacteria. *Mass Spectrom.* 20, 172–194
8. Fenselau, C. and P. Demirev, (2001) Characterization of Intact Microorganisms by MALDI Mass Spectrometry. *Mass Spectrom.* 20, 157–171
9. Wilkins, C. and J.O. Lay, eds., (2005) Identification of Microorganisms by Mass Spectrometry. Wiley Interscience, NJ
10. Valentine, N.B., J.H. Wahl, M.T. Kingsley, and K.L. Wahl, (2002) Direct Surface Analysis of Fungal Species by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 16, 1352–1357
11. Wang, Z., L. Russon, L. Li, D.C. Roser, and S.R. Long, (1998) Investigation of Spectral Reproducibility in Direct Analysis of Bacteria Proteins by Matrix-assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 12, 456–464
12. Valentine, N., S. Wunschel, D. Wunschel, C. Petersen, and K. Wahl, (2005) Effect of Culture Conditions on Microorganism Identification by Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *Appl. Environ. Microbiol.* 71, 58–64
13. Karas, M., D. Bachmann, U. Bahr, and F. Hillenkamp, (1987) Matrix-Assisted Ultraviolet Laser Desorption of Non-Volatile Compounds. *Int. J. Mass Spectrom.* 78, 53–68
14. Karas, M. and U. Bahr, (1991) Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *Mass Spectrom. Rev.* 10, 335–357
15. Hathout, Y., B. Setlow, R.-M. Cabrera-Martinez, C. Fenselau, and P. Setlow, (2003) Small, Acid-Soluble Proteins as Biomarkers in Mass Spectrometry Analysis of *Bacillus* Spores. *Appl. Environ. Microbiol.* 69, 1100–1107
16. Nicholson, W.L. and P. Setlow, (1990) Sporulation, Germination and Outgrowth, in Molecular Biological Methods for *Bacillus*, C.R. Harwood and S.M. Cutting (eds). Wiley, Chichester, England
17. J. Madonna, A., F. Basile, I. Ferrer, M.A. Meetani, J.C. Rees, and K.J. Voorhees, (2000) On-Probe Sample Pretreatment for the Detection of Proteins Above 15 KDa from Whole Cell Bacteria by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 14, 2220–2229
18. Jarman, K.H., D.S. Daly, C.E. Petersen, A.J. Saenz, N.B. Valentine, and K.L. Wahl, (1999) Extracting and Visualizing Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectral Fingerprints. *Rapid Commun. Mass Spectrom.* 13, 1586–1594
19. Jarman, K.H., D.S. Daly, K.K. Anderson, and K.L. Wahl, (2003) A New Approach to Peak Detection. *Chemometrics Intell. Lab. Syst.* 69, 61–76
20. Fergenson, D.P., M.E. Pitesky, H.J. Tobias, P.T. Steele, G.A. Czerwieniec, S.C. Russell, C.B. Lebrilla et al., (2004) Reagentless Detection and Classification of Individual Bioaerosol Particles in Seconds. *Anal Chem.* 76, 373–378
21. Gantt, S.L., N.B. Valentine, A.J. Saene, M.T. Kingsley, and K.L. Wahi, (1999) Use of an Internal Control for Matrix-assisted Laser Desorption /Ionization Time-of-Flight Mass Spectrometry Analysis of Bacteria. *J. Am. Soc* Mass Spectrom. 10, 1131–1137.

Chapter 23

Mass Spectrometric Characterization of Neuropeptides

Stephanie S. Cape, James A. Dowell, and Lingjun Li

Summary

Because of their great biological significance, neuropeptides are the subject of intensive research. Mass spectrometry (MS) is a highly informative and sensitive method used for detecting and characterizing these compounds. Successful MS analysis of neuropeptides is dependent on careful sample preparation. Herein, we present two common sample preparation strategies: direct tissue analysis and pooled tissue extraction coupled with fractionation.

Key words: Neuropeptide, Mass spectrometry, Direct tissue analysis, Extraction, HPLC, De novo sequencing, Sample preparation.

1. Introduction

Neuropeptides function in intercellular communication as neurohormones, neuromodulators, or neurotransmitters. A prepropeptide is synthesized in the rough endoplasmic reticulum from which it is secreted and the signal sequence is removed. The propeptide then transverses the Golgi apparatus and is packaged into vesicles. Within the vesicles, proteolytic cleavage, amidation, glycosylation, phosphorylation, and disulfide bond formation occur to create the bioactive peptide. Often neuropeptides initiate their effects by activating G-protein-coupled receptors. Through this mechanism, they modulate numerous physiological processes such as respiration, feeding, sleep, learning, and memory *(1)*.

Because of their extensive involvement in crucial biological functions, neuropeptides have been the subject of intensive research. Discovery and structural elucidation of this diverse class of signaling molecules is daunting. Traditional methods for neuropeptide detection such as radioimmunoassays and immunocytochemical

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_23

techniques require some prior knowledge of the peptide sequence, and investigation of many peptides simultaneously is challenging if not impossible. Edman degradation offers a means to determine the amino acid sequence of a peptide but requires a large amount of purified sample.

Alternatively, mass spectrometry (MS)-based strategies have become popular in neuropeptide research. MS-based methods are able to detect molecules at low concentrations with molecular specificity in complex matrices. Neuropeptides are especially suitable for study by MS because their bioactive forms are often in the mass range ideal for MS detection. In addition, fragmentation patterns of peptides are particularly informative. However, MS characterization of biological molecules is often complicated by extracellular proteases, high-salt content, lipids, and high sample complexity, which may cause ionization suppression and/or dynamic range issues. Therefore, proper sample handling from animal dissection through analysis is key to obtaining useful and relevant results. As outlined in **Fig. 1**, samples may be prepared by homogenizing the tissue and extracting the neuropeptide content or investigated via direct tissue methods. Because of the wide variation in model systems and in the chemical properties of neuropeptides, it is not practical to present a single extraction scheme that works optimally for all samples. We therefore include two strategies (1) cold acidified methanol extraction (used primarily for invertebrate systems) and (2) boiling acetic acid extraction coupled with filtering (used primarily for mammalian tissues). Extraction is often coupled with fractionation, which improves data quality by reducing sample complexity. However, increased sample handling/processing may lead to sample dilution, loss, and/or chemically induced artifacts. The alternative strategy of direct analysis, which involves desorption of the analyte directly from the tissue, may therefore be more appropriate when limited sample is available.

Fragmentation information can substantially increase the confidence of peptide identification, facilitate characterization of posttranslational modifications, and enable de novo sequencing. The latter aspect is of particular importance when working with neuropeptides isolated from organisms for which genomic sequence is unavailable.

2. Materials

2.1. Materials/ Chemicals

1. Alconox.
2. Acetonitrile, HPLC grade.

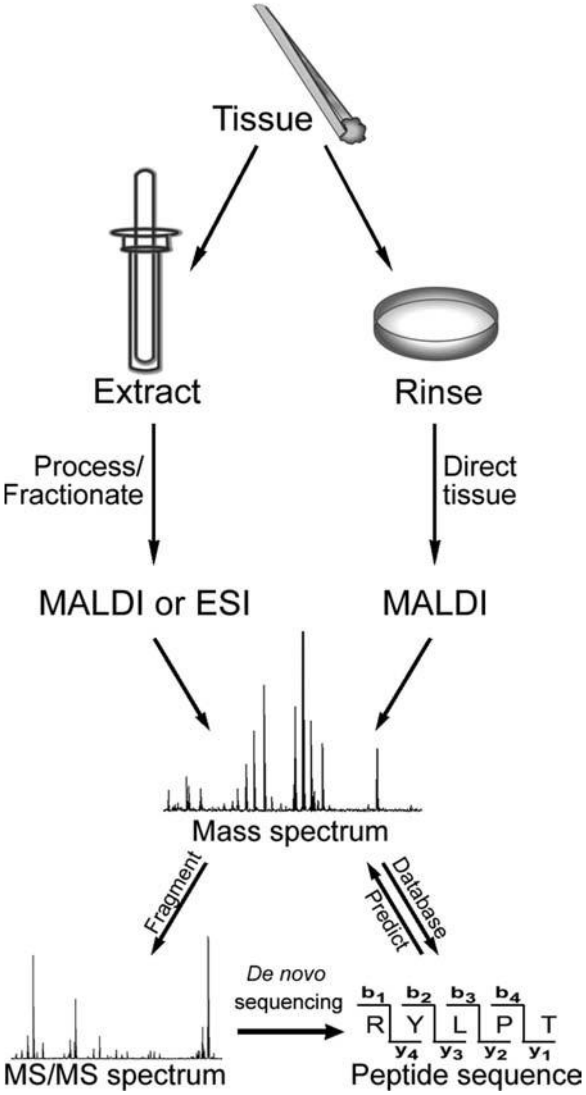

Tissue
Extract
Rinse
Process/
Fractionate
Direct
tissue
MALDI or ESI
MALDI
Mass spectrum
Fragment
Database
Predict
De novo
sequencing
b1 b2 b3 b4
R Y L P T
y4 y3 y2 y1
MS/MS spectrum
Peptide sequence

Fig. 1. Overview of common strategies for mass spectral characterization of neuropeptides. Once the tissue is dissected free of the animal, it can be rinsed and analyzed by direct tissue MALDI analysis, or the neuropeptide content can be extracted from the tissue sample. Extracts can be further processed and/or fractionated prior to MALDI or ESI MS analysis. Once a mass spectrum is obtained the neuropeptide sequence can be determined through database searching (when applicable) or tandem MS and de novo sequencing.

3. Methanol, purge and trap grade.
4. Glacial acetic acid.
5. 2,5-Dihydroxybenzoic acid (DHB), ICN Biomedicals; Aurora, OH.
6. DHB solutions: 10 mg/mL aqueous (dilute), 150 mg/mL of 50:50 water/methanol, v/v (concentrated).

7. Parafilm.
8. Formic Acid.
9. Trifluoroacetic Acid (TFA).
10. Water, double distilled by filtration system (Millipore; Bedford, MA).
11. Acidified methanol: 90% MeOH, 9% acetic acid, 1% H_2O.
12. Molecular weight cut-off tube (MWCO) (10-kDa) (Sartorius; Goettingen, Germany).
13. Microsonicator (Mixsonix; Farmingdale, NY).
14. Handheld ground glass homogenizer (Wheaton Science, Millville, NJ).
15. C18 ZipTips (Millipore, Billerica, MA).

2.2. Instrumentation

2.2.1. Matrix-Assisted Laser Desorption/Ionization (MALDI) Mass Spectrometry

Direct tissue analysis requires a mass spectrometer with a MALDI ionization source. An IonSpec Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (Lake Forest, CA) equipped with a 7-T actively shielded superconducting magnet was used to collect the data shown in Fig. 2. A 337-nm nitrogen

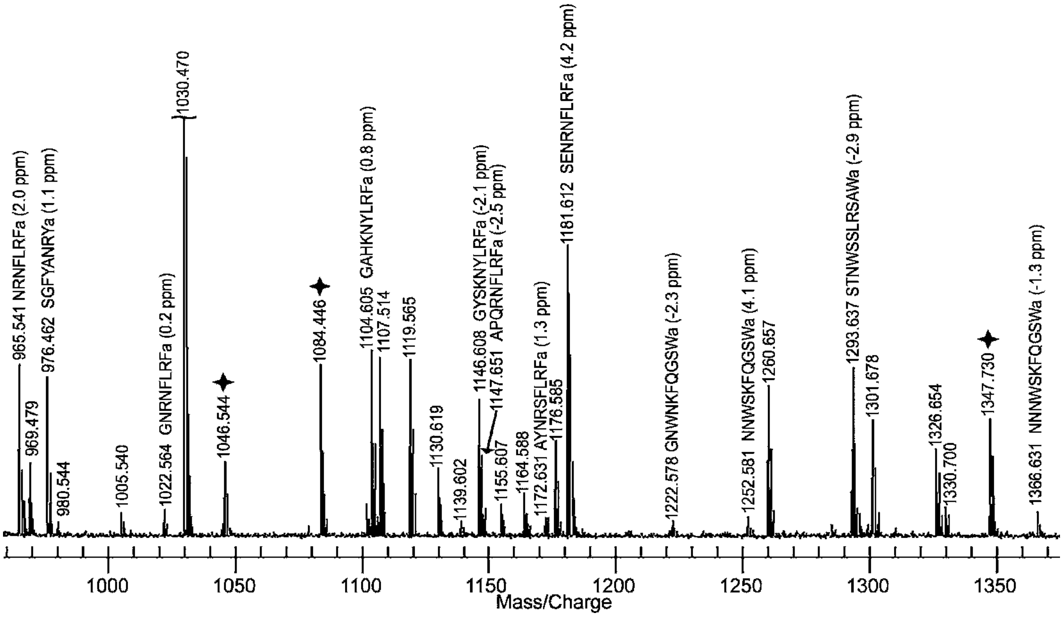


Fig. 2. Direct tissue mass spectrum of an invertebrate (*Cancer borealis*) neurosecretory organ. Many known neuropeptide signals are annotated with the single letter code amino acid sequence and mass measurement error in parts per million (ppm). Lowercase "a" means that the peptide is C-terminally amidated. Stars represent known peptide standards, which were added using the internal calibration by adjacent standards (InCAS) method.

laser (Laser Science, Inc., Franklin, MA) irradiates the matrix-coated tissue to create ions, which are collisionally cooled with nitrogen gas in an external source, and accumulated in a hexapole accumulation chamber before they are transferred through a quadrupole ion guide to the ICR cell. FTICR offers high resolution and high mass measurement accuracy; sensitivity can be increased by utilizing selective ion accumulation in the external accumulation chamber or in the ICR *(2)*. However, many detection methods are suitable for direct tissue analysis.

2.2.2. Electrospray Ionization (ESI) Instrument

A quadrupole time-of-flight (QTOF) mass spectrometer (Waters, QTOF Micro) equipped with a nanoelectrospray ionization source (nESI) is used in our laboratory. Other mass spectrometers capable of automated acquisition of tandem mass spectra are also suitable for peptide identification and sequencing analysis. A capillary HPLC (Waters, Milford, MA) system is used to deliver 6 μL of sample to a trap column (LC Packings PepMap C18, 300-μm i.d. × 5 cm; Amsterdam, The Netherlands) via an isocratic flow of mobile phase A (0.1% formic acid in water) at a rate of 30 μL/min for 3 min. The flow rate is then switched to 250 nL/min and the peptides are flushed onto the analytical column (Microtech C18, 75-μm i.d. × 15 cm; Vista, CA) and eluted over 30 min via a 5–45% gradient of mobile phase B (0.1% formic acid in acetonitrile) into the mass spectrometer.

The peptides are detected in positive mode. Data are collected in both MS-only mode and MS/MS data-dependent acquisition (DDA) mode. The DDA survey and MS-only mass-to-charge range is from *m/z* 400 to 2,000. The intensity threshold for switching from the survey scan to MS/MS is set at 15 counts. The scan time is 0.9 s, interscan time is 0.1 s, capillary voltage is 3,000 V, and cone voltage is 35 V.

2.2.3. Offline HPLC Instrument

A standard system for reverse-phase HPLC is required for offline HPLC separations. A Rainin Dynamax HPLC system equipped with a Dynamax UV-D II absorbance detector (Rainin Instrument Inc., Woburn, MA) is used in our laboratory. The mobile phases include mobile phase A: deionized water containing 0.1% TFA, and mobile phase B: acetonitrile (HPLC grade, Fisher Scientific) containing 0.1% TFA. Twenty microliters of extract is injected onto a Macrosphere C18 column (2.1-mm i.d. × 250-mm length, 5-μm particle size; Alltech Assoc. Inc., Deerfield, IL) or Phenomenex Gemini C18 column (2.1-mm i.d. × 150-mm length). The separations consist of a 120-min gradient of 5–95% mobile phase B. Fractions are automatically collected every 2 min throughout the gradient using a Rainin Dynamax FC-4 fraction collector.

3. Methods

3.1. Direct MALDI MS Analysis of Invertebrate Tissues

Unlike traditional mass spectrometric techniques that involve pooling a large number of ganglia or organs and extracting the peptide content prior to analysis, direct tissue experiments require much less sample preparation and allow comparison between individual samples or animals. This technique is sensitive enough for neuropeptide analysis from single organs *(2, 3)* and even single cells *(4, 5)*. In many cases, sufficient signal is obtained for peptide fragmentation information to be collected *(2, 5–7)*. Rinsing the tissue with dilute matrix solution and acidified methanol extracts peptides to the tissue surface and removes salts, leading to improved detection. Direct tissue analysis of neuropeptides has been applied extensively in invertebrate systems *(8)*, and to a lesser extent in more complex vertebrate systems *(3, 7)*. Figure 2 demonstrates direct tissue MALDI-FTICR analysis of neuropeptides from a neurosecretory organ. Internal calibration was accomplished by spotting a mixture of known neuropeptides next to the tissue and capturing both the internal calibrant and analyte ions in the detection chamber *(9)*.

1. Clean and dry all tools and the MALDI sample plate. Prepare acidified methanol and dilute and concentrated DHB solutions (*see* **Note 1**).
2. Place a small droplet of acidified methanol and a small droplet of dilute DHB on a clean hydrophobic surface such as a piece of parafilm or sylgard-coated Petri dish.
3. The tissue to be analyzed must be of an appropriate size (*see* **Note 2**). For larger tissues, it may be necessary to divide the sample into multiple pieces. Very small tissues can be pooled on the MALDI spot to increase peptide concentration per sample spot.
4. Immediately after tissue is dissected free of the animal, deactivate proteases and extract neuropeptides to the surface of the tissue by briefly dipping the tissue in a small droplet (~10 μL, *see* **Note 3**) of acidified methanol *(2)* (*see* **Note 4**).
5. Rinse the tissue in the droplet of dilute DHB solution to remove salts, and blot the forceps to remove excess liquid (*see* **Note 5**).
6. Place a small droplet of concentrated DHB (0.6 μL, *see* **Note 6**) on the sample facet, and quickly place the tissue in the droplet before the matrix crystallizes. Press the tissue with the forceps or a fine needle to encourage mixing of matrix and analyte and flatten the sample on the plate (*see* **Note 7**).
7. Allow the matrix to crystallize completely before introducing into the mass spectrometer.

8. Perform MALDI analysis on the sample using a laser power 20–30% higher than that used for conventional samples. It is often necessary to focus the laser on multiple locations throughout the tissue to determine which location will produce the best data (*see* **Note 8**).
9. Samples may be stored and reanalyzed. Freeze the tissue on the plate at −80°C and prior to reanalysis allow the plate to warm to room temperature in a desiccator.

3.2. Extraction of Neuropeptides

Neuropeptide extraction procedures require the reduction of protease activity. Postmortem protease activity yields protein degradation products which interfere with the identification of neuropeptides. This is especially true for mammalian tissues which require rapid postmortem protease deactivation to maintain sample integrity.

Protease heat deactivation employing microwave irradiation or boiling extraction buffers are common techniques to reduce protease activity. For mammalian tissues, heat deactivation is achieved by focused microwave irradiation during sacrifice (requires specialized equipment) *(10)*, by conventional microwave irradiation immediately following sacrifice (requires conventional microwave oven) *(11)*, or by boiling the tissue after dissection. Microwave irradiation is generally performed on the entire brain and dissection is performed by hand after irradiation. Alternatively, the whole brain can be snap-frozen immediately following sacrifice and subsequently dissected by a combination of cryostat sectioning and tissue punching. In this manner, the tissue is kept frozen during the dissection process. The proteases are deactivated immediately prior to extraction by placing the tissue punches in a boiling extraction buffer.

As an alternative to heat deactivation, proteases can also be deactivated using acidified organic extraction buffers, including acidified acetone, ethanol, and methanol. These acidic buffers deactivate proteases via denaturation and precipitation. The technique is more commonly used for invertebrate species but is also effective for mammalian tissues.

3.2.1. Acidified Methanol Extraction

1. Add 100–500 μL of ice-cold acidified methanol to pooled tissue (*see* **Note 9**).
2. Extract with a handheld tissue homogenizer, mechanical tissue grinder, or microsonicator on ice (*see* **Note 10**).
3. Carefully transfer the homogenate to a centrifuge tube (*see* **Note 11**).
4. Centrifuge the homogenate for 10–30 min at 15,000 × *g*. Transfer supernatant to a clean tube.
5. Reextract pellet by adding an additional 100–500 μL of acidified MeOH to the pellet, vortexing to resuspend, and then

homogenizing. Centrifuge as before and combine the supernatant with the previously collected supernatant.

6. If necessary, repeat extraction of the pellet and combine the supernatants.
7. Vacuum-dry combined supernatants.
8. Resuspend the dried neuropeptide extract in a minimal volume (20–30 μL) of 0.1% formic acid in water (*see* **Note 12**).
9. Prepare samples using offline HPLC or ZipTip, or directly analyze by MALDI-MS or online LC/MS/MS (*see* **Note 13**).

3.2.2. Aqueous Acetic Acid Extraction

1. For mammalian tissue, which has not been previously heated via microwave irradiation, add 100–500 μL of 90–100°C aqueous acetic acid (0.25%) to pooled tissue and place in a boiling water bath for 10 min. Cool on ice before homogenization. For previously microwave-irradiated mammalian tissue or invertebrate tissue, add 100–500 μL of ice-cold aqueous acetic acid to pooled tissue.
2. Extract tissue on ice with a handheld tissue homogenizer, mechanical tissue grinder, or microsonicator.
3. Carefully transfer the homogenate to a centrifuge tube.
4. Centrifuge the homogenate for 10–30 min at 15,000 × *g*. Transfer supernatant to a clean tube.
5. Reextract pellet by adding an additional 100–500 μL of aqueous acetic acid to the pellet, vortexing to resuspend, and then homogenizing. Centrifuge as before and combine the supernatant with the previously collected supernatant.
6. If necessary, repeat extraction of the pellet and combine the supernatants.
7. Filter supernatants through 10-kDa MWCO tubes. The flow through contains the neuropeptides (*see* **Note 14**).
8. Vacuum-dry MWCO filtered supernatants.
9. Resuspend the dried neuropeptides in a minimal volume (20–30 μL) of 0.1% formic acid in water (*see* **Note 15**).
10. Prepare samples using offline HPLC or ZipTip, or directly analyze by MALDI-MS or online LC/MS/MS (*see* **Note 13**).

3.3. Offline HPLC Fractionation for MALDI Analysis

Fractionation of extracted neuropeptides before MALDI analysis can often increase neuropeptidome coverage.

1. Reconstitute dried-down neuropeptide sample in 0.1% TFA in 95% water/5% acetonitrile.
2. Equilibrate column with 5% mobile phase B for 30 min at a flow rate of 0.2 mL/min.
3. Inject sample and run a gradient of 5–95% B over 120 min at a flow rate of 0.2 mL/min.

4. Collect fractions every 1–4 min, depending on sample complexity.
5. Wash the column and store in 50% acetonitrile/50% water without TFA.
6. Vacuum-dry fractions.
7. Resuspend the dried neuropeptide samples in a minimal volume of 0.1% TFA in water and analyze by MALDI-MS.

3.4. Tandem MS

Fragmentation of neuropeptides is often necessary for confident identification. While postsource decay (PSD), collisionally induced dissociation (CID), and infrared multiphoton dissociation (IRMPD) create primarily b-type and y-type ions, as well as neutral losses, electron-based tandem MS techniques such as electron transfer dissociation (ETD) and electron capture dissociation (ECD) produce mostly c-type and z-type ions while preserving posttranslational modifications. However, ETD and ECD are not used in conjunction with direct tissue analysis because they only produce detectable fragments from multiply charged ions such as those formed by ESI. Figure 3 shows a spectrum produced by CID of a neuropeptide on an ESI QTOF mass spectrometer. This instrumental configuration is advantageous for sequencing analysis not only because of the ESI source, but also because CID coupled with TOF analysis promotes the production and detection of amino acid-specific immonium ions, which are invaluable for determining peptide composition *(12–14)*. Chemical derivatization of specific amino acid residues or the peptide termini can greatly aid in de novo sequencing *(6, 15, 16)*. Some of these derivatization procedures have been shown to effectively label neuropeptides in tissue *(6, 17)* and thus can benefit direct tissue MS/MS analysis as well.

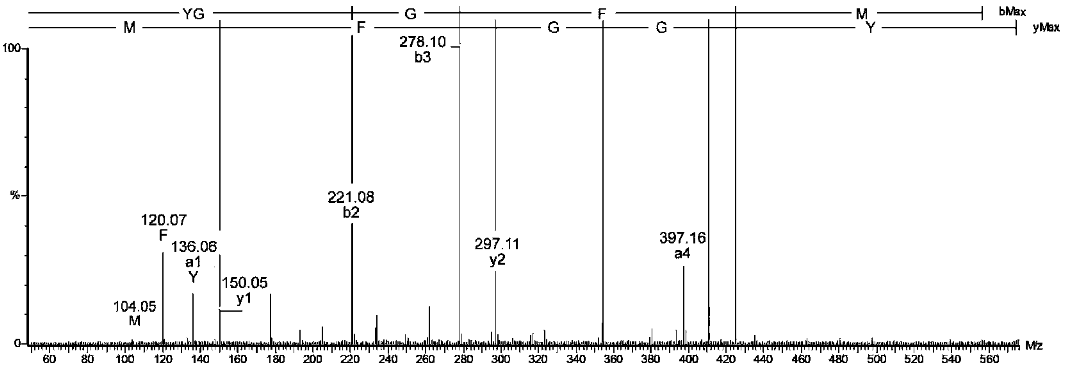


Fig. 3. MS/MS spectrum from methionine–enkephalin extracted from the rat dorsal striatum and analyzed by LC/MS/MS. The amino acid sequence of the peptide was determined by de novo sequencing using Waters' PepSeq software, with major b-type and y-type fragment ions and immonium ions labeled.

3.5. Data Analysis

Identification of neuropeptides can be achieved in a number of ways, depending on the type of MS data and database availability. The simplest way to identify neuropeptides is by manually comparing experimental masses to a list of known neuropeptide masses or by automated database searching using a database such as SwePep (www.swepep.org) *(18)*. This type of analysis is typically done using MALDI data (Fig. 2) and has the advantage of simplicity. However, this method does not allow sequencing of novel neuropeptides. Neuropeptide sequencing requires an instrument capable of producing fragment ions (MS/MS).

Depending on instrument availability and the species of interest, a number of different strategies can be employed for identification of neuropeptides using MS/MS startegies. If the genome is available, a standard protein database search can be performed using proteomic-based search engines, such as Mascot (Matrix Science, http://www.matrixscience.com/) or SEQUEST (Thermo Corp., http://www.thermo.com/) *(19)*. Although these search engines were designed for proteomics applications, they can be easily adapted to neuropeptidomics. However, neuropeptide searches differ from standard proteomics search strategies. Commonly, proteomics experiments employ an enzymatic digestion step (usually with trypsin) to produce peptides from the parent proteins. Many of the digestive enzymes used for proteomics produce a specific C- or N-terminus amino acid in the resulting peptides. This specificity is commonly used as a restrictive search parameter in the database search. Since it is unnecessary to enzymatically digest neuropeptides due to their small size, using proteomic-based search engines requires the enzyme to be specified as *none* or *nonspecific*. This will cause the search engine to interrogate all possible cleavages of a protein instead of just the enzyme-specific cleavages.

Another way to identify neuropeptides employing a database is by sequence tag searching. In contrast to Mascot and SEQUEST that base identifications on parent masses and MS/MS fragmentation patterns, sequence tag searches are performed using small contiguous strings of identified amino acids (sequence tags) compared against sequences in the database *(20)*. Since this technique does not rely on parent mass identification, it is a powerful tool for identifying posttranslationally modified neuropeptides. Unfortunately, the tools for sequence tag searching are not as well developed or automated as the standard proteomics search engines. MS-Seq (Protein Prospector, http://prospector.ucsf.edu/) and Mascot offer sequence tag searching for single tag queries. Both of these interfaces are flexible and powerful but can be time-consuming if analyzing a large volume of data. PEAKS (Bioinformatics Solutions, http://www.bioinformaticssolutions.com/) employs an automated de novo sequencing algorithm (see discussion later) to produce sequence tags and then automatically performs a database search using the de novo derived tags, significantly enhancing sequence tagging search throughput.

If the species of interest does not have a sequenced genome, de novo sequencing of the neuropeptide is required. In contrast to database searching, de novo peptide sequencing is the direct *reading* of the amino acid sequence from the MS/MS spectrum. With a great deal of experience, de novo sequencing can be done manually. However, this is extremely labor intensive and time-consuming. Fortunately, a number of software packages exist, which perform de novo sequencing (*see* Fig. 3), including LuteFisk (http://www.hairyfatguy.com/lutefisk/), PepSeq (packaged with Waters' MassLynx software, http://www.waters.com/), Mascot Distiller (Matrix Science), and PEAKS (Bioinformatics Solutions). A number of these programs, such as PEAKS and Mascot Distiller, include automated de novo processing to enhance sequencing speed and flexibility.

De novo sequencing does not always produce an accurate sequence for the entire neuropeptide. In this case, a BLAST homology search (http://www.ncbi.nlm.nih.gov/BLAST/) can be performed. A BLAST search compares a partial neuropeptide sequence against the database of a closely related species. While a homology search rarely yields a complete sequence, it can provide useful information about evolutionary origins and potential function of the partially sequenced neuropeptide. However, BLAST queries are not optimized for homology searches employing partial de novo sequences and can often lead to erroneous results. More recently, the use of pattern-finding software, such as SPIDER (Bioinformatics Solutions) and MEME (http://meme.sdsc.edu/meme/) has improved partial sequence homology searches *(21)*. These programs are optimized for MS/MS derived data and are more tolerant to sequencing errors than BLAST searches.

4. Notes

1. MALDI targets may be cleaned by gentle scrubbing with Alconox. Subsequent sonication or repeated rinsing with water and methanol to remove any residual compounds is recommended. Tools may be cleaned by gentle scrubbing and/or sonication in the presence of water and/or ethanol. DHB solutions should be prepared fresh daily. Acidified methanol may be stored for later use.
2. Tissue preparation is highly sample dependent. Detection of neuropeptides from nerve tissues often requires the removal of extraneous tissues and/or sheaths from the sample of interest.

If cryostat slicing of the tissue is desired, dissect out the tissue, rinse in water, and embed in gelatin prior to snap-freezing in a dry ice/ethanol bath. Optimal cutting temperature (OCT) media causes mass spectral interference in the mass range of neuropeptides. Slices can then be thaw-mounted on the MALDI plate and covered with matrix.

3. This volume is dependent on tissue size. There must be enough volume to submerge the tissue, but minimal volume is preferable.
4. If the tissue is to be stored prior to analysis, place the tissue in a siliconized tube with enough acidified methanol to submerge it.
5. If different matrix is to be used for analysis, a dilute solution of alternative matrix or water may be substituted *(22, 23)*.
6. The optimal matrix volume will firmly affix the tissue to the plate, but not cause significant analyte dilution. More matrix solution may be added to the top of the tissue as necessary. DHB or α-cyano-4-hydroxycinnamic acid (CHCA) is commonly used as matrices for neuropeptide analysis. DHB, which requires higher laser power than CHCA for the analyte to be desorbed, is characterized by less chemical noise interference than CHCA. However, CHCA forms small relatively homogeneous crystals whereas DHB crystals are long, needle-like, and heterogeneous. As a result, DHB preparations require more extensive searching throughout the sample spot for regions that produce high-quality MS signals. Mixtures of these two common matrices have also shown promise for MALDI peptide analysis *(24)*. Matrix additives such as formic acid (FA) and TFA can also influence the quality of the spectra obtained *(25)*. Specifically, it has been shown that improved reproducibility, signal-to-noise ratio, and reduced ion suppression are obtained when CHCA is prepared with FA and DHB with TFA *(24)*.
7. Thicker samples generally lead to decreased sensitivity and resolution especially in TOF analysis.
8. More ionization events usually correspond to more signal through accumulation of ions such as in a gated-trapping instrument or through signal averaging such as in TOF instruments. However, neuropeptide signal may become depleted through repeat analysis.
9. Amount of tissue depends on availability, size, and neuropeptide content of tissue. Commonly, 20–30 *Cancer borealis* brains are pooled for extraction. Typically, 20–100 mg of mammalian tissue is used for extraction.
10. Microsonication power settings must be optimized to achieve good results. Too much power results in a high degree of

lipid extraction. Too little power does not yield an efficient neuropeptide extraction. About 50% power on a 300-W microsonicator is a good place to start.

11. A Pasteur pipet works well for transferring the homogenate. Because there will be residue remaining in the transfer pipet, it is best to use this same pipet for transferring the homogenate during subsequent pellet extractions.
12. For MALDI-MS analysis, 0.1% TFA in water may be substituted for 0.1% FA.
13. For tissue extracts containing high concentrations of lipids, a simple extraction procedure is sometimes helpful. Remove lipids by liquid–liquid partitioning via sequential extraction into ethyl acetate and then *n*-hexane. After partitioning, collect the lower aqueous phase and vacuum-dry prior to ZipTip cleanup or LC/MS.
14. Prerinse filter twice with aqueous acetic acid before filtering extracts to remove glycerol from MWCO membrane.
15. For MALDI analysis 0.1% TFA in water may be substituted.

Acknowledgments

This work was supported in part by National Science Foundation CAREER Award (CHE-0449991), National Institutes of Health through grant 1R01DK071801, and an Alfred P. Sloan Research Fellowship (L.L.). S.S.C. is supported by an individual NRSA predoctoral fellowship 1F31NS053283. J.A.D. acknowledges an American Foundation for Pharmaceutical Education (AFPE) predoctoral fellowship.

References

1. Strand, F. L. (1994) Models of neuropeptide action. New York Academy of Sciences, New York, NY
2. Kutz, K. K., Schmidt, J. J., and Li, L. (2004) In situ tissue analysis of neuropeptides by MALDI FTMS in-cell accumulation. *Anal. Chem.* 76, 5630–5640
3. Fournier, I., Day, R., and Salzet, M. (2003) Direct analysis of neuropeptides by in situ MALDI-TOF mass spectrometry in the rat brain. *Neuro. Endocrinol. Lett.* 24, 9–14
4. DeKeyser, S. S., and Li, L. (2007) Mass spectrometric charting of neuropeptides in arthropod neurons. *Anal. Bioanal. Chem.* 387(**1**), 29–35 DOI: 10.1007/s00216–006–0596-x
5. Neupert, S., Predel, R., Russell, W. K., Davies, R., Pietrantonio, P. V., and Nachman, R. J. (2005) Identification of tick periviscerokinin, the first neurohormone of ixodidae: Single cell analysis by means of MALDI-TOF/TOF mass spectrometry. *Biochem. Biophys. Res. Commun.* 338, 1860–1864
6. Yew, J. Y., Dikler, S., and Stretton, A. O. (2003) De novo sequencing of novel neuropeptides directly from ascaris suum tissue

using matrix-assisted laser desorption/ionization time-of-flight/time-of-flight. *Rapid Commun. Mass Spectrom.* 17, 2693–2698

7. Jespersen, S., Chaurand, P., van Strien, F. J., Spengler, B., and van der Greef, J. (1999) Direct sequencing of neuropeptides in biological tissue by MALDI-PSD mass spectrometry. *Anal. Chem.* 71, 660–666
8. Hummon, A. B., Amare, A., and Sweedler, J. V. (2006) Discovering new invertebrate neuropeptides using mass spectrometry. *Mass Spectrom. Rev.* 25, 77–98
9. O'Connor, P. B., and Costello, C. E. (2000) Internal calibration on adjacent samples (InCAS) with Fourier transform mass spectrometry. *Anal. Chem.* 72, 5881–5885
10. Svensson, M., Skold, K., Svenningsson, P., and Andren, P. E. (2003) Peptidomics-based discovery of novel neuropeptides. *J.Proteome Res.* 2, 213–219
11. Che, F. Y., Lim, J., Pan, H., Biswas, R., and Fricker, L. D. (2005) Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell. Proteomic.* 4, 1391–1405
12. Baggerman, G., Boonen, K., Verleyen, P., De Loof, A., and Schoofs, L. (2005) Peptidomic analysis of the larval drosophila melanogaster central nervous system by two-dimensional capillary liquid chromatography quadrupole time-of-flight mass spectrometry. *J. Mass Spectrom.* 40, 250–260
13. Fu, Q., Goy, M. F., and Li, L. (2005) Identification of neuropeptides from the decapod crustacean sinus glands using nanoscale liquid chromatography tandem mass spectrometry. *Biochem. Biophys. Res. Commun.* 337, 765–778
14. Mohring, T., Kellmann, M., Jurgens, M., and Schrader, M. (2005) Top-down identification of endogenous peptides up to 9 kDa in cerebrospinal fluid and brain tissue by nanoelectrospray quadrupole time-of-flight tandem mass spectrometry. *J. Mass Spectrom.* 40, 214–226
15. Fu, Q., and Li, L. (2005) De novo sequencing of neuropeptides using reductive isotopic methylation and investigation of ESI QTOF MS/MS fragmentation pattern of neuropeptides with N-terminal dimethylation. *Anal. Chem.* 77, 7783–7795
16. Hsu, J. L., Huang, S. Y., Shiea, J. T., Huang, W. Y., and Chen, S. H. (2005) Beyond quantitative proteomics: Signal enhancement of the a1 ion as a mass tag for peptide sequencing using dimethyl labeling. *J. Proteome Res.* 4, 101–108
17. DeKeyser, S. S., and Li, L. (2006) Matrix-assisted laser desorption/ionization fourier transform mass spectrometry quantitation via in cell combination. *Analyst* 131, 281–290
18. Falth, M., Skold, K., Norrman, M., Svensson, M., Fenyo, D., and Andren, P. E. (2006) SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol. Cell. Proteomic.* 5, 998–1005
19. Ducret, A., Van Oostveen, I., Eng, J. K., Yates, J. R.III, and Aebersold, R. (1998) High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. *Protein Sci.* 7, 706–719
20. Mann, M., and Wilm, M. (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* 66, 4390–4399
21. Baggerman, G., Liu, F., Wets, G., and Schoofs, L. (2005) Bioinformatic analysis of peptide precursor proteins. *Ann. N. Y. Acad. Sci.* 1040, 59–65
22. Neupert, S., and Predel, R. (2005) Mass spectrometric analysis of single identified neurons of an insect. *Biochem. Biophys. Res. Commun.* 327, 640–645
23. Stemmler, E. A., Provencher, H. L., Guiney, M. E., Gardner, N. P., and Dickinson, P. S. (2005) Matrix-assisted laser desorption/ionization fourier transform mass spectrometry for the identification of orcokinin neuropeptides in crustaceans using metastable decay and sustained off-resonance irradiation. *Anal. Chem.* 77, 3594–3606
24. Laugesen, S., and Roepstorff, P. (2003) Combination of two matrices results in improved performance of MALDI MS for peptide mass mapping and protein analysis. *J. Am. Soc. Mass Spectrom.* 14, 992–1002
25. Cohen, S. L., and Chait, B. T. (1996) Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal. Chem.* 68, 31–37

Chapter 24

Peptide and Protein Ion/Ion Reactions in Electrodynamic Ion Traps: Tools and Methods

Scott A. McLuckey

Summary

Gas-phase reactions of oppositely charged ions can play important roles in the analysis of peptides and proteins. Electrospray ionization (ESI) can yield multiply charged versions of gaseous peptide and protein ions that can react with oppositely charged ions via several distinct mechanisms. Reagent ions that react via proton transfer can be used to facilitate protein mixture analysis and the mass assignment of product ions in a tandem mass spectrometry experiment. Proton transfer reactions can also be used to concentrate protein ion signals into one or two charge states and can be used to charge state purify a precursor ion population for subsequent dissociation. Electron transfer reactions have been shown to lead to fragmentation of gaseous protonated peptides and proteins. The extent of sequence information available from an electron transfer reaction is often greater than that obtained via conventional collision-induced dissociation. Electron transfer dissociation is particularly useful in probing the structures of polypeptides with labile posttranslational modifications. We summarize here the tools and general methods for conducting ion/ion reaction studies with emphasis on electrodynamic ion traps as the reaction vessels.

Key words: Ion/ion reactions, Electron transfer dissociation, Quadrupole ion trap, Linear ion trap, Multiply charged ions.

1. Introduction

The advent of electrospray ionization *(1)* (ESI), as applied to peptides and proteins, introduced multiply charged bio-ions to the rich array of charged species amenable to mass spectrometry. The multiple charging phenomenon has numerous consequences both for the chemistry and the mass analysis of the ions. For example, ESI of proteins typically leads to a distribution of charge states that can provide multiple mass measurements per protein

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_24

and yields ions with mass-to-charge values that can be accommodated by mass analyzers with only modest upper *m/z* limits *(2)*. In addition to these generally beneficial consequences of multiple charging, there are also complications associated with the spectral congestion that can occur with the ESI of mixtures and with the determination of ion charge. In addition to issues associated with mass analysis, multiple charging also profoundly influences both the unimolecular and bimolecular gas-phase chemistries of the ions. The unimolecular chemistry of peptide and protein ions is particularly relevant to protein analysis insofar as fragmentation reactions are often used to identify and characterize proteins, usually within the context of a tandem mass spectrometry experiment.

The multiple charging phenomenon associated with ESI has also enabled the study of an essentially new class of chemical reactions, viz., gas-phase bio-ion/ion reactions, which is becoming increasingly recognized as having important practical applications in protein analysis. Gas-phase proton transfer reactions, for example, have been used to simplify ESI mass spectra of protein mixtures *(3, 4)*, facilitate charge state determination of product ions derived from multiply charged precursor ions *(5)*, concentrate and charge state purify ions for subsequent tandem mass spectrometry *(6)*, invert the charge of an ion *(7)*, and increase the absolute charge of an ion *(8, 9)*. Gas-phase electron transfer reactions have been shown to be useful for deriving primary structure information from peptides and proteins with particular utility for characterizing posttranslational modifications *(10)*.

Applications of ion/ion reactions, as well as overviews of fundamental aspects associated with bio-ion/ion reactions have recently been reviewed *(11–13)*. This chapter focuses on the tools and methods used to effect ion/ion reaction experiments with emphasis on experiments conducted using an electrodynamic ion trap as the reaction vessel. However, brief discussion is also provided regarding approaches that employ ion/ion reactions prior to sampling into a mass spectrometer.

2. Tools for Bioion/Ion Reactions

ESI enabled the study of bio-ion/ion reactions with mass spectrometry because the multiple charging phenomenon allowed for the possibility that at least one of the ionic products of an ion/ion reaction retains a non-zero charge. The key criteria for establishing reaction conditions amenable to efficient ion/ion reactions for analytical applications are as follows:

1. Means for forming ions of opposite polarity must be available.
2. Oppositely charged ions must be simultaneously present in overlapping regions of space.
3. The relative velocities of the oppositely charged ions should be minimized to maximize ion/ion reaction cross section *(12, 13)*.

To date, the approaches used to effect bio-ion/ion reactions fall into two general categories. These include techniques that mix oppositely charged ions at atmospheric or near atmospheric pressures prior to sampling into a mass spectrometer and techniques based on mutual storage of ions within an electrodynamic ion trap. The former set of approaches rely on mixing thermalized ions largely in the absence of external electric or magnetic fields in order to meet criteria 2 and 3. The latter set of approaches also rely on the collisional cooling of ions to minimize relative velocities, although at somewhat lower pressures and longer times, to satisfy criterion 3 but differ from the former approaches in that electrodynamic electric fields are employed to meet criterion 2.

2.1. Atmospheric Pressure Ion/Ion Reactions/Sampling into MS

Approaches based on effecting ion/ion reactions prior to sampling products into a mass spectrometer rely on the continuous mixing of the reactant ions and are not subject to the constraint of storing reactants simultaneously, as do the ion trapping methods (see later). Any mass spectrometer platform can be coupled with a reactor of this category, provided it has an atmosphere/vacuum interface. However, if relatively high-mass biomolecules are to be subjected to analysis, it is highly desirable that the mass analyzer be able to accommodate a wide *m/z* range. A limitation of this type of reactor, relative to an ion trap (see later), is that the reactions cannot be effected between stages of mass analysis.

2.1.1. Y-Tube

The first studies involving the reactions of multiply charged bioions with oppositely charged ions were conducted using a Y-tube arrangement *(14, 15)*, where ions of opposite polarity were admitted via separate arms of the Y-tube and were merged in the flow leading to the atmosphere/vacuum interface of the mass spectrometer. A quadrupole mass filter was used to analyze the products of the interaction. Both stainless steel and glass versions of the Y-tube reactor were described. In the latter case, each of the ends was made electrically conductive by depositing a silver coating so that independent potentials could be applied. Three modes of operation were employed (a) single electrospray (no ion/ion reactions) to enable the collection of data in the absence of reactions, (b) dual electrospray with opposite polarity sprays, and (c) electrospray in one arm and discharge ionization in the other arm for formation of singly charged ions.

Advantages of the Y-tube approach include a high degree of flexibility in the combinations of ions that can be subjected to study. For example, any ions that can be formed at atmospheric pressure can be drawn into one of the arms of the Y-tube for study. Singly charged ions formed via discharge-based atmospheric pressure ionization or multiply charged ions formed via a spray ionization method could be used as reagent ions to react with analyte ions. As a result, a variety of reaction phenomenologies were noted with the Y-tube apparatus despite the limited number of reports that described its use. Furthermore, the reactor itself (viz., the Y-tube) is independent of the mass spectrometer such that any type of instrument with an atmosphere/vacuum interface can be used to analyze the products. In the case of the reported studies, a mass filter was used. Unfortunately, the relatively low upper *m/z* limit of the analyzer limited the range of studies that could be pursued for relatively high-mass ions.

A drawback to the Y-tube reactor was the limited ability to define precisely the ion/ion reaction conditions. Generally, mixtures of ions were present in the reaction zone, along with solvent vapors and, in some cases, relatively reactive neutral species formed via the discharge. Imprecise control over the numbers of ions of each reactant and a limited ability to vary reaction time via flow rate also contributed to a relatively high degree of variability in the appearance of the spectra.

2.1.2. ESI/^{210}Po

A second approach to atmospheric pressure ion/ion reactions adapted a technique used to manipulate aerosol charge prior to ion mobility analysis *(16)* whereby ions formed via electrospray were drawn into a weakly ionized plasma arising from ^{210}Po nuclear decay (α-particle emission) in air (*see* **Fig. 1**).

Singly charged ions in the bipolar plasma reduce the charge states of the ions admitted via electrospray. Control over the reaction rates is provided by adjustment of the α-particle flux, via masking of the radioactive element, or by adjustment of the flow rate. Masking the polonium source affects the number of reagent ions whereas adjustment of the flow rate affects residence time of the electrospray ions in the reaction volume. Each measure, or a combination of the two, allows for control of the extent of conversion of multiply charged reactant ions to products *(17)*. This reactor was coupled with orthoganol time-of-flight mass analysis, which provided a wide mass-to-charge range for product ion analysis. This arrangement is restricted to the use of singly charged ions for reactions with multiply charged analyte ions and is particularly well suited to reduction of spectral congestion associated with electrospray of biopolymer mixtures.

2.1.3. ESI/Corona Discharge

A corona discharge has been used to generate singly charged reactant ions in a flow-through ion source analogous to that of

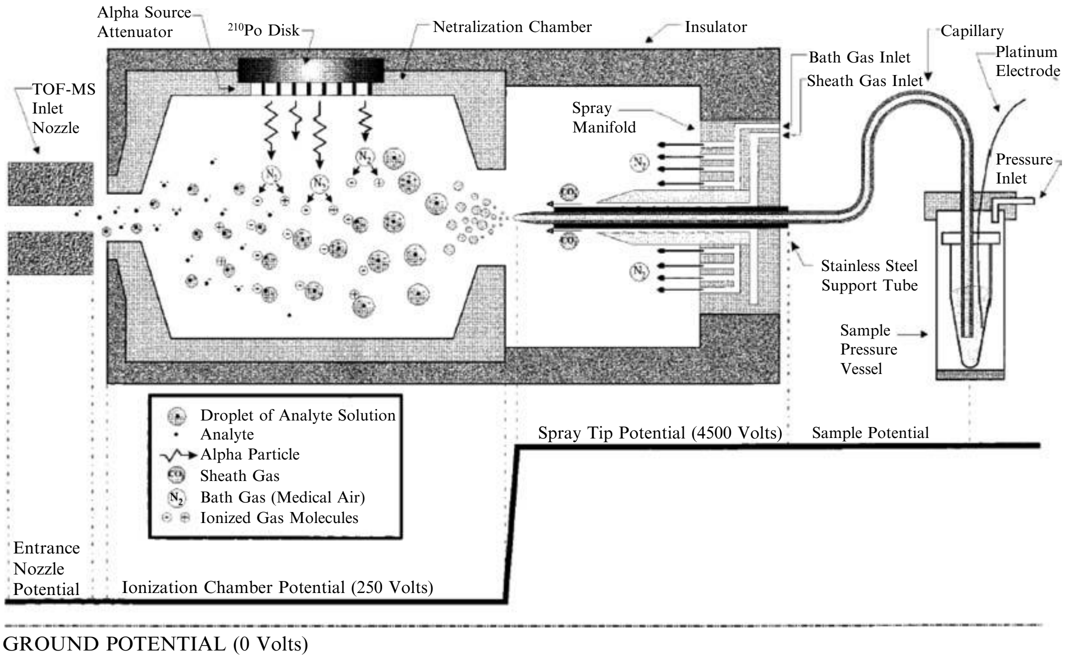

Alpha Source Attenuator
^{210}Po Disk
Netralization Chamber
Insulator
Capillary
Platinum Electrode
Bath Gas Inlet
Sheath Gas Inlet
TOF-MS Inlet Nozzle
Spray Manifold
Pressure Inlet
Stainless Steel Support Tube
Sample Pressure Vessel
Droplet of Analyte Solution
Analyte
Alpha Particle
Sheath Gas
Bath Gas (Medical Air)
Ionized Gas Molecules
Spray Tip Potential (4500 Volts)
Sample Potential
Entrance Nozzle Potential
Ionization Chamber Potential (250 Volts)
GROUND POTENTIAL (0 Volts)

Fig. 1. Schematic of charge neutralization apparatus, which uses an α-particle emitter to generate ions to reduce the charge of ions formed from electrospray ionization. Reprinted with permission from *(17)*. Copyright 2000 American Chemical Society.

the ESI/^{210}Po arrangement *(18)*. The corona discharge current can be used to control the reagent ion flux. A cylindrical capacitor source *(19)* has also been combined with corona discharge for ion/ion reactions studies *(20)* (*see* **Fig. 2**). The cylindrical capacitor source is a variation of conventional ESI that does not require a high potential to be applied to the emitter. This source has been reported to provide stable ion production and is capable of forming ions not readily formed by conventional electrospray *(19, 20)*. A major practical advantage of the corona discharge source is that is obviates use of a radioactive source.

2.2. Ion/Ion Reactions Within Electrodyamic Ion Traps

Electrodynamic ion traps are particularly well suited to serve as reaction vessels for ion/ion reactions. The ability to store oppositely charged ions simultaneously in overlapping regions of space is obviously a key feature of these devices. The fact that ions assume mass-to-charge dependent frequencies of motion in electrodynamic ion traps facilitates the execution of multistage mass spectrometry (MS^n) experiments that employ ion/ion reactions as well as other types of reactions. As a result, a wide variety of ion/ion reaction applications have been developed using ion traps. Most reported studies have employed conventional

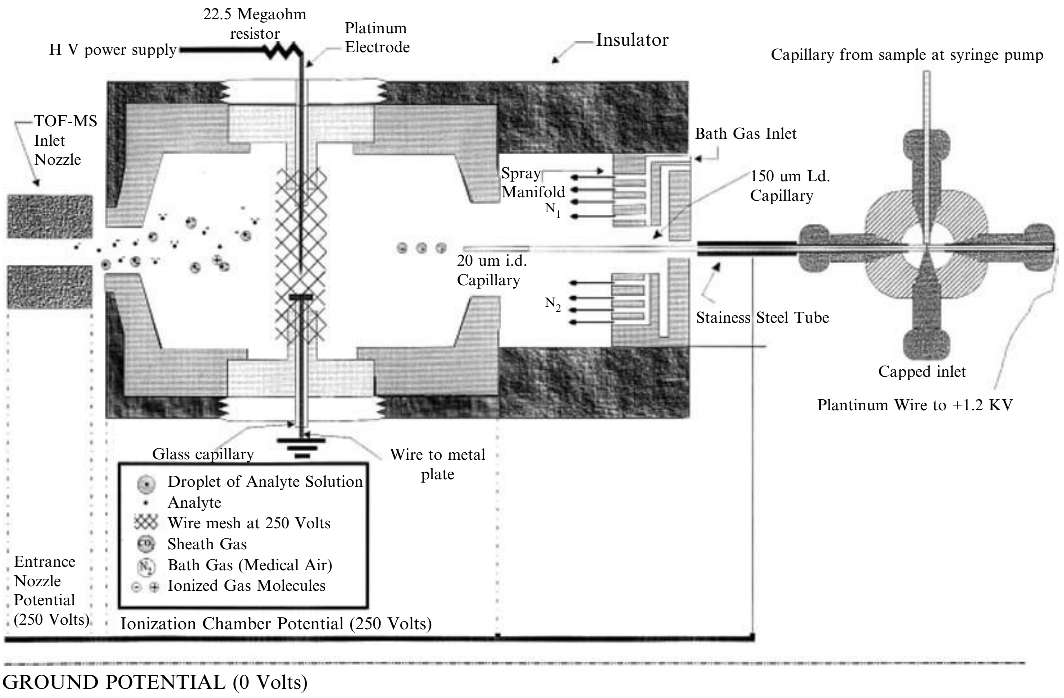

22.5 Megaohm resistor
H V power supply
Platinum Electrode
Insulator
Capillary from sample at syringe pump
TOF-MS Inlet Nozzle
Bath Gas Inlet
Spray Manifold
N_1
150 um I.d. Capillary
20 um i.d. Capillary
N_2
Stainess Steel Tube
Capped inlet
Plantinum Wire to +1.2 KV
Glass capillary
Wire to metal plate
Droplet of Analyte Solution
Analyte
Wire mesh at 250 Volts
Sheath Gas
Bath Gas (Medical Air)
Ionized Gas Molecules
Entrance Nozzle Potential (250 Volts)
Ionization Chamber Potential (250 Volts)
GROUND POTENTIAL (0 Volts)

Fig. 2. Schematic of an inductive charging device coupled with an ion/ion reaction chamber. Reprinted with permission from *(20)*. Copyright 2001 John Wiley and Sons Limited.

three-dimensional (3D) quadrupole ion traps and at least one commercial manufacturer offers a 3D ion trap capable of executing electron transfer ion/ion reaction experiments. Linear ion traps, which have advantages over 3D ion traps in ion storage capacity and ion trapping efficiency *(21, 22)*, have recently been adapted for ion/ion reactions, and commercial instruments capable of executing ion/ion reactions are being marketed. Examples of each type of electrodynamic ion trap modified for ion/ion reactions are discussed.

2.2.1. 3D Ion Traps

No modifications to the electrodes of a conventional 3D ion trap are necessary for conducting ion/ion reaction experiments. Means for forming and admitting ions of opposite polarity, of course, must be in place, which has stimulated the development of a variety of instrument configurations. The schematic of **Fig. 3** serves to summarize a range of approaches that have been taken.

Figure 3 shows openings to the ion trap via holes in the end-cap electrodes and via a hole in the ring electrode. A detector is generally mounted next to one of the end caps, often referred to as the exit end cap, to record ions ejected from the ion trap in the *z*-direction. The other end cap, often referred to as the entrance

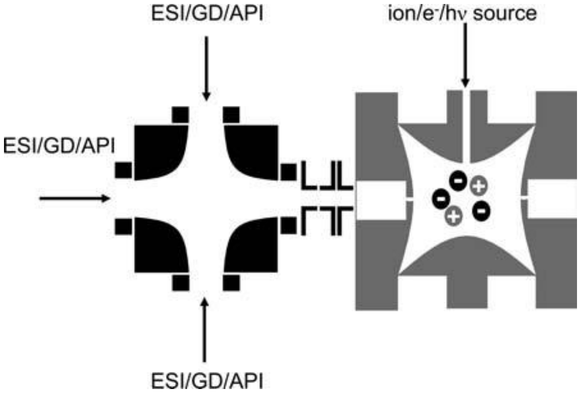


Fig. 3. Schematic of a 3D ion trap with a variety of means for admitting ions of opposite polarity.

end cap, is used to admit multiply charged reactant ions formed via ESI.

In early ion/ion reaction studies, the entrance end cap was used exclusively for the admission of analyte ions derived from ESI. Oppositely charged ions were generated by injection of electrons *(23)*, ions *(24)*, or light *(25)* into the ion trap through the ring electrode. Electron injection was effected by mounting a heated filament and gate electrode opposite the hole in the ring electrode and was used for in situ ionization of vapors present in the ion trap. This approach is most appropriate for the study of multiply charged anions with singly charged cations. The opposite polarity combination is more difficult to study due to the relative inefficiency of in situ negative ion formation via electron irradiation in the ion trap. Ion injection through the ring electrode is a relatively inefficient and energetic process. Fragmentation upon injection is much more extensive via ring-electrode injection than for end-cap injection. This tends to limit the range of reactant species that can be used to those that can be formed in relatively bright beams, to overcome the low ion-capture efficiency, and those that either fragment to give useful reagents or are highly stable. Anions derived from perfluorocarbons, such as perfluoro-1,3-dimethylcyclohexane (PDCH) and perfluoro methyl decalin (PMD), for example, can be readily formed via glow discharge ionization and their fragments react exclusively via proton transfer. Laser light has also been directed through a ring-electrode hole to ablate ions in the trap for subsequent ion/ion reactions *(25)*.

Means for allowing both positive ions and negative ions to be injected in series via the same end-cap electrode have also been described *(26–28)*. A turning quadrupole is shown in **Fig. 3**, which is intended to represent ion deflection techniques used for injection of both analyte and reactant ions into the 3D trap via a single end-cap electrode. In general, the use of ion injection through the end cap for both reactant and analyte ions provides greater flexibility in terms of the types of ions amenable to study and superior ion injection efficiency relative to ring-electrode injection.

2.2.2. Linear Ion Traps

Linear ion traps have become important tools in modern biological mass spectrometry both as stand-alone instruments and as components of hybrid mass spectrometers. For unipolar ion storage, the rod array provides trapping in the x- and y-directions while electrostatic trapping is used to store ions in the z-dimension. A pass-through ion/ion reaction approach has been described wherein the reagent anions were stored in a linear ion trap in the conventional way while analyte cations were transmitted through the device *(29)*. To effect mutual storage of ions of opposite polarity in the z-dimension requires operation of the linear ion trap in a nonconventional way. Two approaches have been demonstrated that result in the creation of a radiofrequency barrier for ion storage in the z-dimension. One involves the direct application of a radiofrequency voltage of amplitude as high as a few hundred volts *(10)* and the other involves the use of unbalanced trapping voltage applied to the quadrupole rod array *(30)*. The effect of using different rf amplitudes on the opposing rods of the quadrupole array is to create an effective rf barrier at the ends of the array. For either approach, it is desirable to be able to form or remove the z-dimension rf barrier in the course of an MS^n experiment.

Two general approaches have been taken to mass analysis from linear ion traps. One involves the radial ejection of ions *(21)* and the other involves axial ejection *(22)*. The radial ejection approach facilitates ion injection from opposing sides of the linear ion trap. Analyte ions can be injected via one side of the array while reagent ions are injected from the opposite side. To date, chemical ionization has been used to form singly charged reagent ions. A schematic of a linear ion trap that uses ion sources on either side (ESI for analyte ions and chemical ionization for reagent ions) of the rod array is shown in **Fig. 4**

The ability to control independently the DC offset to segments of the rod array in the device depicted in **Fig. 4** allows for improved control over timing of ion reactant mixing relative to approaches that permit ion mixing during the accumulation of the second ion population.

The location of the detector in linear ion traps that employ axial ejection can preclude the injection of ions from opposite ends of the rod array. To address this issue, pulsed ESI sources and/or pulsed ESI/pulsed corona discharge API approaches

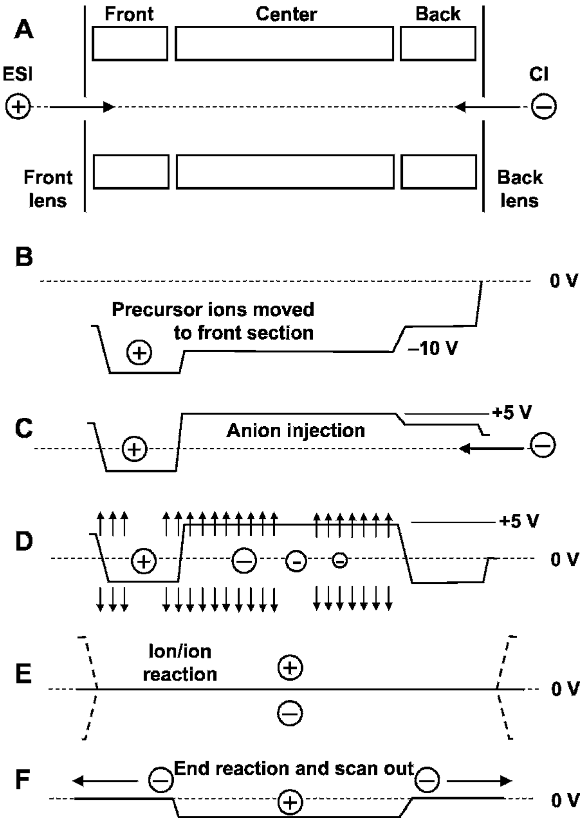

A
Front
Center
Back
ESI
CI
Front lens
Back lens
B
0 V
Precursor ions moved to front section
−10 V
C
Anion injection
+5 V
D
+5 V
0 V
E
Ion/ion reaction
0 V
F
End reaction and scan out
0 V

Fig. 4. Schematic of a linear ion trap with ion sources on either side of the rod array **a** along with the sequence of events for a typical ion/ion reaction experiment: **b** DC potentials applied to the rods to move the ions into a limited region of the rod array, **c** anion accumulation from the chemical ionization source, **d** isolation of the oppositely charged reactants with a tailored waveform, **e** mixing of reactants and mutual storage time, and **f** removal of anions prior to mass analysis. Reprinted with permission from *(10)*. Copyright 2004 National Academy of Sciences, USA.

have been implemented with a hybrid triple quadrupole/linear ion trap to allow for ion/ion reaction studies in either the relatively high-pressure Q2 collision cell (4–8 mtorr) or in the Q3 (<0.1 mtorr) linear ion trap. A schematic of the instrument with pulsed ESI sources is shown in Fig. 5. The advantage of pulsed sources is that the same ion path can be used for all ions. In the case of the hybrid triple quadrupole/linear ion trap instrument, the ion path includes a mass filter (Q1) that can be used for mass selection of both analyte and reagent ions. A key requirement for the use of pulsed sources in this fashion is the provision for software control of the various voltages required along the ion path during the course of a multistage experiment.

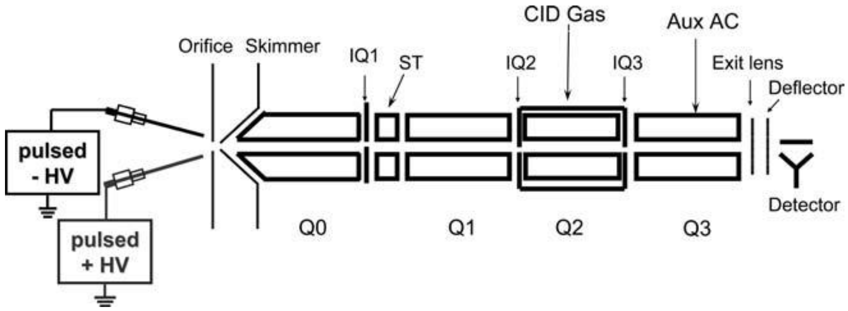


Fig. 5. Schematic of a hybrid triple quadrupole/linear ion trap mass spectrometer adapted with pulsed ESI sources and modified for ion/ion reaction studies.

3. Ion/Ion Reaction Methods in Electrodynamic Ion Traps

In contrast to the approaches described in Subheading 2.1, the *m*/*z* range over which ions can be stored simultaneously is an important issue in any ion trap-based ion/ion reaction experiment. The following equations are based on 3D ion traps *(31)* but analogous arguments hold for linear ion traps. It is a requirement that all ions, reactants, and products be stored below their respective exclusion limits,

$$q_z = \frac{8ZeV}{m\left(r_0^2 + 2z_0^2\right)\Omega^2} \leq 0.908, \qquad (1)$$

where r_0 and z_0 are the inscribed radius of the ring electrode and minimum spacing between an end-cap electrode and the center of the trap, respectively, m is the ion mass, Ze is the ion charge, Ω is the angular frequency of the drive or trapping voltage (typically applied to the ring electrode), and V is the amplitude of the drive voltage. The values of r_0, z_0, and Ω are fixed with most ion traps such that V is the experimental parameter that is used to control the q_z values of the ions. For a given set of values of V, r_0, z_0, and Ω, the mass-to-charge ratio at which $q_z = 0.908$ is often referred to as the low-mass cutoff (LMCO). Hence, the *m*/*z* ratios of all reactants and products must remain higher than the LMCO during the ion/ion reaction period. Care is also required during ion accumulation steps prior to the ion/ion reaction that no reactants fall below the LMCO in the course of accumulating or isolating another reactant.

The lower limit to the *m*/*z* range for mutual storage is established by the lowest *m*/*z* reactant and, in practice, is somewhat lower than the LMCO of this species. This is due to the fact that power absorption from the drive voltage can become significant as the ion approaches the exclusion limit and the resulting acceleration of the ion can inhibit the ion/ion reaction rate. The upper limit to the *m/z* range for mutual ion storage is less well-defined than the lower limit but it generally depends upon the effective trapping well, sometimes referred to as the pseudopotential well, for high *m/z* ions. The pseuopotential well *(31)* in the *z*-dimension, D_z, is given by

$$D_z = \frac{q_z V}{8}. \tag{2}$$

The pseudopotential trapping well in the z-dimension is twice that in the radial dimension, i.e.,

$$D_z = 2D_r. \tag{3}$$

The significance of the trapping wells is that they effectively define the highest kinetic energy that an ion can have without escaping the trapping field. For a given *V*, the trapping wells decrease with increasing *m/z* such that at some point even low-energy (e.g., thermal energy) ions can evaporate from the ion trap. Hence, for a given ion/ion reaction experiment, *V* must be sufficiently low to maintain a LMCO somewhat lower than the *m/z* ratio of the lowest *m/z* reactant and sufficiently high to store the highest *m/z* product. In practice, a mutual storage range as high as *m/z* 300–150,000 has been noted *(32)*. The *trapping by proxy* approach described later provides an exception to the considerations just mentioned and provides a means for use of relatively low *m/z* reagents under specialized circumstances.

Many ion/ion reaction experiments involve analyte species of relatively modest mass (e.g., <3,000 Da) and such experiments pose no issues for the mass analysis step. Ion/ion reactions can, however, give rise to products with mass-to-charge ratios in excess of 100,000. Commercial ion traps do not support *m/z* ranges in excess of about 10,000 but this is not due to a limitation in the physics of the device. Upper *m/z* limits in excess of 50,000 have been routinely accessed via use of resonance ejection at low q_z values *(33)*.

3.1. Single Proton Transfer Ion/Ion Reaction Experiments

Many variations of key experiments can be executed with the range of tools available for the study of ion/ion reactions. In this section, generic experiments are described with values that are typical of most instruments. In the case of proton transfer experiments, species that react exclusively by proton transfer are assumed. Relatively high-mass anions (e.g., >330 Da) derived from perfluorocarbons are particularly useful as proton transfer reagents because they are readily formed in high numbers, they react with peptides and proteins via proton transfer, and they allow for storage of high mass-to-charge ratio products.

3.1.1. Protein and Product Ion Mixtures

Proton transfer reaction experiments are particularly useful for simplifying ESI spectra of protein mixtures where spectral congestion can complicate mass assignments. They have also been applied to mixtures of product ions derived from the dissociation of a multiply charged parent ion to facilitate product ion mass assignment. In both cases, an excess of reagent ion charge (usually >10×) is accumulated in the ion trap to maximize reaction rates and to ensure that all analyte ions react. The following steps summarize the generic protein mixture experiment in an electrodynamic ion trap:

1. Analyte ion accumulation at optimum LMCO for analyte ions (10–1,000 ms)
2. Reagent ion accumulation at a useful LMCO (10–50 ms)
3. Mutual storage period at as high a useful LMCO as possible (100 ms)
4. Removal of residual anions (via a ramp of the LMCO or resonance ejection)
5. Mass analysis

The MS/MS experiment is similar but also includes analyte ion isolation and activation steps:

1. Positive ion accumulation at optimum LMCO for analyte (10–1,000 ms)
2. Analyte ion precursor selection
3. Ion activation/dissociation (CID, IRMPD, ETD…)
4. Anion accumulation at a useful LMCO (10–50 ms)
5. Mutual storage period at as high a useful LMCO as possible (100 ms)
6. Removal of residual anions (via a ramp of the LMCO or resonance ejection)
7. Mass analysis

Note that the removal of residual anions is usually a necessary step because the excess charge associated with the reagent ions can influence the motions of the product analyte ions and thereby interfere with the mass analysis step *(34)*.

The LMCO during the ion/ion reaction period strongly influences the ability to store high *m/z* ions, as discussed earlier. However, if a low-mass reagent ion can be accumulated in the ion trap rapidly, to the point where the electric field associated with the reagent ion can prevent high-mass analyte product ions from escaping the ion trap, the upper *m/z* limit defined by the well depth is no longer relevant because it is the electric field of the reagent ion that retains the analyte ion. This phenomenon, which has been referred to as *trapping by proxy (35)*, relies on the electrodynamic field of the ion trap to store the low-mass ions and the electric field of the stored low-mass ions to trap the high-mass ions, rather than requiring the electrodynamic field to perform both functions. This technique has been demonstrated with mixtures of oligonucleotides in the negative ion mode where $O_2^{+\cdot}$ and $C_4H_9^+$ ions, which can be formed in very bright beams via glow discharge, were demonstrated as reagent ions for electron transfer and proton transfer, respectively. The steps of a trapping by proxy experiment are similar to those for any other mixture analysis scenario except that the reagent ion accumulation must be fast and the reagent ion removal step must include a rapid increase in V to provide an adequate well depth for the analyte ions once the reagent ions have been removed.

3.1.2. Ion Parking

It is possible to inhibit selectively the ion/ion reactions that take place in an ion trap by accelerating one of the reactants in a process referred to as *ion parking (36)*. Several variations of ion parking have been described. In the case of proteins with multiple charge states, it is possible to selectively inhibit the reaction of a selected charge state by applying a relatively low amplitude supplementary frequency in resonance, or nearly so, with the fundamental z-dimension secular frequency of a charge state of interest. It is important that the reagent ion numbers be maintained below the point where the electric field of the reagent ions significantly affects the analyte ion motion. The reaction rates of all other charge states will be unaffected under these conditions. If a single analyte ion species is subjected to ion acceleration in this way and there are higher charge states present in the ion trap, the analyte ions of the higher charge states will eventually react to the point where they will fall into resonance with the applied signal. In this way, it is possible to concentrate much of the signal originally dispersed over multiple charge states largely into a single charge state. The steps for a single ion parking experiment are as follows:

1. Analyte ion accumulation at optimum LMCO for analyte (10–1,000 ms).
2. Reagent ion accumulation at a useful LMCO (5–20 ms). It is important to avoid accumulation of sufficient reagent ions numbers to affect the analyte ion motion.
3. Mutual storage period at as high a useful LMCO as possible with simultaneous application of a tailored waveform to effect ion parking (reaction time = 10–100 ms). For single-frequency ion parking, a low-amplitude voltage (a few tens of millivolts) is used to inhibit the reaction rate.
4. Removal of residual anions (via a ramp of the LMCO or resonance ejection).
5. Mass analysis.

In the case of complex protein mixture analysis, it is possible that a single ion parking step will accumulate two or more proteins with charge states that share the same *m/z* ratio. In this scenario, it is possible to effect a charge state purification step with a second ion parking experiment that inhibits the reaction of a lower charge state of the protein of interest. Any other proteins present in the original parked population will not be parked in the second step. The sequential ion parking experiment allows for the gas-phase concentration and charge state purification of a protein for subsequent ion activation *(6)*.

The use of tailored waveforms *(37)* applied to the end-cap electrodes, a single high-amplitude voltage applied to the end-cap electrodes *(38)*, or the application of a DC potential across the end-cap electrodes to separate oppositely charged ion clouds *(39)* can be used to inhibit the reaction rates of multiple analyte ions simultaneously. These supplementary voltages would be applied in earlier **step 3**. The simultaneous inhibition of the reaction rates of ions of different *m/z* values is referred to as *parallel ion parking*. Several applications have been described for parallel ion parking *(37, 38)*.

3.2. Electron Tansfer

The experimental steps associated with an electron transfer ion/ion reaction are the same as those for proton transfer. The difference between the experiments is in the nature of the reagent ion *(40)*. In the case of multiply protonated peptides and proteins, proton transfer to a singly charged negative ion is always a potentially competitive reaction channel. The criteria for a reagent to show significant probability for transfer of an electron to a multiply protonated peptide or protein are favorable Franck-Condon factors associated with the transition from the anionic to neutral form of the reagent species and an electron affinity that is not too high (i.e., greater than about 3 eV) *(41)*. Anionic species from a variety of reagent molecules have been shown to transfer an electron to a polypeptide cation with anions derived from fluoranthene and from diazobenzene showing a strong preference for electron transfer over proton transfer. These anions, and most others shown to react via electron transfer are formed either by conventional negative chemical ionization or negative atmospheric pressure chemical ionization. The steps for a typical ETD experiment are as follows:

1. Positive ion accumulation at optimum LMCO for analyte (10–1,000 ms)
2. Analyte ion precursor selection/isolation
3. Anion accumulation/isolation at a useful LMCO (10–50 ms)
4. Mutual storage period (100 ms)
5. Removal of residual reagent ions

6. Possible post-ion/ion activation step for ETD survivor ions
7. Possible reagent ion accumulation step for proton transfer
8. Mutual ion storage for ion/ion proton transfer involving ETD products
9. Removal of residual reagent ions
10. Mass analysis

Steps 1–5 and 7 comprise a standard ETD experiment. In **Fig. 4**, parts b–f describe how some of these steps are executed in the linear ion trap of **Fig. 4**. The procedure summarized earlier highlights the MS^n capabilities of modern electrodynamic ion traps by including a possible subsequent ion activation step for electron transfer survivor peptides *(42)* and a second ion/ion reaction period to subject the ETD products to ion/ion proton transfer reactions. A combination of ETD and ion/ion proton transfer reactions for top–down protein identification *(43)* has been described that employed many of the steps listed earlier.

3.3. Other Reaction Types

The methods associated with single proton transfer and single electron transfer are emphasized here for their clear utility in protein and peptide analysis. A number of other types of ion/ion reactions have also been observed that promise to be useful in the context of peptide and protein analysis. These include multiple proton transfers in a single ion/ion collision that leads to charge inversion of the peptide or protein ion *(7–9)* and metal ion transfer reactions that replace protons with a metal ion *(44–46)*. Reagents for charge inversion and metal ion transfer have been formed via ESI, and experiments taking advantage of these phenomena can be conducted with instruments of the type depicted in **Figs. 3** and **5** . Experimental procedures similar to those described earlier apply to charge inversion and metal transfer experiments.

4. Prognosis

The remarkably high experimental flexibility of electrodynamic ion traps along with the high dimensionality associated with gas-phase ion/ion reactions make it likely that many new types of ion/ion reaction experiments will be developed in the coming years. The facility with which linear ion traps can be coupled with other MS technologies makes straightforward the development of interesting hybrid instrumentation. The implementation of ion/ion reactions on such hybrid systems, the hybrid triple quadrupole/linear ion trap and quadrupole/linear ion trap/time-of-flight instruments *(47)* are just two examples, will add to the range of applications to which ion/ion reactions can be directed.

A key feature associated with the implementation of ion/ion reactions in electrodynamic ion traps is the control over when and where the reactions can take place within the context of an MS^n experiment. The robust nature of the chemistry makes practical the use of multiple ion/ion reactions in a single experiment. Hence, the potential number of experiments is remarkably large.

Acknowledgment

This work was supported by the National Institutes of Health, Institute of General Medical Sciences, under Grant GM 45372.

References

1. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M. (1990) Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* 246, 64–71
2. FennJr., J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M. (1990) Electrospray Ionization – Principles and Practice. *Mass Spectrom. Rev.* 9, 37–70
3. Stephenson, J.L., McLuckey, S.A. (1996) Ion/Ion Proton Transfer Reactions for Protein Mixture Analysis. *Anal. Chem.* 68, 4026–4032
4. Scalf, M., Westphall, M.S., Krause, J., Kaufman, S.L., Smith, L.M. (1999) Controlling Charge States of Large Ions. *Science* 283, 194–197
5. Stephenson, J.L., McLuckey, S.A. (1998) Simplification of Product Ion Spectra Derived from Multiply-Charged Parent Ions via Ion/Ion Chemistry. *Anal. Chem.* 70, 3533–3544
6. Reid, G.E., Shang, H., Hogan, J.M., Lee, G.U., McLuckey, S.A. (2002) Gas-Phase Concentration, Purification, and Identification of Whole Proteins from Complex Mixtures. *J. Am. Chem. Soc.* 124, 7353–7362
7. He, M., Emory, J.F., McLuckey, S.A. (2005) Reagent Anions for Charge Inversion of Polypeptide/Protein Cations in the Gas Phase. *Anal. Chem.* 77, 3173–3182
8. He, M., McLuckey, S.A. (2003) Two Ion/ion Charge Inversion Steps to form a Doubly-protonated Peptide from a Singly-protonated Peptide in the Gas Phase. *J. Am. Chem. Soc.* 125, 7756–7757
9. He, M., McLuckey, S.A. (2004) Increasing the Negative Charge of a Macro-Anion in the Gas-Phase Via Sequential Charge Inversion Reactions. *Anal. Chem.* 76, 4189–4192
10. Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F. (2004) Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. USA* 101, 9528–9533
11. McLuckey, S.A.; StephensonJr., J.L. (1998) Ion/Ion Chemistry of High-Mass Multiply Charged Ions. *Mass Spectrom. Rev.* 17, 369–407
12. Pitteri, S.J., McLuckey, S.A. (2005) Recent Developments in the Ion/Ion Chemistry of High Mass Multiply Charged Ions. *Mass Spectrom. Rev.* 24, 931–958
13. McLuckey, S.A. (2006) Biomolecule Ion–Ion Reactions. In: “Principles of Mass Spectrometry Applied to Biomolecules” J. Laskin (Ed.), Wiley, Hoboken, NJ, Chapter 15
14. Loo, R.R.O., Usdeth, H.R., Smith, R.D. (1991) Evidence of Charge Inversion in the Reaction of Singly Charged Anions with Multiply Charged Macro-ions. *J. Phys. Chem.* 95, 6412–6415
15. Loo, R.R.O., Udseth, H.R., Smith, R.D. (1992) A New Approach for the Study of Gas-Phase Ion–Ion Reactions Using Electrospray Ionization. *J. Amer. Soc. Mass Spectrom.* 3, 695–705
16. Kaufman, S.L., Skogen, J.W., Dorman, F.D., Zarrin, F., Lewis, K.C. (1996) Macromolecule analysis based on electrophoretic mobility in air: Globular proteins. *Anal. Chem.* 68, 1895–1904
17. Scalf, M., Westphall, M.S., Smith, L.M. (2000). Charge Reduction Electrospray Mass Spectrometry. *Anal. Chem.* 72, 52–60

18. Ebeling, D.D., Westphall, M.S., Scalf, M., Smith, L.M. (2000) Corona Discharge in Charge Reduction Electrospray Mass Spectrometry. *Anal. Chem.* 72, 5158–5161
19. Wang, H., Hackett, M. (1998) Ionization within a Cylindrical Capacitor: Electrospray without an Externally Applied High Voltage. *Anal. Chem.* 70, 205–212
20. Ebeling, D.D., Westphall, M.S., Scalf, M., Smith, L.M. (2001) A Cylindrical Capacitor Ionization Source: Droplet Generation and Controlled Charge Reduction for Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 15, 401–405
21. Schwartz, J.C., Senko, M.W., Syka, J.E.P. (2002) A Two Dimensional Ion Trap Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* 13, 659–669
22. Hager, J.W. (2002) A New Linear Ion Trap Mass Spectrometer. *Rapid Commun. Mass Spectrom.* 16, 512–526
23. Herron, W.J., Goeringer, D.E., McLuckey, S.A. (1995) Ion–Ion Reactions in the Gas-Phase: Proton Transfer Reactions of Protonated Pyridine with Multiply-Charged Oligonucleotide Anions. *J. Am. Soc. Mass Spectrom.* 6, 529–532
24. StephensonJr., J.L., McLuckey, S.A. (1997) Adaptation of the Paul Trap for Study of the Reaction of Multiply Charged Cations with Singly Charged Anions. *Int. J. Mass Spectrom. Ion Processes* 162, 89–106
25. Payne, A.H., Glish, G.L. (2001) Gas-Phase Ion/Ion Interactions Between Peptides or Proteins and Iron Ions in a Quadrupole Ion Trap. *Int. J. Mass Spectrom.* 204, 47–54
26. Wells, J.M., Chrisman, P.A., McLuckey, S.A. (2002) Dueling Electrospray: Instrumentation to Study Ion/Ion Reactions of Electrospray-generated Cations and Anions. *J. Am. Soc. Mass Spectrom.* 13, 614–622
27. Badman, E.R., Chrisman, P.A., McLuckey, S.A. (2002) A Quadrupole Ion Trap Mass Spectrometer with Three Independent Ion Sources for the Study of Gas-Phase Ion/Ion Reactions. *Anal. Chem.* 74, 6237–6243
28. Hartmer, R.; Lubeck, M. (2005) New Approach for Characterization of Post-Translationally Modified Peptides using Ion Trap MS with Combined ETD-CID Fragmentation. *LCGC – The Application Notebook* 0905, 14
29. Wu, J., Hager, J.W., Xia, Y., Londry, F.A., McLuckey, S.A. (2004) Positive Ion Transmission Mode Ion/Ion Reactions in a Hybrid Linear Ion Trap. *Anal. Chem.* 76, 5006–5015
30. Xia, Y., Wu, J., Hager, J.W., Londry, F.A., McLuckey, S.A. (2005) Mutual Storage Mode Ion/Ion Reactions in a Hybrid Linear Ion Trap. *J. Am. Soc. Mass Spectrom.* 16, 71–81
31. March, R.E., Todd, J.F.J. (2005) Quadrupole Ion Trap Mass Spectrometry, 2nd Edition, Wiley-Interscience, New York, 2005, Chapters 2 and 3
32. Reid, G.E., Wells, J.M., Badman, E.R., McLuckey, S.A. (2003) Performance of a Quadrupole Ion Trap Mass Spectrometer Adapted for Ion/Ion Reaction Studies. *Int. J. Mass Spectrom.* 222, 243–258
33. Kaiser, R.E., Cooks, R.G., Stafford, G.C., Syka, J.E.P., Hemberger, P.H. (1991) Operation of a Quadrupole Ion Trap Mass Spectrometer to Achieve High Mass-to-Charge Ratios. *Int. J. Mass Spectrom. Ion Processes* 106, 79–115
34. StephensonJr.,, J.L., McLuckey, S.A. (1997) Anion Effects on Storage and Resonance Ejection of High Mass-to-Charge Cations in Quadrupole Ion Trap Mass Spectrometry. *Anal. Chem.* 69, 3760–3766
35. McLuckey, S.A., Wu, J., Bundy, J.L., StephensonJr., J.L., Hurst, G.B. (2002) Oligonucleotide Mixture Analysis via Electrospray and Ion/Ion Reactions in a Quadrupole Ion Trap. *Anal.Chem.* 74, 976–984
36. McLuckey, S.A., Reid, G.E., Wells, J.M. (2002) Ion Parking During Ion/ion Reactions in Electrodynamic Ion Traps. *Anal. Chem.* 74, 336–346
37. Chrisman, P.A., Pitteri, S.J., McLuckey, S.A. (2005) Parallel Ion Parking: Improving Conversion of Parents to First Generation Products in Electron Transfer Dissociation. *Anal. Chem.* 77, 3411–3414
38. Chrisman, P.A., Pitteri, S.J., McLuckey, S.A. (2006) Parallel Ion Parking of Protein Mixtures. *Anal. Chem.* 78, 310–316
39. Grosshans, P.B., Ostrander, C.M., Walla, C.A. Methods and Apparatus to Control Charge Neutralization Reactions in Ion Traps. *U.S. Patent* 6,674, January 6, 2004067 B2
40. Coon, J.J., Syka, J.E.P., Schwartz, S.C., Shabanowitz, J., Hunt, D.H. (2004) Anion Dependence in the Partitioning between Proton Transfer and Electron Transfer in Ion/Ion Reactions. *Int. J. Mass Spectrom.* 236, 33–42
41. Gunawardena, H.P., He, M., Chrisman, P.A., Pitteri, S.J., Hogan, J.M., Hodges, B.D.M., McLuckey, S.A. (2005) Electron Transfer versus Proton Transfer in Gas-Phase Ion–Ion Reactions of Polyprotonated Peptides. *J. Am. Chem. Soc.* 127, 12627–12639

42. Swaney, D.L., McAlister, G.C., Wirtala, M., Schwartz, J.C., Syka, J.E.P., Coon, J.J. (2007) A Supplemental Activation Method for High Efficiency Electron Transfer Dissociation of Doubly Protonated Peptide Precursors. *Anal. Chem.* 79, (2) 477–485
43. Coon, J.J., Ueberheide, B., Syka, J.E.P., Dryhurst, D.D., Ausio, J., Shabanowitz, J., Hunt, D.F. (2005) Protein Identification Using Sequential Ion/Ion Reactions and Tandem Mass Spectrometry. *Proc. Natl Acad. Sci. USA* 102, 9463–9468
44. Newton, K.A., McLuckey, S.A. (2003) Gas-Phase Peptide/Protein Cationizing Agent Switching Via Ion/Ion Reactions. *J. Am. Chem. Soc.* 125, 12404–12405
45. Newton, K.A., Amunugama, R., McLuckey, S.A. Gas-Phase Ion/Ion Reactions of Multiply Protonated Polypeptides with Metal Containing Anions. J. Phys. Chem A. 109, 3608-3616
46. Gunawardena, H., O'Hair, R.A.J., Mcluckey, S.A. (2006) Selective Disulfide Bond Cleavage in Gold (I) Cationized Polypeptide Ions formed via Gas-Phase Ion/Ion Cation Switching. *J. Proteome Res.* 5, 2087–2092
47. Xia, Y., Chrisman, P.S., Erickson, D.E., Liu, J., Liang, X., Londry, F.A., Yang, M.J., McLuckey, S.A. (2006) Implementation of Ion/Ion Reactions in a Quadrupole/Time-of-Flight/Tandem Mass Spectrometer. *Anal. Chem.* 78, 4146–4154

Chapter 25

Electron Capture Dissociation LC/MS/MS for Bottom–Up Proteomics

Roman A. Zubarev

Summary

ECD is a fragmentation technique that exhibits unusual properties. It is useful in bottom–up proteomics as a complement to collisional dissociation for deriving additional sequence information, locating posttranslational modifications, and revealing racemization of individual amino acids in a peptide sequence.

Key words: Peptide fragmentation, Protein identification, Posttranslational modifications.

1. Introduction

Electron capture dissociation (ECD) is a method of deriving primary structure of polypeptides in the gas phase by reacting their multiply charged cations with low-energy electrons *(1)*. The advantages of performing ECD vs. conventional low-energy collisional dissociation on tryptic peptides are the following. First, ECD breaks N–C_α bond instead of the peptide C–N bond, providing abundant and complementary backbone cleavages *(2)*. Second, ECD has different amino acid preferences than collisional activation and thus yields complementary sequence information *(3)*. Third, ECD is a *soft* fragmentation that preserves labile posttranslational modifications, such as phosphorylation *(4)*, glycosylation *(5)*, and sulfation *(6)*, thus revealing their position in the sequence. At the same time, radical character of C-terminal ECD fragments leads to secondary reactions that include specific losses from Leu and Ile by means of which these isomeric residues can be distinguished *(7)*. ECD breaks disulfide bonds specifically, unlike collisional dissociation *(8)*. ECD tends to neutralize the positive charge located in the least basic

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_25

site, and can reveal both charge location and the basicity order of the charged sites *(9)*. Finally, ECD is sensitive to secondary structure and thus can distinguish partially racemized sequences from their monochiral analogues *(10)*. In principle, ECD can work with any type of mass spectrometer. In practice, ECD is easiest to implement in Fourier transform ion cyclotron resonance mass spectrometry (FT ICR). ECD is fast enough to work on chromatographic mass scale *(11)*. In proteomics, ECD is recommended for use as a complement to collisional dissociation for protein identification *(12)* and de novo sequencing *(13)*.

2. Materials

2.1. Mass Spectrometry

1. Finnigan LTQ FT mass spectrometer with a 7-T superconducting magnet (Thermo Fisher Scientific, Inc.).
2. Nanoelectrospray ion source from Proxeon Biosystems (Odense, Denmark).
3. Electron source for ECD from Thermo based on an indirectly heated dispenser cathode (Heatwave Labs, Watsonville, CA).

2.2. Liquid Chromatography

1. 1100 HPLC system (Agilent, CA) including a solvent degasser, nanoflow pump, and thermostated microautosampler.
2. Prepare analytical column by packing 15-cm fused silica emitter (75-μm I.D., 375-μm O.D.; Proxeon Biosystems) with a methanol slurry of reversed-phase, fully end-capped Reprosil-Pur C18-AQ 3-μm resin (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany), using a pressurized *packing bomb* (Proxeon Biosystems) at 50–60 bar of He pressure.
3. Mobile phase (A) 0.5% acetic acid and 99.5% water (v/v); (B) 0.5% acetic acid and 10% water in 89.5% acetonitrile (v/v).

3. Methods

Inclusion of ECD in a bottom–up LC/MS/MS experiment performed on a modern FT ICR mass spectrometer is straightforward. It is recommended to spend some time on off-line ECD efficiency optimization and then running ECD LC/MS/MS on a known mixture (e.g., BSA digest) before analytical work on unknown mixtures (*see* **Note 1**). Data processing depends upon the search engine used. Optimal case is submission of collisional activation and ECD datasets for simultaneous searching in the database.

3.1. Data-Dependent Analysis

1. ECD event duration: 75 ms; *amplitude* – 5.
2. Add ECD to data-dependent acquisition script that includes one survey scan (resolution = 100,000) and two double MS/MS events for most abundant ions with z 2: first ECD, then collisional dissociation. Resolution in both MS/MS events 25,000 (*see* **Note 2**).
3. Put the mass of the fragmented ion on the exclusion list for 60 s.
4. Load 8 µL of digested peptide mixture onto the column and rinse for 20 min in 4% buffer B at a flow rate of 500 nL/min.
5. Run a 90-min gradient from 4 to 45% buffer B at a constant flow rate of 200 nL/min.

3.2. Data Processing

1. Convert MS/MS data from the raw LC/MS/MS datafile to .dta files using BioWorks 3.2 (Thermo).
2. Merge collisional activated MS/MS data and ECD data for the same peptide into a single .dta file as described in *(12)* (*see* **Note 3**).
3. Submit merged .dta files to Mascot search engine (v. 2.1, Matrix Science, London, UK). Search parameters: trypsin specificity, mass tolerance 5 ppm for MS and ±0.02 Da for MS/MS; instrument setting *ESI-FTICR*, type of fragmentation *CID*; possible fragments: *b* and *y*.

4. Notes

1. Other parameters of the electron source are best tuned using Xcalubur software (Thermo) with 2+ ions of substance P as test species. Efficiency of ECD should be >12%.
2. Two microscans (integrated transients before FT) in ECD identify more peptides than one microscan.
3. As an alternative to merged MS/MS data, ECD data can be submitted as *ECD* type of fragmentation. Include *a, c, y, z*, and *w* ions in the list of possible fragments.

Acknowledgments

The author is grateful to Michael L. Nielsen and Mikhail M. Savitski for help in implementing ECD in proteomics. This work was supported by the Knut and Alice Wallenberg Foundation and Wallenberg Consortium North (WCN2003-UU/SLU-009 and instrumental grant to RZ and Carol Nilsson) as well as the Swedish research council (621-2004-4897, 621-2002-5025, and 621-2003-4877).

References

1. Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron capture dissociation of multiply charged protein cations. A non-ergodic process. *J. Am. Chem. Soc.* 120, 3265–3266
2. McLafferty, F. W., Fridriksson, E. K., Horn, D. M., Lewis, M. A., and Zubarev, R. A. (1999) Biomolecule mass spectrometry. *Science*, 284, 1289–1290
3. Savitski, M. M., Kjeldsen, F., Nielsen, M. L., and Zubarev, R. A. (2006) Complementary sequence preferences of electron capture dissociation and vibrational excitation in fragmentation of polypeptide poly-cations. *Angew. Chem.* 45, 5301–5303
4. Stensballe, A., Jensen, O. N., Olsen, J. V., Haselmann, K. F., and Zubarev, R. A. (2000) Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Commun. Mass Spectrom.* 14, 1793–1800
5. Mirgorodskaya, E., Roepstorff, P., and Zubarev, R. A. (1999) Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal. Chem.* 71, 4431–4436
6. Haselmann, K. F., Budnik, B. A., Olsen, J. V., Nielsen, M. L., Reis, C. A., Clausen, H., et al. (2001) Advantages of external accumulation for electron capture dissociation in Fourier transform mass spectrometry. *Anal. Chem.* 73, 2998–3005
7. Kjeldsen, F., Budnik, B. A., Haselmann, K. F., Jensen, F., and Zubarev, R. A. (2002) Dissociative capture of hot (3–13 eV) electrons by polypeptide polycations: an efficient process accompanied by secondary fragmentation. *Chem. Phys. Lett.* 356, 201–206
8. Zubarev, R. A., Kruger, N. A., Fridriksson, E. K., Lewis, M. A., Horn, D. M., Carpenter, B. K., et al. (1999) Electron capture dissociation of gaseous multiply-charged proteins is favored at disulfide bonds and other sites of high hydrogen atom affinity. *J. Am. Chem. Soc.* 121, 2857–2862
9. Kjeldsen, F., Adams, C. M., Savitski, M. M., and Zubarev, R. A. (2006) Localization of positive charges in gas-phase polypeptide polycations by tandem mass spectrometry. *Int. J. Mass Spectrom.* 252, 204–212
10. Adams, C. and Zubarev, R. A. (2005) Distinguishing and quantification of peptides and proteins containing d-Amino acids by tandem mass spectrometry. *Anal. Chem.* 77, 4571–4580
11. Palmblad, M., Tsybin, Y. O., Ramstrom, M., Bergquist, J., and Hakansson, P. (2002) Liquid chromatography and electron-capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* 16, 988–992
12. Nielsen, M. L., Savitski, M. M., and Zubarev, R. A. (2005) Improving protein identification using complementary fragmentation techniques in Fourier transform mass spectrometry. *Mol. Cell. Proteomics*, 6, 835–845
13. Savitski, M. M., Nielsen, M. L., Kjeldsen, F., and Zubarev, R. A. (2005) Proteomics-grade de novo sequencing approach. *J. Proteome Res.* 4, 2348–2354

Chapter 26

Two-Dimensional Ion Mobility Analyses of Proteins and Peptides

Alexandre A. Shvartsburg, Keqi Tang, and Richard D. Smith

Summary

Ion mobility separations have emerged as a major tool for mass spectrometry of proteins and peptides. The high speed of ion mobility spectrometry (IMS) compared with chromatography enables accelerating proteomic analyses at same separation power or raising the peak capacity at equal throughput. Of interest to structural biology, tractable physics of ion transport in gases permits characterizing the structure of macromolecules by matching measured mobilities to values calculated for candidate geometries. The two known experimental methods are drift-tube IMS based on absolute mobility and field asymmetric waveform IMS (FAIMS) based on differential mobility as a function of electric field. Here, we describe combining them into 2D separations coupled to time-of-flight MS, a development made practical by electrodynamic ion funnel interfaces that effectively convey ions in and out of IMS, including "hourglass" funnels to accumulate ions filtered by FAIMS between pulsed injections into IMS. For peptide separations, the peak capacity of FAIMS/IMS is ~500 and potentially higher, a metric close to that of top capillary LC systems. In structural investigations, FAIMS/IMS allows more protein conformers to be distinguished than either stage alone, and extends the dynamic range of detection by an order of magnitude over 1D IMS. A controlled heating of ions by rf field over a variable time in the funnel trap between FAIMS and IMS stages allows following the evolution of selected isomers in both thermodynamic and kinetic aspects, which opens a new approach to mapping the pathways and energy surfaces of protein folding.

Key words: Ion mobility spectrometry, FAIMS, 2-D separations, Time-of-flight MS, Ion funnels.

1. Introduction

Analytical methods based on ion transport in gases under the influence of electric field, known collectively as ion mobility spectrometry (IMS), add major new capabilities to mass-spectrometric investigations. Separation of ion mixtures and characterization of ions by IMS are of particular utility in biological mass

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_26

spectrometry (MS), where the sample complexity creates a virtually open-ended demand for separation power *(1)* while the ion size normally rules out structural elucidation by spectroscopic means. The use of IMS in mass spectrometry of proteins and peptides has expanded dramatically since the previous edition of this volume in 2000.

One major development was the advent of field asymmetric waveform IMS (FAIMS) as a practical bioanalytical tool, enabled by coupling FAIMS to electrospray ionization (ESI) sources *(2)*. In conventional IMS, ions are pulled through a tube filled with a nonreactive buffer gas by a moderate electric field, E, established by voltages on surrounding electrodes. This separates ions by the quantity known as mobility (K), defined as

$$K = v/E, \tag{1}$$

where v is the ion drift velocity *(3)*. The dependence of K on the gas number density (N) is removed by introducing the reduced mobility, K_0, defined as K for standard N (at 0°C and 1 atm.), N_0. The mobility of any ion may be expressed *(3)* as a series over even powers of E:

$$K(E) = K(0) \times (1 + aE^{2+} bE^{4} + \ldots). \tag{2}$$

In IMS, E is usually low enough that in essence $K(E) = K(0)$. The difference becomes significant at high E (in general, $E/N > 40$ Townsend, meaning ~1 kV/mm at N_0). This provides the basis for FAIMS (also called differential mobility spectrometry – DMS, or ion mobility increment spectrometry – IMIS) that sorts ions by the difference between K at high and low E *(4)*. In FAIMS, ions entrained in a gas pass through a gap between two electrodes carrying a periodic waveform $U_D(t)$ with peak voltage (the dispersion voltage, DV) creating E (the dispersion field, DF) sufficient for K to materially differ from $K(0)$. The $U_D(t)$ is asymmetric: while its mean is zero, positive and negative segments create different average E in the gap *(4)*. Then the ion displacements during those segments are unequal, and all ions in the gap move toward an electrode where they are neutralized. Selected ion species can be prevented from drifting toward either electrode and hence permitted to pass FAIMS by a constant compensation field (CF) due to compensation voltage (CV) applied to one of the electrodes *(4)*. Scanning CV yields a FAIMS spectrum of ions present. The FAIMS performance strongly depends on the gap geometry, and cylindrical, planar, and spherical geometries have been implemented *(2, 4–6)*.

Depending on the signs of coefficients a and b in **Eq. 2**, the slope of $K(E)$ can be positive (termed A-type ions) or negative (C-type ions) *(2)*. For any ion, those coefficients and thus the type depend on the composition and temperature of buffer gas *(7)*. In general, CF does not correlate with the ion mass/charge (*m/z*) ratio tightly *(8, 9)*. That renders FAIMS and MS dimensions substantially orthogonal, which is a major attraction of FAIMS/MS platform for analyses of complex samples. In contrast, *K* is closely related to the size and thus mass of an ion, hence a correlation between the IMS and MS dimensions for ions of a particular *z* – a fundamental limitation of IMS/MS *(10)*. As a result, FAIMS can often separate isomers not distinguishable by IMS *(6, 11)*. This capability is important to the MS studies of proteins and peptides, which increasingly focus on the variations of tertiary/quaternary structure and biomedically significant misfolding such as implicated in Alzheimer's and Parkinson's diseases, prion diseases, and other neurological disorders *(12, 13)*.

Another seminal step was the development of IMS/time-of-flight (ToF) MS systems. While IMS/(quadrupole MS) platforms had been used and even commercialized since 1970s *(14)*, they were analytically impractical because of low-duty cycle of quadrupole MS as a scanning technique, exacerbated by huge ion losses in the IMS stage. In the IMS/ToF arrangement, usually MS analyses take ~0.1 ms and IMS separations are ~10^2–10^3 times longer. Considering that the resolving power (*R*) of IMS is <150, the MS dimension may be folded into the IMS with ToF mass-dispersing ion packets as they appear from IMS at its normal speed *(15)*. This provides simultaneous IMS and MS analyses of all sampled ions.

Rapid technological improvement and expanding applications of FAIMS and IMS had culminated in the introduction of several FAIMS/MS and IMS/ToF products by major vendors in 2005–2007. These products included a FAIMS unit configurable with any mass spectrometer, triple quadrupole (TSQ Quantum) and ion trap (LTQ) systems with intergrated FAIMS upfront (Thermo Fischer Scientific, San Jose, CA) *(16)* and the Synapt HDMS ToF comprising an IMS stage (Waters, Milford, MA) *(17)*. The emergence of commercial FAIMS/MS and IMS/ToF systems, though short of the top benchmarks in terms of ion mobility resolution, evidences the maturity of FAIMS and IMS for regular laboratory use and will accelerate their adoption by mainstream analytical community.

1.1. Constraints on the Separation Power of IMS and FAIMS

Commercial FAIMS/MS systems employ cylindrical FAIMS geometries, which limits *R* to ~10–20 *(5, 6, 18, 19)*. For example, the CF range for tryptic peptides (in 1:1 He/N_2 gas at DF = 2 kV/mm) is –(3–17) V/mm *(8, 18–20)* with the mean peak width (full width at half maximum – fwhm) of ~0.4–0.7 V/mm *(20)*, depending on *R*. This translates into a peak capacity of ~20. New planar FAIMS devices have recently reached *R*~ 40, increasing the peak capacity two- to fourfold *(6)*. While IMS has achieved a higher *R* of ~100–150 *(21–23)*, tryptic digests occupy only approximately one-third of the IMS separation space and the peak capacity is similar to that of high-resolution FAIMS, ~35–50. The highest current resolving power of FAIMS or IMS is certainly not the physical limit and may be increased, e.g., by extending the ion residence in planar FAIMS *(6)* or by using higher drift voltages and longer drift tubes or higher gas pressures in IMS. However, those approaches are associated with a severe sensitivity loss or major technical difficulties.

Hence, the peak capacities of either FAIMS or IMS are insufficient for most biological samples, and are far below ~10^3 achieved by capillary liquid chromatography (LC) *(24)*. Of course, ions move much faster in gases than in liquids, and IMS methods are far superior to condensed-phase alternatives in terms of peak capacity production rate *(25)*. In particular, IMS (and, to a lesser extent, FAIMS) analyses could be folded into an LC separation like MS could be folded into IMS: ~10^2–10^3 IMS cycles fit within the elution time of a single feature in LC (~10–30 s). This makes LC/IMS *(26)* and LC/FAIMS *(19)*, with the LC stage raising total peak capacity to meet the analytical requirements, viable strategies for separations of complex mixtures. However, an LC step significantly reduces the throughput and negates the advantages of IMS methods with regard to reproducibility, simplicity, and automation.

A limited resolving power of FAIMS or IMS is yet more of a challenge for separation and characterization of gas-phase protein conformations, where a preionization separation step is not an option. For example, for ubiquitin and cytochrome *c*, FAIMS found some conformers but not others distinguished by H/D exchange *(27, 28)*. Conventional IMS provides a priori structural elucidation of ions by relating their geometries to mobilities using first-principles calculations *(29, 30)*. That is not currently possible for FAIMS because no way to connect CF values to ion geometries exists. That limits the utility of FAIMS in biological MS and mutes its advantage over IMS in terms of higher orthogonality to the MS dimension.

1.2. Two-Dimensional IMS

1.2.1. The Case for 2D IMS

The progress of liquid-phase separations illustrates that combining stages operating on different principles is usually more effective than pushing to increase the resolution of any one stage *(31)*. The major limitations of IMS and FAIMS noted earlier could similarly be overcome in 2D FAIMS/IMS. First, the peak capacity can be greatly increased, formally reaching ~50 (for FAIMS) × 100 (for IMS) = 5,000. The effective peak capacity of a multidimensional separation is always below the formal value, because of nonzero correlation(s) between the dimensions. So the key issue is the extent of correlation between FAIMS and IMS separation properties for the analytes of interest. As the values of a function and its derivative are generally unrelated, IMS and FAIMS dimensions could be largely orthogonal. That proved to be the case for tryptic peptides *(8)*. On the other hand, CF values are somewhat correlated to ion mass, as evidenced by species with $m > \sim 400$ Da (except some proteins with $m > \sim 25$ kDa) being C-type and those with $m < 100$ Da being A-type in N_2 *(32)*. Since m is also correlated to mobility, FAIMS and IMS cannot be fully independent. However, the orthogonality is significant, making the combination attractive. For example, even with the orthogonality of 20%, the 2D peak capacity would reach $\sim 10^3$, far beyond the limits of FAIMS or IMS alone and on par with high-end LC that is slower by orders of magnitude. An IMS separation with $R > 100$ could be completed in ~0.1 s, and the ion residence time in high-resolution planar FAIMS is about the same. Hence, a FAIMS/IMS analysis with 50 CF steps could, in principle, take just ~10 s vs. ~10 h ($> 10^5$ s) for an LC separation with $R \sim 10^3$ *(24)*.

In the structural biology context, the FAIMS/IMS combination would also increase the number of resolvable conformers and additionally enable characterization of species preseparated by FAIMS. That essentially turns FAIMS into a new structural elucidation method for gas-phase ions that often distinguishes isomers with minor geometry differences.

1.2.2. Approach to an Effective FAIMS/IMS/ToF MS Platform

While FAIMS/MS and IMS/MS systems are available as described earlier, FAIMS, IMS, and MS have not been joined in any commercial product. Theoretically, one could create a FAIMS/IMS/MS capability by attaching a FAIMS unit (Thermo Fischer Scientific) *(16)* to the inlet of IMS/ToF system (Waters) *(17)*. However, the IMS stage in that instrument uses a traveling wave rather than a constant electric field, which provides a low $R < 10$ and allows no *a priori* measurement of absolute mobilities needed for structural elucidation. Those drawbacks exclude or greatly limit

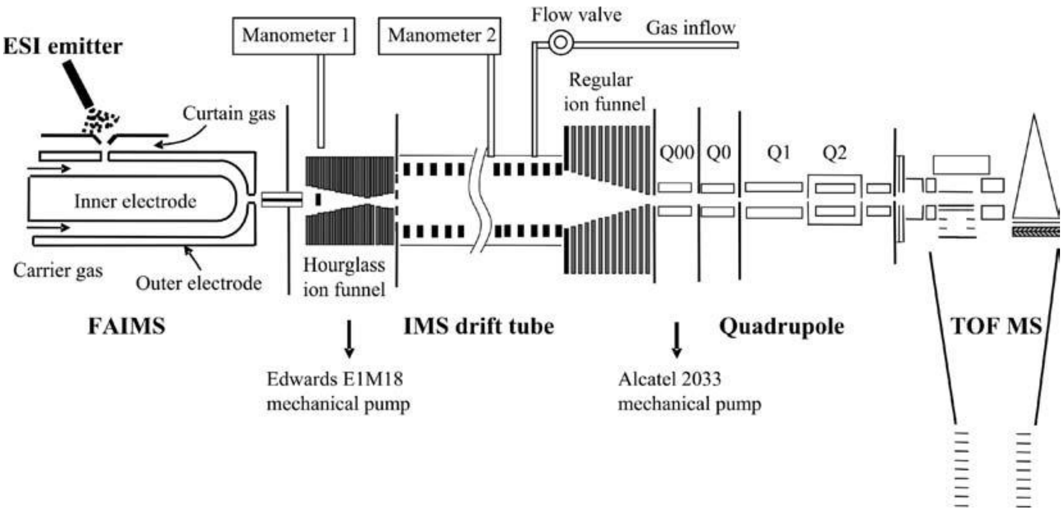

ESI emitter
Manometer 1
Manometer 2
Flow valve
Gas inflow
Regular ion funnel
Curtain gas
Inner electrode
Carrier gas
Outer electrode
Hourglass ion funnel
Q00
Q0
Q1
Q2
FAIMS
IMS drift tube
Quadrupole
TOF MS
Edwards E1M18 mechanical pump
Alcatel 2033 mechanical pump

Fig. 1. Diagram of the ESI/FAIMS/IMS/ToF system for 2D ion mobility analyses.

most applications of either analytical or structural nature, and much more powerful systems with $R > 100$ could be constructed using standard IMS separations by a uniform dc field. Here, we describe the FAIMS/IMS/MS technology developed at Pacific Northwest National Laboratory *(20)*, which involves coupling a commercial FAIMS to custom drift tube IMS and further to a commercial ToF employing electrodynamic ion funnel interfaces *(33)*, Fig. 1. The discussion focuses on details critical for the integration of such a system, examining the rationales for specific technical choices (**Subheading 3.1**). Then we illustrate the use of FAIMS/IMS/MS for analyses on peptide and protein level (**Subheadings 3.2** and **3.3**).

2. Materials

2.1. Major Instrumentation Parts

1. Commercial ToF (*see* **Subheading 3.1.1** and **Note 1** for selection considerations).
2. Time-to-digital converter, TDC (Ortec 9353, Oak Ridge, TN), I/O control board (National Instruments PCI-6711), control PC (Dell with the Windows XP operating system).
3. Materials and rf power supply for a quadrupole (*see* **Subheading 3.1.1** and **Note 2**).
4. Two mechanical pumps with capacities of >5 and >10 L/s (*see* **Note 3**), insulated pump hose.
5. Regular and "hourglass" electrodynamic ion funnels (**Subheadings 3.1.3** and **3.1.5**), rf power supplies: power amplifier (ENI2100L, Electronic Navigation Industries, Rochester, NY) and a custom high-Q head with a built-in waveform generator and rf power amplifier, dc power supplies (TD-9500, Spectrum Solutions, Russellton, PA).
6. Materials for the drift tube (**Subheading 3.1.4**): nonmagnetic steel tubes, blocks, and flanges, one high-voltage feed-through, three gas feed-throughs, ceramic rods, Teflon spacers, sheet copper, PEEK sheets (1-in. thick), electrical resistors and connecting wire, and insulating bolts and nuts.
7. Hygroscopic filters (5188-5248, Agilent Technologies, Santa Clara, CA).
8. High-transmission wire grid (20 lines/in., Buckbee-Mears, St. Paul, MN).
9. Selectra or other commercial FAIMS systems suitable for coupling with MS (*see* **Note 4**).
10. High-voltage dc power supplies: 1) EW60R10, 0–60 kV (Glassman, High Bridge, NJ) or others providing a lower

voltage, depending on the IMS drift voltage; 2) C60, 0–6 kV (EMCO High Voltage Corporation, Sutter Creek, CA).

11. High-voltage isolation transformer: AD6386 (Del Electronics, Valhalla, NY)
12. Two capacitance manometers (baratrons) for 0–20 Torr pressure, preferably with 0.1% accuracy (MKS, Wilmington, MA).
13. Gas lines with insulating breaks, needle flow valve (Swagelok, Solon, OH).
14. Optical communication system: ME540A-ST (Black Box Network Services, Lawrence, PA)
15. Fused-silica capillary (50 μm i.d./150 μm o.d., Polymicro Technologies, Phoenix, AZ).
16. Gas-tight microsyringe (Hamilton, Reno, NV), syringe pump (KD Scientific, Holliston, MA).
17. C_{18} SPE column (Supelco, Bellefonte, PA).
18. Adjustable-height laboratory table (Vestil Manufacturing, Angola, IN).

2.2. Reagents

1. Proteins and peptides in **Table 1** (Sigma, St. Loius, MO), trypsin (Promega, Madison, WI).
2. Proteins: bovine ubiquitin, equine cytochrome *c* (Sigma, St. Loius, MO).
3. Solvents: HPLC grade methanol, acetic acid (Fisher, Fairlawn, NJ), deionized water.

3. Methods

3.1. Constructing an ESI/FAIMS/ToF MS System

3.1.1. Selection and Modification of a ToF Platform

In an IMS/ToF system, ToF must detect ions while preserving their spatial spread established by the IMS stage. Given that the typical IMS drift time (t_D) for peptide and protein ions generated by ESI is ~30–100 ms (**Subheading 3.1.2**), the ion flight time through the whole ToF instrument (t_F) should preferably be <~1 ms. Otherwise, the diffusion, Coulomb repulsion, and mass dispersion in the ToF champer will materially remix ions separated by IMS, reducing and eventually obliterating the IMS separation. A long t_F would also increase the measured t_D and thus decrease the apparent K, though that could be corrected by measuring t_D as a function of IMS drift voltage, U, or pressure, P *(34)*. The pusher frequency of typical commercial ToFs is ~10 kHz, thus ions spend <~100 μs in the ToF chamber and the main issue is the transit time through preceding rf ion guides (quadrupoles or octopoles). That is not a factor in other MS and LC/MS analyses and some otherwise excellent instruments feature an extended

Table 1
Components for the preparation of tryptic digest used to evaluate the application of FAIMS/IMS for bottom–up proteomic separations

Proteins	Sigma catalog no.	Molecular mass (Da)	Concentration (mg/mL)
Bovine heart cytochrome *c*	C2037	12,327	0.198
Bovine milk α-lactalbumin	L6010	14,178	0.195
Equine skeletal muscle myoglobin	M0630	16,951	0.200
Bovine milk β-lactoglobulin	L3908	19,883	0.196
Bovine erythrocyte carbonic anhydrase	C3934	28,982	0.200
Rabbit glyceraldehyde-3-phosphate dehydrogenase	G2267	35,688	0.198
Chicken egg white ovalbumin	A2512	44,287	0.198
Bovine serum albumin	A7638	~66,430	0.195
Bovine holo-transferrin	T1408	~77,750	0.194
Rabbit muscle phosphorylase b	P6635	~97,200	0.198
E. coli β-galactosidase	G5635	~116,300	0.204
Peptides			
Bradykinin fragment 2–9	B1901	904.0	0.003
[des-Pro^{3}, Ala^{2,6}]-bradykinin	B4791	921.1	0.018
Bradykinin	B3259	1,060.2	0.048
Tyr-bradykinin	B4764	1,223.4	0.012
Human angiotensin I	A9650	1,296.5	0.024
Human osteocalcin fragment 7–19	O3632	1,407.6	0.006
Human vasoactive intestinal peptide fragment 1–12	V0131	1,425.5	0.014
Syntide 2	S2525	1,507.8	0.028
Human leptin fragment 93–105	L7288	1,527.7	0.015
Pro$_{14}$–Arg	P2613	1,532.9	0.005
Human fibrinopeptide A	F3254	1,536.6	0.006
Porcine dynorphin A fragment 1–13	D7017	1,604.0	0.062
Neurotensin	N6383	1,672.9	0.017
[Ala^{92}]-peptide 6	P7967	2,123.4	0.008

(continued)

Table 1
(continued)

Proteins	Sigma catalog no.	Molecular mass (Da)	Concentration (mg/mL)
Human diazepam binding inhibitor fragment 51–70	G9898	2,150.4	0.009
Presenilin-1 N-terminal peptide	P2490	2,227	0.214
Epidermal growth factor receptor fragment 661–681	E9520	2,318.6	0.006
3X FLAG peptide	F4799	2,861.9	0.003
Human Tyr-C-Peptide	C9781	3,780.2	0.018

Concentrations refer to the values for peptides and protein digests in the final solution

accumulation of ions prior to injection into the ToF chamber, e.g., for collisional cooling purposes (*see* **Note 1**).

The transit time of ions through the quadrupole chambers of Q-ToF roughly scales with the gas pressure, hence the front chamber with the highest gas load and pressure (Q0) is of primary concern. That pressure depends on the gas flow from IMS that is proportional to P and the area of smallest aperture (*conductance limit*, Θ) between the IMS and ToF stages. While one could reduce Θ or P until the delay of ions in Q-ToF is acceptable, the former would drastically reduce the ion transmission and the latter would degrade the resolution for a drift tube of practical length. This problem may be solved by adding a differentially pumped stage (Q00) before Q0 (*see* **Notes 2 and 3**). The pressure in Q00 is ~0.1 P, with the gas inflow into and the pressure in Q0 decreasing in proportion. This allows effective IMS operation at $P \leq 5$ Torr (for N_2 gas). For higher P, two or more identical differentially pumped stages may be inserted in front of Q0 (*see* **Note 5**). However, the acceptance diameter of reasonable quadrupoles still limits Θ to ~2 mm.

3.1.2. Axial Ion Focusing in IMS: Fundamental Considerations

Ions in a uniform electric field undergo free thermal diffusion, producing parabolically divergent beams inside IMS. As the diffusion in a weak field is isotropic, initially small ion packets expanding inside IMS remain spherical, with equal dimensions along and across the drift direction. Hence, the beam diameter (W) at the end of drift tube of length L is ~L/R, which favors minimizing L, but achieving a specific R at a particular P requires a certain minimum L. As $U = EL$ and $K = N_0 \times K_0/N$ and $t_D = L/v$, **Eq. 1** yields:

$$L = (N_0 K_0 t_D U / N)^{1/2}. \tag{3}$$

For ESI-generated peptide and protein ions, K_0 changes little over a wide size range because larger ions tend to carry higher z, and typical values in N_2 at room temperature are ~0.75–1.25 cm^2/(V × s) (*see* **Note 6**). The ranges of t_D and U may be determined considering that the maximum (diffusion-limited) IMS resolving power depends on U and gas temperature (T) as

$$R_d = [Uze / (16kT \ln 2)]^{1/2}, \tag{4}$$

where e is the elementary charge and k is the Boltzmann constant *(35)*. However, in practice R is limited by finite temporal widths of the ion pulse injected into IMS (t_P) and the detector acquisition window (t_A), defined as fwhm *(35)*:

$$R = t_D / [t_P^2 + t_A^2 + (t_D / R_d)^2]^{1/2}. \tag{5}$$

Here, t_A is the ToF cycle duration, which typically equals ~0.1 ms (**Subheading 3.1.1**). The minimum t_P permitted by limitations of ion injection process is also ~0.1 ms. Though that value could be decreased somewhat at a major cost in sensitivity, the resolution by **Eq. 5** is controlled by the higher of t_P and t_A and reducing t_P much below t_A would not be useful. Hence $t_P = t_A$ ~0.1 ms is close to the optimum (*see* **Note 7**). Then, to avoid a significant (>10%) loss of IMS resolution, the quantity t_D/R_D in **Eq. 5** must exceed ≈0.3 ms, which means $t_D \geq 35$ ms to achieve $R > 100$. By **Eqs. 4** and **5**, establishing $R > 100$ (for all z) also requires $U \geq 4$ kV. Substituting those inequalities for t_D and U into **Eq. 3** at $P = 5$ Torr at $T = 25$°C yields $L \geq 170$ cm, assuming $K_0 \leq 1.25$ cm^2/(V × s) to cover most biological ions.

At $R = 100$, $L \geq 170$ cm means $W > 17$ mm. As Θ must be ≤ 2 mm (**Subheading 3.1.1**), the ion transmission efficiency at IMS exit (S) would be <$(2/17)^2$ ~1%. In actual systems subject to other design constraints, S was less than that optimized value, e.g., ~0.1% in the high-resolution IMS with $L = 63$ cm *(36)*. By **Eq. 3**, L and thus W are proportional to $1/\sqrt{P}$, hence S scales as $1/P$. However, the maximum P is limited by the pumping capacity at IMS/ToF interface (**Subheading 3.1.1**). A sensitivity loss by ~10^2–10^3 times prohibiting real proteomic analyses, effective axial ion focusing in IMS is a prerequisite for practical IMS/MS systems.

3.1.3. Lossless IMS/MS Interfacing Provided by an Ion Funnel

The need for ion focusing during or after IMS separation in IMS/MS has been widely recognized. Implementing IMS in an axially segmented quadrupole *(37)* or ion tunnel *(17)* provides high sensitivity because of rf focusing over the whole "drift tube", but poor resolution because of low operating pressure and limited

IMS length. Efforts to focus ions by shaped dc electric fields at the end of usual drift tubes were only partly successful *(38)*. FAIMS may focus ions at ambient conditions *(39)*, but does not accept wide ion beams. However, beams of any diameter could be focused effectively at reduced gas pressure using ion funnels *(33)*.

Electrodynamic ion funnels combine dc and rf fields to compress ion beams with ~100% efficiency: the dc field pulls ions through the funnel while the effective potential due to rf repels ions from the walls. The initial beam diameter is constrained only by the funnel orifice. The minimum final diameter is limited by space charge, but even for extreme ion currents ~1 nA could be <2 mm, i.e., within the acceptance diameter of standard quadrupoles. The strength of ion confinement in a funnel is proportional to P^{-2}, which sets the maximum P*(40)*. Presently used funnels are effective up to ~8 Torr (in N_2), which exceeds the 5-Torr limit imposed by the pumping capacity of a single Q00 stage (**Subheading 3.1.1**). However, the funnel operating pressure could be elevated to ~30 Torr, and there is a strong motivation to so do (*see* **Note 8**).

A regular ion funnel comprises a stack of insulated electrodes with coaxial orifices of decreasing diameter, **Fig. 2a, b**. A uniform

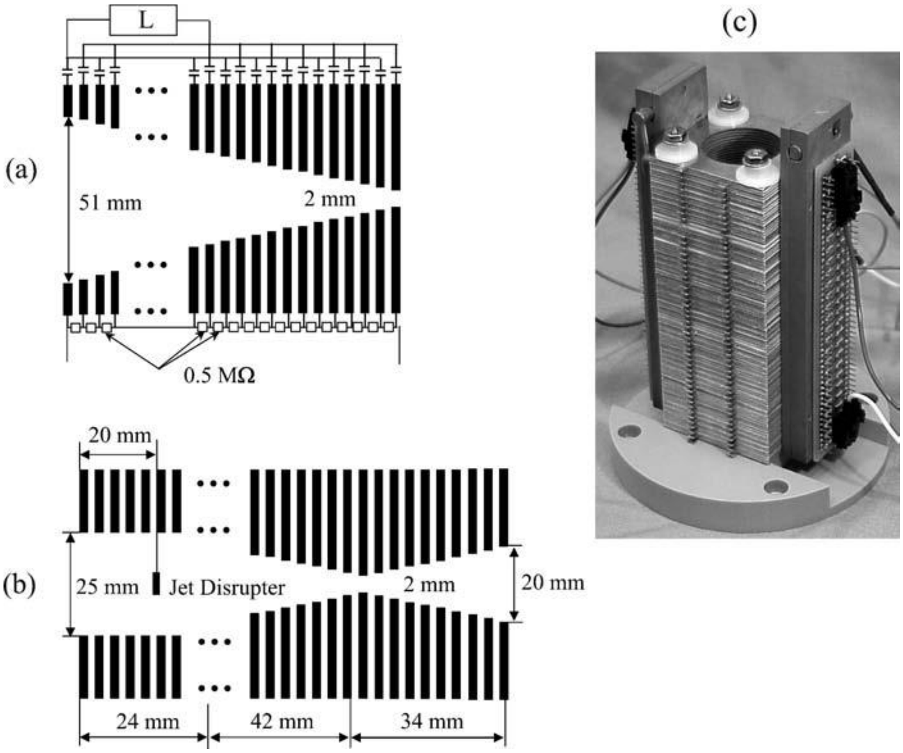


Fig. 2. Ion funnels: schemes of a regular (**a**) and hourglass (**b**) funnels. A photo of a funnel with 25-mm entrance orifice is shown in (**c**).

axial dc field is produced by a ladder of voltages established on electrodes by a voltage divider. The rf voltage is applied antiphase to adjacent electrodes. The number, dimensions, and spacing of electrodes may differ depending on the gas pressure, ion beam width, rf power supply limitations, and other factors. A sufficient axial field being critical to move ions through the funnel, the electrode spacing must be at most ~1/2 of the smallest funnel aperture. Given the exit aperture of 2.2 mm, we used 0.5-mm thick brass electrodes spaced by 0.5-mm thick Teflon sheets. The entrance orifice must be large enough to capture the ion beam coming from IMS, including the tails of its radial intensity distribution. As that distribution is nearly Gaussian, essentially all ions are found within $2W$ off the axis. The 51-mm orifice diameter equals $2.4W$ for $L = 210$ cm (**Subheading 3.1.4**). The conical angle of the funnel should be small enough for gradual beam compression, which defines the minimum funnel length for a given entrance orifice and thus the number of electrodes needed for a given spacing. We used 84 electrodes, for the conical angle of 32°. The last electrode carries no rf and an adjustable dc voltage to assist ion extraction from the funnel.

A stack of electrodes spaced by insulators is a capacitor. A lower capacitance decreases the rf current through the funnel, reducing the power requirements and Ohmic heating of electrodes. The funnel capacitance is proportional to the electrode surface and the dielectric constant (ε) of material between them. Hence, one should choose the smallest possible electrodes and the insulator with lowest ε. The minimum ε of solid materials belongs to Teflon ($\varepsilon = 1.5$), and we use it for the insulating sheets between electrodes (*see* **Note 9**). The capacitance of terminal IMS funnel in present system is 8 nF.

As any capacitor, the funnel may be a part of resonating LC circuit. Resonance conditions minimize the rf power requirements and are preferred, but the resonance frequency may not be optimum for ion focusing and the funnel may have to be run off-resonance. This is the case with IMS exit funnel, driven by the power amplifier at 500 kHz with peak amplitude of 20–40 V. Different electrode pairs have unequal surface areas and thus different capacitances. To effectively eliminate that difference and to isolate the dc load from the rf circuit, each electrode is connected to that circuit through a sufficiently large capacitor, 10 nF (**Fig. 2a**).

Ion packets focused in a funnel are axially broadened by diffusion and Coulomb repulsion, which may mix species separated by IMS and thus worsen the resolution. Hence, the transit time of ions through the funnel should be short, which would also reduce space-charge effects that impede ion focusing. That time is inversely proportional to the axial field and may be reduced by raising the dc voltage across the funnel. However, a too high field may heat ions to the point of isomerization or dissociation, and excessive axial velocity does not allow enough time for ion focusing. A dc field matching that in IMS appears close to optimum and also permits treating the funnel as a part of the drift tube, which simplifies the interpretation of IMS data *(23)*.

3.1.4. Modular IMS Drift Tube

As a funnel may focus ion beams of any diameter, it allows IMS drift tubes of virtually unlimited length. Based on calculations for minimum L (**Subheading 3.1.2**), we have chosen $L = 210$ cm. Unlike for previous IMS instruments with $L < 60$ cm, a single-piece drift tube of such length would be impractical. A modular IMS comprising identical units (here 10) is flexible, easily serviceable, and may be extended at will *(23)*. Each unit consists of a 20-cm long cylindrical chamber (i.d. = 20 cm) made of steel (*see* **Note 10**), housing a stack of 21 copper ring electrodes, **Fig. 3**. Units are electrically insulated by 1-cm-thick disks machined from PEEK (*see* **Note 11**), vacuum-sealed by O-rings, and joined using insulating bolts and nuts (reinforced fiberglass). The disks feature center holes for ion passage (i.d. 6.5 cm) and 0.5-cm-deep

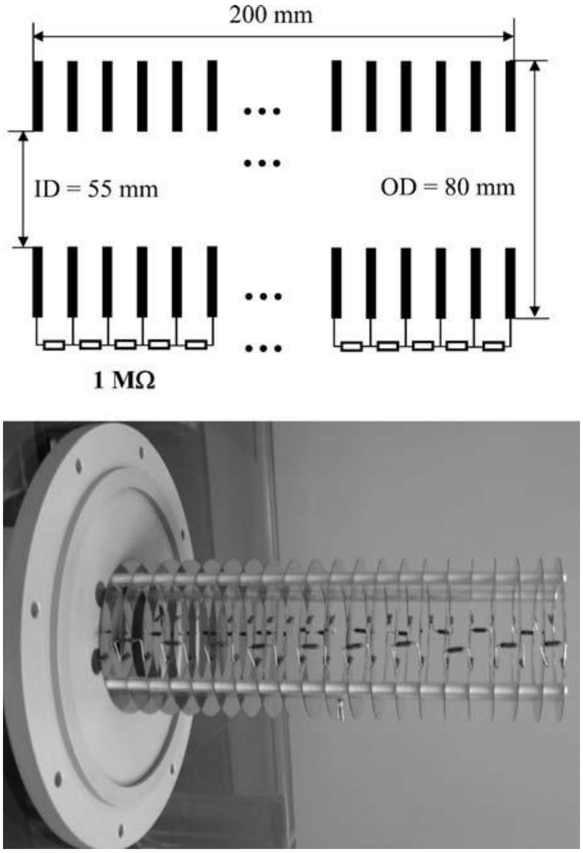


Fig. 3. Scheme and photo of a single IMS unit.

blind holes on both sides for mounting four 21-cm long solid ceramic rods (o.d. 3.2 mm) supporting electrodes inside each unit. The electrodes are spaced by 10-mm long plastic tubes that freely slide on the rods (*see* **Note 12**).

All electrodes in a tube are consecutively joined by high-precision 1-MΩ resistors, dividing the drift voltage along the electrode stack (*see* **Note 13**). The median resistor of each unit is connected to the chamber (*see* **Note 14**). The drift voltage is set by controlling the potentials on front and back electrodes of assembled tube. The first is generated by the 60-kV dc power supply, loaded using an electrical feed-through on the chamber wall, and monitored by a high-voltage probe. The second is defined by another dc power supply referenced to ground, and selected to match the axial fields in the drift tube and terminal ion funnel (**Subheading 3.1.3**). In typical experiments, the dc potential ranges along IMS and ion funnel were 4,000–240 V and 240–90 V, respectively, providing equal axial field of 17.9 V/cm.

The buffer gas continuously flows into IMS through a line in the last unit, filling a toroidal space between a steel insert and chamber walls and then leaking into the drift tube through a slit along the insert circumference that ensures a radially uniform inflow. The gas pressure in IMS is monitored by a baratron connected to one of the chambers. As all chambers are at a voltage, gas lines leading inside feature insulating breaks. The gas flow (and thus the IMS pressure) is controlled by a needle valve on the distribution line. The purity of buffer gas is paramount to the reproducibility and accuracy of IMS data. We use UHP gases dried by hygroscopic filters.

3.1.5. Effective Ion Injection into IMS Using Hourglass Funnels

Regular ion funnel (**Subheading 3.1.3**) were designed for steady-state ion focusing into MS systems and thus provide no ion storage space. As IMS is a pulsed technique, ions arriving from a continuous source (e.g., ESI or ESI/FAIMS) between pulses must be accumulated or discarded. For t_P ~ 0.1 ms and $t_D \geq 35$ ms (**Subheading 3.1.2**), rejecting ions that approach a closed IMS gate means a duty cycle of <0.3%, which is inadequate for proteome analyses. Here, ions appearing from FAIMS are focused and then accumulated using an "hourglass funnel" *(23)* that is similar to the regular one but has a neck, creating a storage volume of 3.5 cm^3 contiguous with the drift tube (**Fig. 2b**). In this volume, ions are confined radially by the rf field and axially by dc barrier of 12 V (biased to the last funnel electrode) on the wire grid located 2 mm after that electrode. The entrance orifice (25 mm) is one-half that in the IMS exit funnel, which is sufficient to capture the ion plume expanding from FAIMS exit.

All funnel voltages are provided by dc and ac power supplies floated on the front IMS voltage (**Subheading 3.1.4**), with ac fed from an isolation transformer raising the line ac power to high voltage. The dc voltage range across the funnel is 40–190 V (*see* **Note 15**). The rf waveform is operated at 500 kHz and 10–40 V: higher amplitudes strengthen ion focusing and thus maximize ion transmission, but field heating may cause the isomerization of fragile ions that affects IMS separation properties (**Subheading 3.3.2**). This is not a consideration after the IMS stage, so the IMS exit funnel may use a stronger rf. The collisional fragmentation is still an issue after IMS analyses, but usually requires a stronger rf field than needed for effective ion focusing. Ions are pulsed into IMS by periodically dropping the grid voltage by 27 V, using a two-channel voltage switch. As the switch is electrically floated, the pulse frequency and duration are controlled via a trigger signal sent through an optical cable.

While the gas in IMS exit funnel is still, the hourglass funnel faces a supersonic jet caused by abrupt pressure drop from ~1 atm in FAIMS. The resulting gas dynamics that interferes with ion focusing is mitigated using a jet disrupter *(42)* – a brass disk (o.d. 6.5 mm) held on the funnel axis by insulated metal cross wires (**Fig. 2b**). Its voltage is typically biased at ~25 V above the adjacent electrodes, but could be varied to control the ion transmission or remove parasitic ions with low *m/z* that take up the charge capacity of ion funnel trap *(43, 44)*.

The hourglass funnel is housed in a machined steel chamber biased at 200 V above the front IMS voltage. The chamber is evacuated through an insulating reinforced plastic hose, with the pressure adjusted by a flow valve on the hose and monitored by a baratron. This pressure should be below the pressure in IMS, so that the gas flows through the funnel out of IMS. That is critical to bar neutrals produced by ESI and carried by gas through FAIMS from IMS where they could charge-exchange with ions, resulting in wrong and irreproducible IMS data *(14)*. However, too fast flow would overcome the dc field pulling ions into IMS, especially for ions with lower *K*. So the pressure differential between IMS and source chamber should be tuned carefully, watching for discrimination based on mobility (*see* **Note 16**). The gas flow out of IMS also allows independent gas compositions in IMS and FAIMS, which is often important.

ESI sources are typically coupled to ion funnels using a heated capillary where droplets generated by ESI are thermally desolvated. Ions filtered by FAIMS are already desolvated and heating is redundant, but one still needs a conductance limit between the atmospheric-pressure FAIMS and ion funnel chamber. This is provided by a 0.43-mm aperture drilled in a 0.2-mm-thick steel plate, which avoids ion losses to the walls that occur in narrow capillaries(6). The gas conductance of this inlet is ~1 L/min.

3.1.6. Coupling the FAIMS Stage to IMS Inlet

The FAIMS Selectra system is coupled to IMS using a PEEK adapter that secures the FAIMS exit ~0.5 mm away from the inlet aperture. The gas flow through FAIMS gap is approximately proportional to the total carrier gas consumption, and, for the usual FAIMS inflow of ~1.5–4 L/min, varies in the ~0.4–1 L/min range. Though this value roughly matches the inlet conductance (**Subheading 3.1.5**), the flows in FAIMS and IMS inlet should be decoupled to allow changing the FAIMS gas flow without affecting the ion funnel pressure and to prevent the inlet suction from disturbing smooth gas flow inside FAIMS. Hence, the adapter is not gas-tight and the gap between FAIMS and IMS inlet equilibrates with atmosphere.

The FAIMS power supply unit has to be in a constant communication with the control PC. As the PC is grounded, the electrical wiring in Selectra was replaced by optical cables.

3.1.7. ESI Emitter

The ESI emitter (*see* **Note 17**) is mounted on an *X*–*Y* translation stage allowing fine adjustment, and positioned a few mm away from the FAIMS inlet aperture. Samples are infused to the emitter through a steel union using a microsyringe driven by a pump. The union is biased at 2–2.5 kV above the inlet using the 6-kV dc power supply floated on the IMS front voltage.

3.1.8. Overall Arrangement and Safety Enclosure

The IMS drift tube resides on plastic supports secured to a commercial laboratory table with height adjustment to facilitate coupling to ToF. The ESI emitter and syringe pump sit on a separate plastic table. The FAIMS power supply is enclosed in a plastic container. For electrical safety reasons, the whole ESI/FAIMS/IMS system is surrounded by a grounded wire cage fastened to steel beams, with hinged access doors fitted with electrical safety interlocks.

3.1.9. Instrument Control and Data Acquisition System

The ion detector in Q-Star has four channels, with counts from each streamed into a dedicated TDC. To improve performance and support the needed new software (below), that TDC was upgraded to a commercial single-channel 10-GHz TDC, with counts from all channels combined after the preamplifier. The timing sequences of IMS pulsing and MS data acquisition are set by I/O control board. Both TDC and that board are installed in a PC that also runs the instrument control software.

The Q-Star operating software (Sciex Analyst) averages ToF spectra over a preset time, which would obliterate the IMS separation, and cannot control IMS operation or coordinate it with MS analyses. Present system is run by custom software that controls the instrument and stores and processes FAIMS/IMS/MS data. However, ToF is operated as usual, with Analyst setting t_A depending on the upper m/z limit desired, for instance 0.10 ms for ~1 kDa and 0.14 ms for ~2 kDa. To synchronize IMS and MS operations, the IMS pulsing period is divisible by the pusher period (e.g., 500 ToF spectra with $t_A = 0.14$ ms are acquired during $t_D = 70$ ms).

The software was designed in Visual Basic 6 and implemented on top of the ActiveX controls running the TDC. The system collects a set number of sequential 2D maps, each derived by adding a certain number of elementary IMS-MS frames. The acquisition of first frame starts from a ToF pusher pulse that triggers a call to open the IMS ion gate (**Subheading 3.1.5**) and start the data collection. The software then maps the ion counts recorded by the TDC onto a 2D frame, filling a sequence of "0" and "1" in each t_D segment equal to t_A. This continues for a set number of periods up to the desired maximum t_D. In the end, each frame pixel contains "1" if an ion has arrived within corresponding bins of t_D and t_A, and "0" otherwise. The program then resets the experiment and starts acquiring the next frame coincidentally with a ToF pusher pulse. All following frames up to a predefined number (typically 100–1,000) are histogrammed on top of the first, transforming a binary series in each IMS drift time segment into a ToF spectrum. This spectrum is converted into a mass spectrum using calibration and standard formulae relating m/z of ions to their flight times.

FAIMS/IMS/MS analyses involve collecting IMS/MS data at a sequence of CVs set from FAIMS control PC (**Fig. 4**). This process creates a 3D data set that may be visualized as a series of IMS/MS maps (above) or FAIMS/IMS palettes showing the isomers for a particular m/z.

3.2. 2D Ion Mobility Separations in Bottom–Up Proteomics

3.2.1. Proteolytic Digestion

The capabilities of FAIMS/IMS/ToF system are illustrated using a proteomic standard that comprises pooled digests of 11 proteins (distributed over the mass range representative of typical proteomes, 12–116 kDa) and 19 peptides with masses of 0.9–3.8 kDa, **Table 1** *(45)*. The protein mixture was denatured in urea and thiourea, reduced by dithiothreitol for 30 min at 60°C, diluted 10× in 100 mM NH_4HCO_3, and digested with trypsin in a 1:50 enzyme/protein ratio (w:w) for 3 h at 37°C. Ideally, this would produce 401 tryptic peptides. The digest was cleaned up using the SPE column, combined with a solution of peptides, and dissolved in 50:49:1 methanol/deionized water/acetic acid to the concentrations listed in **Table 1**.

3.2.2. Analysis of a Peptide Mixture Using FAIMS/IMS/MS

Setting a FAIMS/IMS/MS experiment involves defining several FAIMS parameters: the carrier gas, DV, the CV range to be covered, instrumental resolution, and CV increment. We use 1:1 He/N_2 gas, which works for tryptic digests better than N_2 *(18)*. FAIMS separations generally (though not always) improve with increasing DF, and the allowed maximum is usually a good choice. Peptides are C-type ions that are analyzed using negative DV and CV, known as the P2 mode *(18, 19)*. For Selectra with a 2-mm gap, the maximum DF of 2 kV/mm and CF of –(3–17) V/mm (**Subheading 1.1**) translate into DV = –4 kV and CV = –(6–34)

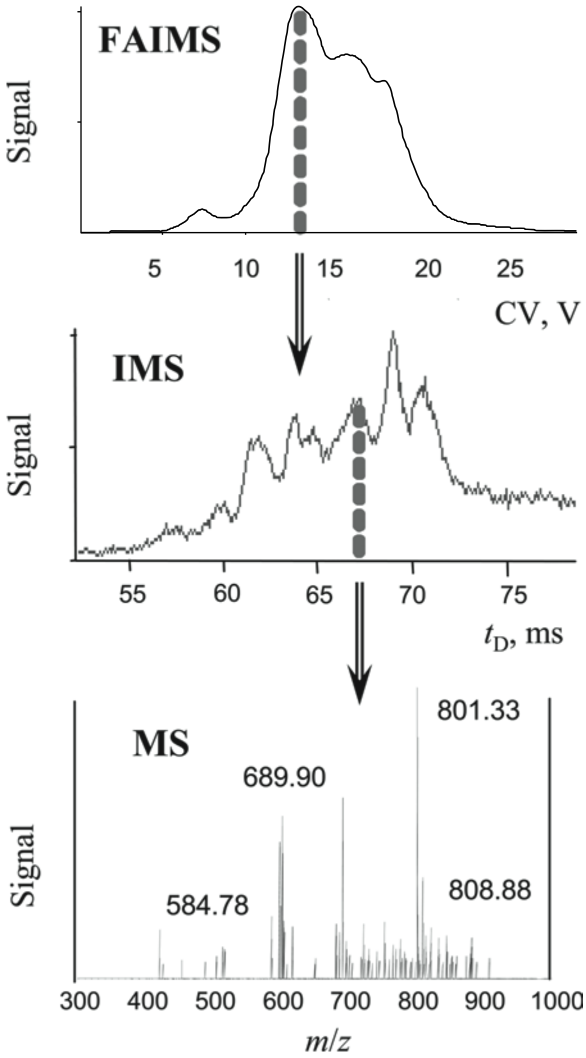


Fig. 4. Experimental sequence of FAIMS/IMS/ToF analyses: a segment of FAIMS CV spectrum is selected for nested IMS/MS analyses, in which slices of IMS drift time distribution are mass-dispersed.

V (*see* **Notes 18** and **19**). The ion signal at highest absolute CVs was too low for quality data, and measurements were limited to CV = −(7–22) V. The resolving power is set by adjusting the axial gap width in hemispherical region in 0.1-mm increments from 1.7 to 2.7 mm (*see* **Note 4**). Here, we used the 2.3-mm setting that provides a fair resolution/sensitivity balance *(46)*. The CV increment should be selected to enable detection of the species with narrowest CV peaks while minimizing wasteful overlaps. Since the fwhm of peptide peaks in FAIMS spectra obtained using Selectra is typically 1 V or greater *(19)*, we adopted a fixed CV step of ~1 V.

The data for a tryptic digest (**Subheading 3.2.1**) exhibit ~360 unique (out of >600 total) "spots" among IMS/MS maps at 16 different CVs (Fig. 5a), *(20)*. Similarly to previous findings, the FAIMS and MS separations are orthogonal (Fig. 5b). The FAIMS and IMS dimensions exhibit a modest correlation (Fig. 5c), because both *K* and absolute CV tend to increase for peptides with higher *z*. For ions with any particular *z*, FAIMS and IMS dimensions are virtually independent. The effective 2D peak capacity of this FAIMS/IMS separation is ~500, and could be doubled by replacing Selectra with high-resolution planar

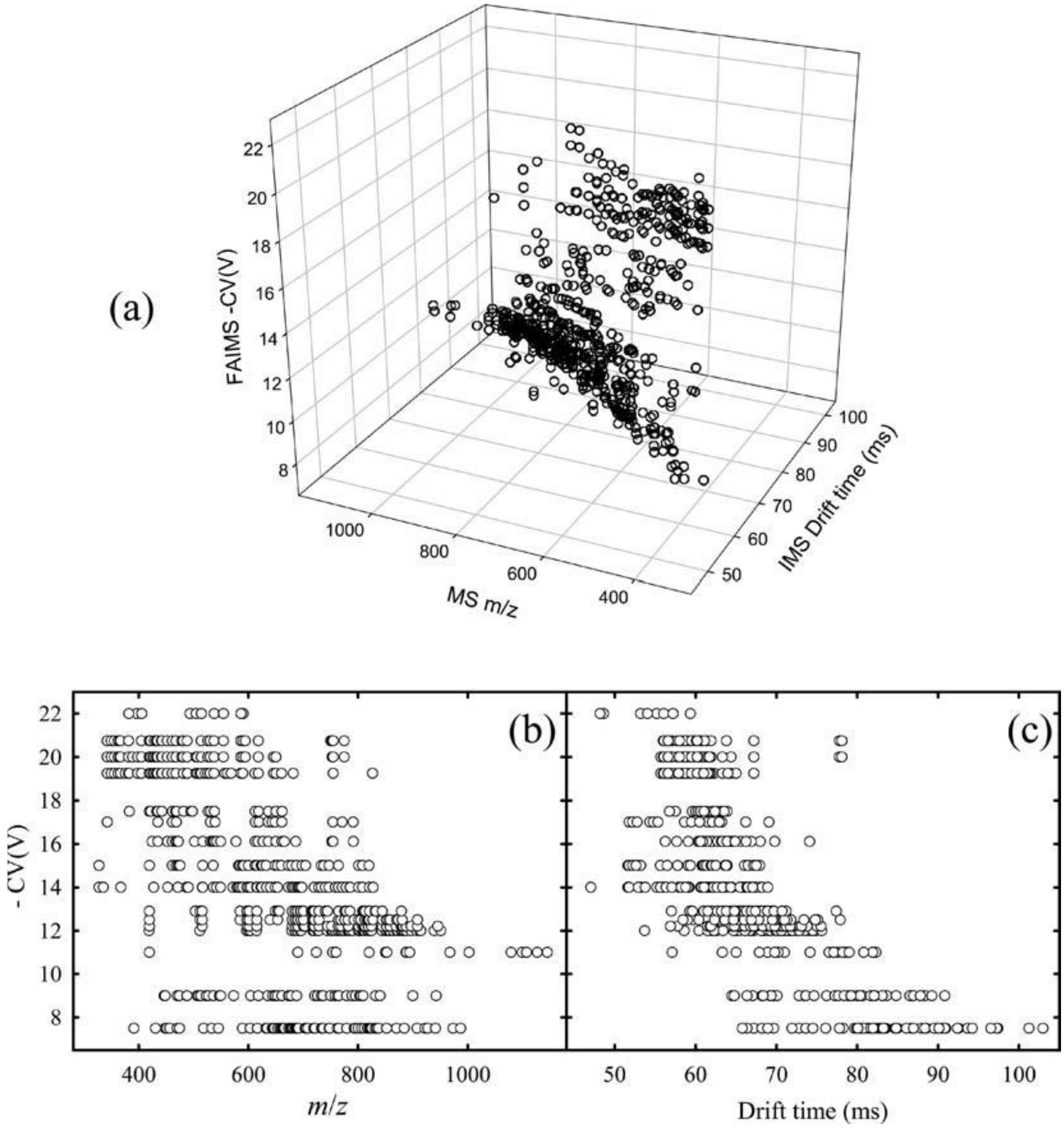


Fig. 5. Dispersion of tryptic peptides in the 3D FAIMS/IMS/MS space (**a**) and in the FAIMS/MS (**b**) and FAIMS/IMS (**c**) planes.

FAIMS *(6)*. So the projected value of ~10^3 (**Subheading 1.2.1**) is attainable with existing technology.

To maximize the throughput without affecting proteome coverage, the CV step should be varied in proportion to mean or minimum peak fwhm. In cylindrical FAIMS, peaks tend to widen at higher absolute CV because of stronger ion focusing *(5)*. Here, fwhm and their standard deviations at CV of −7.5, −13, and −20.5 V (typical values for peptides with $z = 1$, 2, and 3) are, respectively, 0.65 ± 0.15, 1.1 ± 0.15, and 2.5 ± 0.7 V (20): fwhm roughly scales with CV. Hence, the overlap between data separated by a fixed CV increases at higher absolute CVs, e.g., the IMS/MS maps at CV of −19.2 and −20.7 V share 65% of spots vs. 10% in maps at −7.5 and −9 V. Therefore, employing a constant relative CV step appears a reasonable strategy.

3.3. Characterization of Protein Conformers Using FAIMS/IMS

3.3.1 2D Mapping of Macromolecular Isomers

Any complex mixture separated by FAIMS/IMS (**Subheading 3.2.2**) could also be addressed, albeit much slower, by LC or gel methods. What those methods are incapable of is structural elucidation, because chromatographic retention times are not tightly connected to molecular geometry. In contrast, a relative simplicity of ion motion in gases permits quantitatively relating mobilities to 3D ion geometry, allowing structural characterization using IMS and FAIMS/IMS (**Subheadings 1.1** and **1.2.1**). The major issue with those approaches to conformational analyses of biomolecules is the degree of similarity between solution- and gas-phase structures. Though these are certainly not identical, soft ESI/MS processes apparently retain the key elements of biologically relevant folding under mild solution conditions *(47)*.

Here, we show the conformer separation by FAIMS/IMS for ubiquitin – a common model protein (8,565 Da) previously probed by many analytical techniques, including FAIMS and IMS individually *(27, 48)*. A 50-μM solution of ubiquitin in 50:49:1 methanol/water/acetic acid was infused at 0.5 μL/min. To minimize ion activation in the FAIMS/IMS ion funnel that may distort protein geometries between the two stages, the peak rf voltage was 10 V, i.e., the lowest value providing fair ion signal. The charge distribution of protonated ubiquitin ions comprised z = 6–14, and FAIMS/IMS palettes were measured for each z *(11)*. These data reveal more isomers than FAIMS or IMS alone for every z (**Fig. 6**), e.g., seven (A–G) for 6+ vs. three or four in IMS and two or three in FAIMS. The total number of species distinguished for all z was 40 using FAIMS/IMS vs. 14–18 using either FAIMS or IMS. Many newly detected conformers are traces engulfed by dominant peaks in 1D IMS. These data establish the dynamic range of ~10^3 for isomeric protein separations using FAIMS/IMS vs. <10^2 using IMS *(11)*. In other cases, peak shoulders or tails in IMS reveal the presence of another isomer

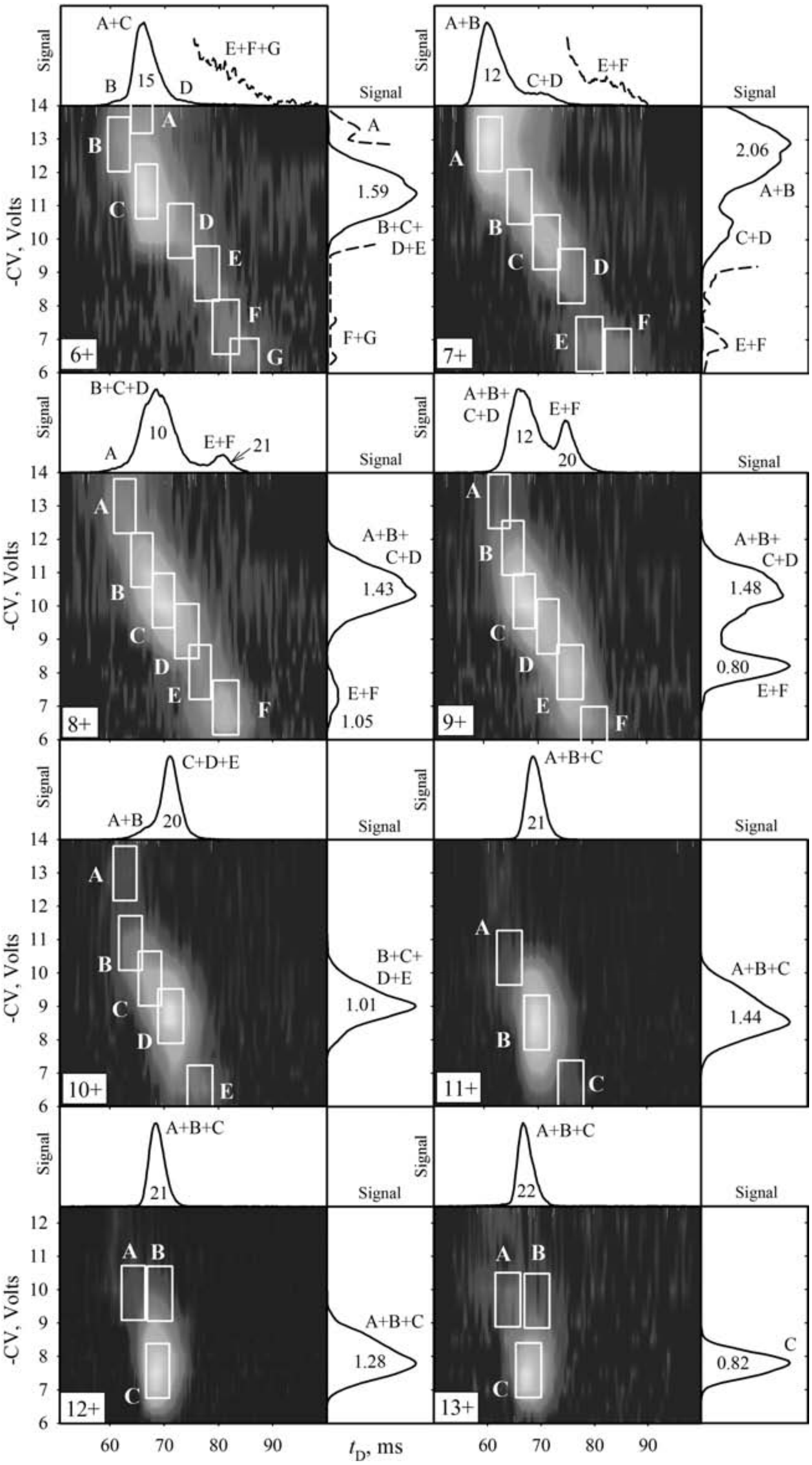


Fig. 6. FAIMS/IMS palettes for ubiquitin ions ($z = 6$–13), with peak rf amplitude of 10 V on the ion funnel at IMS entrance. The measured FAIMS-only spectrum is on the right of each palette. The composite IMS spectrum from the projection of palette on the t_D axis is on the top of each palette. All plots are scaled to equal dominant peak intensity. Boxes in palettes define distinct conformers marked in 1D and 2D spectra by letters. The apparent *R* for features in IMS spectra and fwhm (V) of features in FAIMS spectra are indicated.

but not its accurate t_D value. Such conformers are often fully separated by FAIMS/IMS, allowing t_D to be determined. Hence, the capability of FAIMS/IMS to "dig" deeper into isomeric mixtures is valuable for structural proteomics and potentially for analyses of biomedically relevant misfolded proteins that first emerge as traces on the background of normal conformers. However, the utility of FAIMS/IMS is not limited to characterization of minor isomers. For highly charged cytochrome *c* ions, 2D palettes *(11)* reveal multiple unfolded conformers with abundances of >20% each that have nearly equal mobilities and thus were not discovered by IMS alone *(49)*.

3.3.2. Probing the Dynamics of Conformational Transitions

A core proteomic technique is MS/MS, where ions with specific m/z selected in the first MS stage are dissociated, and products are dispersed by m/z in the second MS stage. Unlike for peptides (**Subheading 3.2.2**), for protein ions FAIMS and IMS dimensions seem substantially correlated (Fig. 6). While this limits the gain in peak capacity provided by 2D separations, it enables an ion mobility analog of MS/MS where specific conformers filtered by FAIMS are isomerized, and resulting structures are dispersed by K in IMS *(11)*. Ions may be isomerized between FAIMS and IMS stages by field heating in the funnel that approximately scales as (rf amplitude)^2 (*see* **Note 20**). For example, (at least partly) folded ubiquitin ions progressively unfold when the peak rf voltage is ramped from 10 to 40 V *(11)*, producing spots on 2D palettes to the right of their parents (Fig. 7). The unfolding follows the general route from compact to partly folded to unfolded families observed in IMS experiments using thermal or collisional activation *(48)*. However, by probing the evolution of specific precursors, FAIMS/IMS disentangles the reaction pathways mixed in IMS and thus paints a more quantitative and detailed picture of protein folding dynamics. For example, the {compact → partly folded} and {partly folded → unfolded} transitions are coupled in IMS spectra. These processes are decoupled in FAIMS/IMS (e.g., for 7+ {A → C, C → F} is replaced by {A → A1, C → C2}), which enables accurate measurement of pertinent thermodynamic and kinetic properties. This capability is particularly useful to map the dynamics of minor conformers that are hidden under major peaks in IMS *(11)*.

The capability to follow the structural evolution of specific conformers is also provided by coupling two IMS drift tubes using an ion funnel, where ions of specific mobility selected in the first IMS stage are isomerized by heating in the funnel, and the products are characterized using the second IMS *(50)*. Drawing parallels to MS techniques, that platform is analogous to ToF–ToF while the present FAIMS/IMS is analogous to Q-ToF. As with their MS analogs, the duty cycle is much higher for FAIMS/IMS than IMS/IMS, because the duty cycle of FAIMS in selected

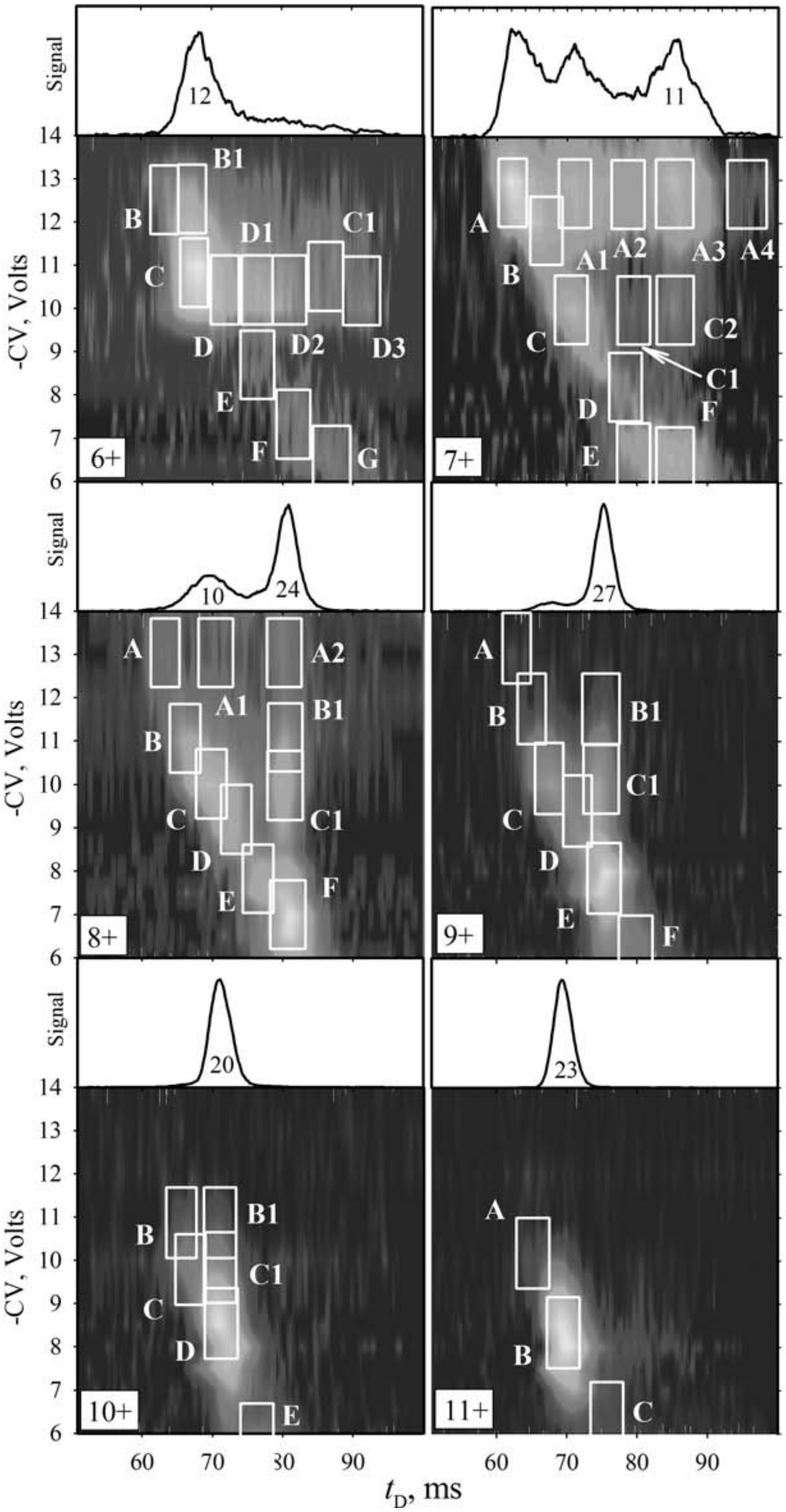


Fig. 7. Same as Fig. 6 for the peak rf of 40 V. FAIMS spectra do not depend on heating in the funnel (positioned after FAIMS) and thus do not differ from those in Fig. 6.

ion monitoring mode is 100% and that of IMS in the double-gate mode is ~1%.

The difference from the comparison between Q-ToF and ToF–ToF is that FAIMS/IMS and IMS/IMS separations of protein conformers appear to be complementary rather than competing techniques. For example *(11)*, FAIMS resolves multiple isomers of ubiquitin or cytochrome *c* with exactly equal mobilities that will not be distinguished by IMS/IMS whereas the t_D width of each spot in 2D palettes indicates the presence of >10 unresolved conformers that should be separable by IMS/IMS. This calls for developing 3D FAIMS/IMS/IMS platforms to study free proteins with maximum depth and specificity.

4. Notes

1. Since ions are delayed (and lost) in each ion guide preceding the ToF chamber, single-stage ToFs with a minimum number of ion guides (usually three) are the best platforms for IMS/ToF systems. We used ESI-ToF (Agilent), which has an excellent MS resolving power (~10,000) and mass accuracy (~3 ppm), or a Q-ToF (Q-Star Pulsar, Sciex, Concord, Canada) with resolving power of ~8,000. Instruments that effect rf ion trapping in the axial direction, for example employing an ion tunnel such as Micromass Ultima (Waters, Milford, MA), have longer transit times (often >10 ms!) and are not suitable for the present application.
2. The Q00 stage houses an rf-only quadrupole (transmitting all ions) of minimum length needed for effective differential pumping (~1 in.). Circular quadrupole rods of 6.3-mm diameter were machined from steel and fixed in a PEEK holder, spaced apart by 5.6 mm. The pumping capacity should preferably be >10 L/s (*see* **Note 3**). The rf waveform on the quadrupole had a frequency of ~912 kHz and peak amplitude of 100 V. To minimize t_F, the dc voltage between the quadrupole and following lens should be as high as possible without causing dissociation of ions via collisional heating. We used an offset of 10 V.
3. We used rotary vane pumps E1 M18 (BOC Edwards, Wilmington, MA) providing 5.7 L/s and Alcatel 2033 (Adixen, Hingham, MA) providing 12.8 L/s (**Fig. 1**).
4. Of the two cylindrical FAIMS geometries commercialized by Thermo Fischer Scientific, we used the longitudinal configuration (where gas flows along the axis) with hemispherical exit cap. In this system, the resolution is adjusted by varying the

gap in spherical region via translation of the inner electrode. The transverse configuration, where resolution is controlled by thermal gradient between electrodes, should also be suitable.

5. The pressure drop inside each short quadrupole is limited by the dimensions of gaps between the rods. Hence, increasing the pumping capacity on each stage by adding multiple pumps is not particularly effective, and reducing the pressure by over an order of magnitude requires multiple differentially pumped stages.
6. For example, the measured K_0 (in N_2 at $T = 25°C$) are ~1.2 $cm^2/(V \times s)$ for $(H^+)_2$bradykinin comprising nine residues (m = 1,060 Da), ~0.8–1.2 $cm^2/(V \times s)$ for ubiquitin (z = 6–14) with 76 residues (m = 8.6 kDa), and ~0.8–1.0 $cm^2/(V \times s)$ for cytochrome c (z = 7–18) with 104 residues (m = 12.3 kDa).
7. Smaller bench-top ToF instruments often operate at higher frequency up to ~100 kHz and thus have lower t_A values down to ~10 μs, which would allow reducing t_P. However, such systems with a short ion flight path suffer from low mass resolution and accuracy, which precludes many important bionalytical applications.
8. The strength of ion focusing in a funnel also scales as (rf amplitude)^2, so P could be elevated roughly in proportion to the rf voltage. Increasing the peak rf amplitude to ~80 V has allowed effective ion focusing at ~30 Torr in N_2 *(40)*. Using high-pressure ion funnels in IMS/MS systems requires multiple differential pumping stages in front of Q0 (**Subheading 3.1.1**).
9. The lowest possible ε (~1.0) are found for gases, and air gaps between electrodes are preferable to any solid. Thus, to the extent permitted by structural rigidity, Teflon insulating sheets may be reduced to washers so that the space between electrodes is mostly empty.
10. Steel chambers may be conveniently built by welding flanges onto cut pieces of tubing. Aluminum chambers are lighter, but do not allow welding gas and electrical feed-throughs. One may choose steel for units with those ports (first, last, and possibly other depending on the design), and aluminum for other units.
11. High-density polyethylene is an economical alternative.
12. The total length of tubular spacers equals the chamber length, but the electrode stack is slightly longer because of finite electrode thickness (0.38 mm). Assembling the unit compresses the spacers, firmly securing all electrodes inside.
13. This resistor chain draws a leak current that depends on the resistances and drift voltage. This current should vastly

exceed the ion current entering IMS, else the electrode voltages would be unstable. Assuming the maximum (pulsed) ion current of 10 nA, the voltage stability to <0.1% calls for the leak current of >10 mA. For $U \geq 4$ kV, this limits the total resistance to <400 MΩ. However, the current should be not as high as to tax the power supply or appreciably heat the drift tube. For the power release of <1 W at $U \leq 10$ kV, the resistance should be >100 MΩ. This guided the choice of 1 MΩ per resistor for the present IMS comprising 210 electrodes.

14. A gas at a reduced pressure inside IMS is a poor electrical insulator: Paschen curves (showing the electrical breakdown field as a function of gas pressure) usually have minima at a few Torr *(41)*. That is a particular problem for the He gas typically used in structural characterization work. While the electrical resistance of buffer gas sets the physical maximum of IMS drift field, reaching it requires that nowhere inside IMS the electric field substantially exceeds the drift field. So the voltage on each IMS chamber should match that of electrodes inside as closely as possible. With each chamber shorted to its median electrode, the voltage between chambers and electrodes inside is, at most, ½ the voltage between neighboring chambers. Still, the distance between electrodes and chamber walls in the present design (6 cm) is less than ½ of the chamber length, and the discharge from electrodes to chambers (or between adjacent chambers through the holes in the insulating disk inside) is the preferred electrical breakdown mode that limits the drift voltage in present IMS system.
15. The resulting dc field of 15 V/cm is lower than that in the exit IMS funnel. This is due not to a different funnel geometry, but to the maximum dc bias of FAIMS (**Subheading 3.1.6**). Floating of FAIMS on the hourglass funnel (in addition to their present floating on the IMS front voltage) would allow increasing the dc field in the funnel to its optimum level.
16. This is convenient to do using the valve on gas line into IMS, which provides a fine control of the drift tube pressure. The pressure is first raised to suppress all ion signals, then gradually decreased to the point where measured IMS/MS spectra cease to change.
17. Nano-ESI emitters were made by pulling sections of silica capillary with a butane torch. As with any ESI application, other emitters including commercial ones may be substituted depending on the solution flow rate and other factors.
18. The gap in transverse-geometry FAIMS (*see* **Note 4**) is 2.5 mm, and same values of DF and CF translate into DV = –5 kV and CV = –(7.5–42.5) V, respectively.

19. For an unknown sample, the CV range may be rapidly determined by a FAIMS scan with ToF in the total ion count (TIC) mode and IMS gate permanently open to provide continuous ion transmission.
20. Ions could also be heated by increasing the dc field in the funnel, but that decreases focusing (and thus ion signal) while raising the rf field improves focusing.

Acknowledgments

The development of 2D FAIMS/IMS separations at Pacific Northwest National Laboratory (PNNL) has greatly benefited from collaboration with and contributions by Dr. Fumin Li, David Prior, Michael Buschbach, Gordon Anderson, and Heather Mottaz. Portions of this work were supported by the PNNL Laboratory Directed Research and Development Program, Battelle Industrial Research Development Program, and the National Institute of Health (NCRR Grant RR018522 and NCI Grant CA126191). The research was performed using EMSL, a national scientific user facility located at PNNL and sponsored by the U.S. Department of Energy's Office of Biological and Environment Research).

References

1. Aebersold, R. and Mann, M. (2003) Mass-spectrometry based proteomics. *Nature* 422, 198–207
2. Purves, R. W. and Guevremont, R. (1999) Electrospray ionization high-field asymmetric waveform ion mobility spectrometry – mass spectrometry. *Anal. Chem.* 71, 2346–2357
3. Mason, E. A. and McDaniel, E. W. (1988) *Transport Properties of Ions in Gases.* Wiley, New York
4. Buryakov, I. A., Krylov, E. V., Nazarov, E. G., and Rasulev, U. K. (1993) A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude asymmetric strong electric field. *Int. J. Mass Spectrom. Ion Processes* 128, 143–148
5. Shvartsburg, A. A., Tang, K., and Smith, R. D. (2004) Modeling the resolution and sensitivity of FAIMS analyses. *J. Am. Soc. Mass Spectrom.* 15, 1487–1498
6. Shvartsburg, A. A., Li, F., Tang, K., and Smith, R. D. (2006) High-resolution field asymmetric waveform ion mobility spectrometry using new planar geometry analyzers. *Anal. Chem.* 78, 3706–3714
7. Barnett, D. A., Ells, B., Guevremont, R., Purves, R. W., and Viehland, L. A. (2000) Evaluation of carrier gases for use in high-field asymmetric waveform ion mobility spectrometry. *J. Am. Soc. Mass Spectrom.* 11, 1125–1133
8. Guevremont, R., Barnett, D. A., Purves, R. W., and Vandermey, J. (2000) Analysis of a tryptic digest of pig hemoglobin using ESI-FAIMS-MS. *Anal. Chem.* 72, 4577–4584
9. Shvartsburg, A. A., Mashkevich, S. V., and Smith, R. D. (2006) Feasibility of higher-order differential ion mobility separations using new asymmetric waveforms. *J. Phys. Chem. A* 110, 2663–2673
10. Ruotolo, B. T., McLean, J. A., Gillig, K. J., and Russell, D. H. (2004) Peak capacity

of ion mobility mass spectrometry: the utility of varying drift gas polarizability for the separation of tryptic peptides. *J. Mass Spectrom.* 39, 361–367

11. Shvartsburg, A. A., Li, F., Tang, K., and Smith, R. D. (2006) Characterizing the structures and folding of free proteins using 2-D gas-phase separations: observation of multiple unfolded conformers. *Anal. Chem.* 78, 3304–3315; *ibid* 8575
12. Bernstein, S. L., Wyttenbach, T., Baumketner, A., Shea, J. E., Bitan, G., Teplow, D. B., and Bowers, M. T. (2005) Amyloid β-protein: monomer structure and early aggregation states of Aβ42 and its Pro^{19} alloform. *J. Am. Chem. Soc.* 127, 2075–2084
13. Borysik, A. J. H., Read, P., Little, D. R., Bateman, R. H., Radford, S. E., and Ashcroft, A. E. (2004) Separation of β_2-microglobulin conformers by high-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled to electrospray ionisation mass spectrometry. *Rapid Commun. Mass Spectrom.* 18, 2229–2234
14. Carr, T. W. (1984) *Plasma Chromatography*. Plenum, New York
15. Hoaglund, C. S., Valentine, S. J., Sporleder, C. R., Reilly, J. P., and Clemmer, D. E. (1998) Three-dimensional ion mobility/TOFMS analysis of electrosprayed biomolecules. *Anal. Chem.* 70, 2236–2242
16. Kapron, J. T., Jemal, M., Duncan, G., Kolakowski, B., and Purves, R. (2005) Removal of metabolite interference during liquid chromatography/tandem mass spectrometry using high-field asymmetric waveform ion mobility spectrometry. *Rapid Commun. Mass Spectrom.* 19, 1979–1983
17. Pringle, S. D., Giles, K., Wildgoose, J. L., Williams, J. P., Slade, S. E., Thalassinos, K., Bateman, R. H., Bowers, M. T., and Scrivens, J. H. (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass Spectrom.* 261, 1–12
18. Barnett, D. A., Ells, B., Guevremont, R., and Purves, R. W. (2002) Application of ESI-FAIMS-MS to the analysis of tryptic peptides. *J. Am. Soc. Mass Spectrom.* 13, 1282–1291
19. Venne, K., Bonneil, E., Eng, K., and Thibault, P. (2005) Improvement in peptide detection for proteomics analyses using nanoLC-MS and high-field asymmetry waveform ion mobility mass spectrometry. *Anal. Chem.* 77, 2176–2186
20. Tang, K., Li, F., Shvartsburg, A. A., Strittmatter, E. F., and Smith, R. D. (2005) Two-dimensional gas-phase separations coupled to mass spectrometry for analysis of complex mixtures. *Anal. Chem.* 77, 6381–6388
21. Srebalus, C. A., Li, J., Marshall, W. S., Clemmer, D. E. (1999) Gas-phase separations of electrosprayed peptide libraries. *Anal. Chem.* 71, 3918–3927
22. Asbury, G. R. and Hill, H. H. (2000) Evaluation of ultrahigh resolution ion mobility spectrometry as an analytical separation device in chromatographic terms. *J. Microcolumn. Sep.* 12, 172–178
23. Tang, K., Shvartsburg, A. A., Lee, H. N., Prior, D. C., Buschbach, M. A., Li, F., Tolmachev, A. V., Anderson, G. A., and Smith, R. D. (2005) High-sensitivity ion mobility spectrometry/mass spectrometry using electrodynamic ion funnel interfaces. *Anal. Chem.* 77, 3330–3339
24. Smith, R. D., Shen, Y., and Tang, K. (2004) Ultrasensitive and quantitative analyses from combined separations-mass spectrometry for the characterization of proteomes. *Acc. Chem. Res.* 37, 269–278
25. McLean, J. A., Ruotolo, B. T., Gillig, K. J., and Russell, D. H. (2005) Ion mobility – mass spectrometry: a new paradigm for proteomics. *Int. J. Mass Spectrom.* 240, 301–315
26. Taraszka, J. A., Kurulugama, R., Sowell, R., Valentine, S. J., Koeniger, S. L., Arnold, R. J., Miller, D. F., Kaufman, T. C., and Clemmer, D. E. (2005) Mapping the proteome of drosophila melanogaster: analysis of embryos and adult heads by LC-IMS-MS methods. *J. Proteome Res.* 4, 1223–1237
27. Robinson, E. W. and Williams, E. R. (2005) Multidimensional separations of ubiquitin conformers in the gas phase: relating ion cross sections to H/D exchange measurements. *J. Am. Soc. Mass Spectrom.* 16, 1427–1437
28. Purves, R. W., Ells, B., Barnett, D. A., and Guevremont, R. (2005) Combining H-D exchange and ESI-FAIMS-MS for detecting gas-phase conformers of equine cytochrome *c*. *Can. J. Chem.* 83, 1961–1968
29. Mesleh, M. F., Hunter, J. M., Shvartsburg, A. A., Schatz, G. C., and Jarrold, M. F. (1996) Structural information from ion mobility measurements: effects of the long-range potential. *J. Phys. Chem.* 100, 16082–16086
30. Shvartsburg, A. A., Liu, B., Siu, K. W. M., and Ho, K. M. (2000) Evaluation of ionic mobilities by coupling the scattering

on atoms and on electron density. *J. Phys. Chem. A* 104, 6152–6157

31. Washburn, M. P., Wolters, D., and Yates, J. R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19, 242–247
32. Shvartsburg, A. A., Bryskiewicz, T., Purves, R. W., Tang, K., Guevremont, R., and Smith, R. D. (2006) Field asymmetric waveform ion mobility spectrometry studies of proteins: dipole alignment in ion mobility spectrometry? *J. Phys. Chem. B* 110, 21966–21980
33. Kim, T., Tolmachev, A. V., Harkewicz, R., Prior, D. C., Anderson, G., Udseth, H. R., Smith, R. D., Bailey, T. H., Rakov, S., and Futrell, J. H. (2000) Design and implementation of a new electrodynamic ion funnel. *Anal. Chem.* 72, 2247–2255
34. Jarrold, M. F. and Bower, J. E. (1992) Mobilities of silicon cluster ions: the reactivity of silicon sausages and spheres. *J. Chem. Phys.* 96, 9180–9190
35. Siems, W. F., Wu, C., Tarver, E. E., Hill, H. H., Larsen, P. R., and McMinn, D. G. (1994) Measuring the resolving power of ion mobility spectrometers. *Anal. Chem.* 66, 4195–4201
36. Dugourd, Ph., Hudgins, R. R., Clemmer, D. E., and Jarrold, M. F. (1997) High-resolution ion mobility measurements. *Rev. Sci. Instrum.* 68, 1122–1129
37. Javahery, G. and Thomson, B. (1997) A segmented radiofrequency-only quadrupole collision cell for measurements of ion collision cross section on a triple quadrupole mass spectrometer. *J. Am. Soc. Mass Spectrom.* 8, 697–702
38. Hoaglund-Hyzer, C. S., Lee, Y. J., Counterman, A. E., and Clemmer, D. E. (2002) Coupling ion mobility separations, collisional activation techniques, and multiple stages of MS for analysis of complex peptide mixtures. *Anal. Chem.* 74, 992–1006
39. Guevremont, R. and Purves, R. W. (1999) Atmospheric pressure ion focusing in a high-field asymmetric waveform ion mobility spectrometer. *Rev. Sci. Instrum.* 70, 1370–1384
40. Ibrahim, Y., Tang, K., Tolmachev, A. V., Shvartsburg, A. A., and Smith, R. D. (2006) Improving mass spectrometer sensitivity using a high-pressure electrodynamic ion funnel interface. *J. Am. Soc. Mass Spectrom.* 17, 1299–1305
41. Meek, J. M. and Craggs, J. D. (1978) *Electrical Breakdown of Gases.* Wiley, New York
42. Kim, T., Tang, K., Udseth, H. R., and Smith, R. D. (2001) A multicapillary inlet jet disruption electrodynamic ion funnel interface for improved sensitivity using atmospheric pressure ion sources. *Anal. Chem.* 73, 4162–4170
43. Page, J. S., Bogdanov, B., Vilkov, A. N., Prior, D. C., Buschbach, M. A., Tang, K., and Smith, R. D. (2005) Automatic gain control in mass spectrometry using a jet disrupter electrode in an electrodynamic ion funnel. *J. Am. Soc. Mass Spectrom.* 16, 244–253
44. Page, J. S., Tolmachev, A. V., Tang, K., and Smith, R. D. (2005) Variable low-mass filtering using an electrodynamic ion funnel. *J. Mass Spectrom.* 40, 1215–1222
45. Purvine, S., Picone, A. F., and Kolker, E. (2004) Standard mixtures for proteome studies. *OMICS* 8, 79–92
46. Guevremont, R., Thekkadath, G., and Hilton, C. K. (2005) Compensation voltage (CV) peak shapes using a domed FAIMS with the inner electrode translated to various longitudinal positions. *J. Am. Soc. Mass Spectrom.* 16, 948–956
47. Ruotolo, B. T. and Robinson, C. V. (2006) Aspects of native proteins are retained in vacuum. *Curr. Opin. Chem. Biol.* 10, 402–408
48. Li, J., Taraszka, J. A., Counterman, A. E., and Clemmer, D. E. (1999) Influence of solvent composition and capillary temperature on the conformations of electrosprayed ions: unfolding of compact ubiquitin conformers from pseudonative and denatured solutions. *Int. J. Mass. Spectrom.* 185/186/187, 37–47
49. Shelimov, K. B., Clemmer, D. E., Hudgins, R. R., and Jarrold, M. F. (1997) Protein structure *in vacuo*: gas-phase conformations of BPTI and cytochrome *c. J. Am. Chem. Soc.* 119, 2240–2248
50. Koeniger, S. L., Merenbloom, S. I., and Clemmer, D. E. (2006) Evidence for many resolvable structures within conformation types of electrosprayed ubiquitin ions. *J. Phys. Chem. B* 110, 7017–7021

Chapter 27

Proteomics for Validation of Automated Gene Model Predictions

Kemin Zhou, Ellen A. Panisko, Jon K. Magnuson, Scott E. Baker, and Igor V. Grigoriev

Summary

High-throughput liquid chromatography mass spectrometry (LC-MS)-based proteomic analysis has emerged as a powerful tool for functional annotation of genome sequences. These analyses complement the bioinformatic and experimental tools used for deriving, verifying, and functionally annotating models of genes and their transcripts. Furthermore, proteomics extends verification and functional annotation to the level of the translation product of the gene model.

Key words: Protein, Peptide, Proteomics, Mass spectrometry, LC-MS, Gene model, Genome, Proteome, Annotation, Splice site, Intron, Exon, BLAST, FASTA, Eukaryote, *Phanerochaete chrysosporium*.

1. Introduction

The explosion of genome sequencing projects in the last decade has been an enormous boon to biological researchers. However, the quantity of DNA sequence data presents a challenge for annotation and usage. The development and refinement of algorithms for annotation of genome sequences have been crucial for providing in silico predicted gene models. This is especially true for eukaryotic genomes with the additional complexity of introns and exons. Expressed sequence tag (EST) data are critical for experimental verification of predicted transcripts from the

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_27

gene models. The placement of splice sites and information about transcription start sites and untranslated regions of the gene can be derived from the EST data.

Analogous to the use of EST data, the use of high-throughput LC-MS-based proteomic data to experimentally verify and adjust predicted models for translation products of the gene models has begun to emerge. An organism can be grown under diverse sets of conditions and the extracted proteins pooled before analysis in order to maximize the proportion of the proteome that is observed. Alternatively, specific growth conditions at specific developmental stages or time points can be analyzed separately to extend the functional annotation of a proportion of the proteome, e.g., associate a subset of proteins with a particular metabolic state. Similar to EST data, proteomics data can be used to verify splice sites in transcripts where the identified peptides span an exon/intron/exon boundary. But proteomics data add an additional dimension of functional annotation, as it can be used to verify predicted translation start and stop sites, signal peptide cleavage sites, and posttranslational protein modifications.

2. Materials

1. *NCBI BLAST*: The programs needed for BLAST-based mapping of peptides are available for a variety of operating systems for free download from the NCBI (www.ncbi.nlm.nih.gov/BLAST/download.shtml) *(1)*.
2. *Genome sequence*: A complete or high-coverage draft genome sequence.
3. *Gene models*: Gene models are generated for a given genome by automated gene calling software, such as Genewise *(2)* or Fgenesh *(3)*.
4. *Peptide sequence data from global proteomic experiment(s)*: Must be in a format, such as FASTA, that is compatible with command-line batch BLAST analysis.

3. Methods

3.1. Proteomic Analysis

Global proteomic analysis is used to determine the sequences of peptides present in a protein sample (*see* **Notes 1** and **2**).

3.2. Peptide Mapping

Peptides that have been identified in the previous step are mapped to the genome sequence using the tblastn tool with the following options: tblastn -e 1,000-W2-FF-f 6-K50.

The expectation value (-e option) is given the large value of 1,000 in order to collect short matches to the genome. A very short word size of 2 (-W option) is used to increase the sensitivity of the blast algorithm. Turning off the filtering (-F option) is also essential for successfully matching short peptides to the genome. A lower threshold value is used to extend the length of the matches (-f option). Because the lower threshold generates a lot of matches, and only one or two are true matches, the number of alignments reported is lowered to 50 (-K option).

3.3. Quality Assessment

After the peptides are mapped to the genome, a best match is selected for each peptide. The quality of the peptide matches is assessed with respect to percent sequence identity and percent coverage. The peptide is further categorized based on the presence or absence of a gap. Based on the collective criteria the peptides are assigned to one of the following categories:

Perfect match	100% Sequence identity without gaps in the alignment covering more than 98% of the peptide sequence
Split match	Two nonoverlapping fragments of a peptide exhibit perfect matches to two genomic locations in close proximity. Total coverage is equal to 100% of the peptide
Imperfect match	80–99% sequence identity, 90–98% coverage
Imperfect split	Two fragments of a peptide with maximum overlap of two amino acids have imperfect matches to two regions of genome sequence in close proximity. Total coverage is close to the entire peptide length
Uncertain	The remainder of the mapped peptides

Several top matches of the same quality are retained to reflect gene duplication or nonunique peptides in the genome sequence (*see* **Note 1** and **Fig. 1** for an example).

3.4. Validation of Predicted Genes

Comparison of the coordinates of mapped peptides with the coordinates of predicted gene models in the genome sequence provides experimental support for the predicted genes. Given sufficient peptide coverage of a predicted gene model, the boundaries of the protein coding portion of the gene and splice sites can be verified by the experimental proteomics data (*see* **Subheading 4** for an example).

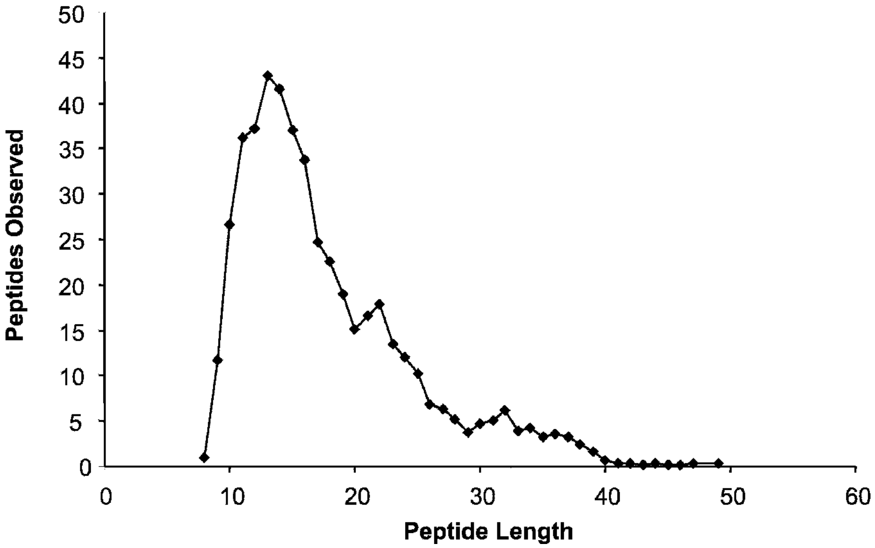


Fig. 1. Peptide length distribution for the set of *Phanerochaete chrysosporium* peptides in the example (*see* Note 1).

4. Notes

1. *Example*. 4,825 peptides were used for mapping to the version 2.0 genome assembly of *Phanerochaete chrysosporium* and for validations of predicted gene models (v2.1) for this assembly. The average peptide length was 14 amino acids. The distribution of peptide lengths is shown in **Fig. 1**. The peptides were mapped to 5,135 locations on the genome including 4,149 (81%) perfect matches (**Fig. 2**). The difference between the number of locations and the number of total peptides reflects the existence of nonunique peptide matches and split matches. A total of 224 split matches were detected with 107 perfect and 117 imperfect split matches providing support for splice boundaries in genes containing two or more exons. Excluding the uncertain peptides, support was provided for 2,795 exons representing 1,440 of 10,048 predicted gene models.
2. *Database searching*. There are a number of different algorithms available for searching mass spectrometry generated peptide data against protein databases. In the example in **Note 1**, Sequest was used with the following cutoffs for fully tryptic peptides: charge state 1: Xcorr ≥ 1.9; charge state 2: Xcorr ≥ 2.2; charge state 3: Xcorr ≥ 3.75; all must have a $\Delta CN^2 \geq 0.1$. Xcorr is a statistical estimate of the

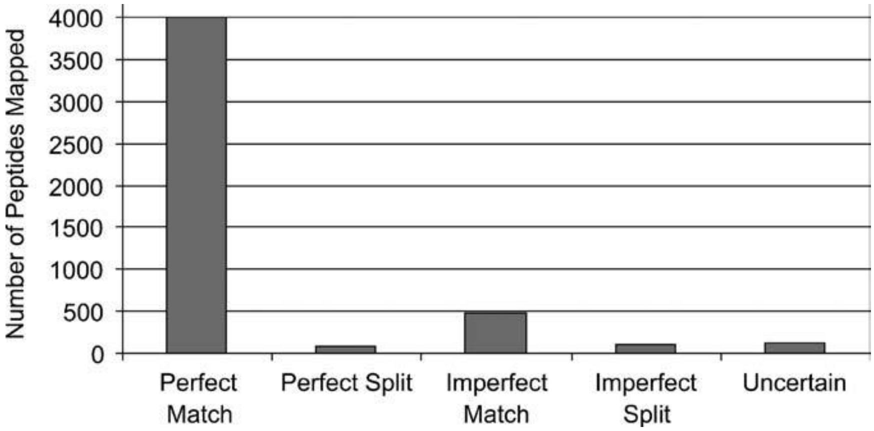
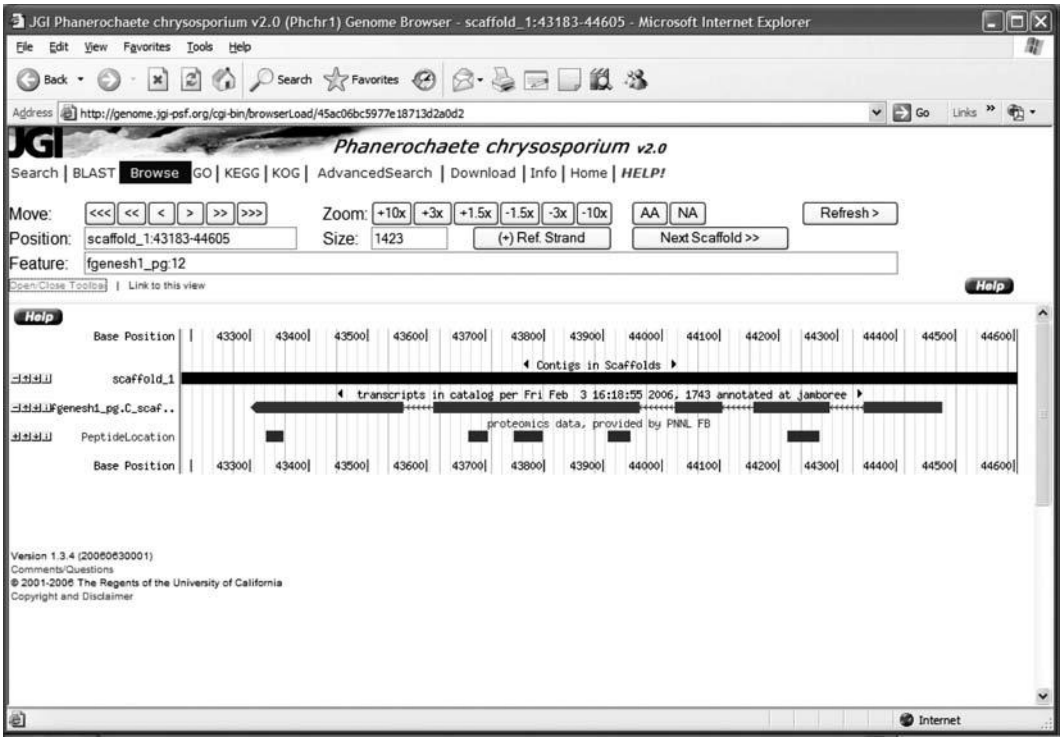


Fig. 2. Distribution of mapped *P. chrysosporium* peptides according to the quality categories described in Subheading 3.3.


JGI Phanerochaete chrysosporium v2.0 (Phchr1) Genome Browser - scaffold_1:43183-44605 - Microsoft Internet Explorer

Address: http://genome.jgi-psf.org/cgi-bin/browserLoad/45ac06bc5977e18713d2a0d2

Position: scaffold_1:43183-44605

Size: 1423

Feature: fgenesh1_pg:12

Version 1.3.4 (20060630001)
Comments/Questions
© 2001-2006 The Regents of the University of California
Copyright and Disclaimer

File Edit View Favorites Tools Help
Back
Search
Favorites
Go
Links
JGI
Phanerochaete chrysosporium v2.0
Search | BLAST | Browse | GO | KEGG | KOG | AdvancedSearch | Download | Info | Home | HELP!
Move: <<< << < > >> >>>
Zoom: +10x +3x +1.5x -1.5x -3x -10x
AA NA
Refresh >
(+) Ref. Strand
Next Scaffold >>
Open/Close Toolbar | Link to this view
Help
Help
Base Position 43300 43400 43500 43600 43700 43800 43900 44000 44100 44200 44300 44400 44500 44600
Contigs in Scaffolds
scaffold_1
transcripts in catalog per Fri Feb 3 16:18:55 2006, 1743 annotated at Jamboree
fgenesh1_pg.C_scaf..
proteomics data, provided by PNNL FB
PeptideLocation
Base Position 43300 43400 43500 43600 43700 43800 43900 44000 44100 44200 44300 44400 44500 44600
Internet

Fig. 3. Display of peptide data on the JGI *P. chrysosporium* Genome Browser. Genome scaffold *black*, gene model *red*, mapped peptides *green* (*see* Note 3). (*See Color Plates*)

cross-correlation sequence of a random process. In this application, it is a measure of the quality of the match of the peptide derived from the mass spectral data to the peptide determined from translation of the genome sequence.

In general, the most important aspect of accurate peptide identification is the use of stringent parameters, regardless of the algorithm.

3. *Data display*. The procedure is designed for work with the JGI Genome Portal (**Fig. 3**; genome.jgi-psf.org) and the underlying mySQL database. However, this procedure can easily be modified for any genome browser that displays *features*, such as gene models, on *tracks* mapped back to genome sequence contigs or scaffolds. The generic model organism database (GMOD) project has produced a mostly open-source visual genome database that displays information *tracks* (www.gmod.org). *(4)* Visual display is critical for manual genome curation. At a genome-wide level, text-based lists of gene models with associated numbers and coordinates of peptides can be used effectively for analysis of automated annotation.

References

1. Altschul, SF, Madden, TL, Schäffer, AA, Zhang, J, Zhang, Z, Miller, W, Lipman, DJ. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402
2. Birney, E. Durbin, R. 2000 Using GeneWise in the *Drosophila* annotation experiment. *Genome Res* 10, 547–548
3. Salamov AA Solovyev VV. 2000 Ab initio gene finding in *Drosophila* genomic DNA. *Genome Res.* 10, 516–522
4. Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, Lewis S. 2002. The generic genome browser: a building block for a model organism system database. *Genome Res.* 12(10):1599–610

Chapter 28

Support Vector Machines for Improved Peptide Identification from Tandem Mass Spectrometry Database Search

Bobbie-Jo M. Webb-Robertson

Summary

Accurate identification of peptides is a current challenge in mass spectrometry (MS)-based proteomics. The standard approach uses a search routine to compare tandem mass spectra to a database of peptides associated with the target organism. These database search routines yield multiple metrics associated with the quality of the mapping of the experimental spectrum to the theoretical spectrum of a peptide. The structure of these results make separating correct from false identifications difficult and has created a false identification problem. Statistical confidence scores are an approach to battle this false positive problem that has led to significant improvements in peptide identification. We have shown that machine learning, specifically support vector machine (SVM), is an effective approach to separating true peptide identifications from false ones. The SVM-based peptide statistical scoring method transforms a peptide into a vector representation based on database search metrics to train and validate the SVM. In practice, following the database search routine, a peptide is denoted in its vector representation and the SVM generates a single statistical score that is then used to classify presence or absence in the sample.

Key words: Peptide identification, Database search routine, Statistics, Machine learning, SVM.

1. Introduction

Intrinsic to protein identification via MS-based proteomics is the comparison of tandem mass spectrometry (MS/MS) fragmentation patterns to a known genome to provide the basic peptide identification upon which all other measurements are based *(1–3)*. These database search routines are known to return both correct identifications against the experimental spectra, as well as a similar number of false positives. Complicating the identification process is multiple metrics of identification associated with peptide identifications. For example, the algorithm SEQUEST

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*
© Humana Press, a part of Springer Science+Business Media, LLC 2009
DOI: 10.1007/978-1-59745-493-3_28

(3) returns values such as *Xcorr*, *delCN*, and *RSp*, and the algorithm NWPolygraph *(1)* yields scores such as *Entropy* and *Likelihood*. Rule-based approaches are one simple method by which to filter true from false identifications but lack statistical rigor and are difficult to quantify. The current approach to improve accuracy of peptide identifications after the database search stage is to integrate the multiple metrics of a peptide identification algorithm into a single discriminant-type score *(4–7)*. The approaches are varied, but primarily fall into three categories: *(1)* Bayesian statistics, *(2)* traditional hypothesis tests, and *(3)* machine learning – specifically support vector machines (SVM). In most cases, availability of code to generate these statistical measures of confidence does not exist, with the exception of the popular Protein Prophet *(5)* software publicly available at http://sourceforge.net/projects/proteinprophet. However, even this code is specific to the SEQUEST algorithm metrics used for generating the distributions and hence cannot be used with any other database search routine without developing a new discriminant function.

SVMs are an excellent supervised learning approach to classification that have the ability to separate nonlinear data and are highly robust to overtraining *(89)*. This has led to an explosion of their use in computational biology *(10)*. We have demonstrated that higher accuracy can be achieved if a peptide-specific SVM score is used in lieu of NWPolygraph database search metrics *(1)*. A similar approach has been demonstrated for SEQUEST *(4)*. Similar to Protein Prophet, each of these approaches are also linked to a specific database search algorithm, but the method is easily generalized.

2. Materials

Generating a SVM classifier and subsequent scoring of peptides identified by MS/MS is a purely computational task, and thus the only physical material needed is a computer with sufficient computational power and memory. In addition, there are two computational based materials needed, the SVM code and the benchmark dataset.

2.1. Svm Code Access

To generate and validate the SVM model, code to perform the optimization and classification tasks must be available. Any SVM code can be used. A popular option is the gradient assent algorithm implementation GIST, http://svm.sdsc.edu. It offers both, a web-interface that can be used from any operating system as well as the C code that can be compiled on UNIX-based machines and run for more efficient and memory intensive applications. Additionally, it offers code separate from the training stage to run the SVM model.

2.2. Benchmarking Dataset

SVM learning requires example data from which to build a classifier. In this case it requires a set of peptides that are both *correct* and *incorrect* identifications for a given database search routine. The metrics of the database search routine and other properties of the peptides (or spectra) are then compiled into a dataset that will be used for differentiation. The dataset used in the Polygraph implementation *(1)* is available at http://www.systemsbiology.org/protein_mixture.html *(11)*. Briefly, a standardized mixture of peptides from the organism *Deinococcus radiodurans* was developed and analyzed repeatedly. SEQUEST was used to identify peptides using a stringent selection criterion. The standard mixture was also run against a decoy database of 88,000 peptides derived from the human genome. Identifications against this independent set gave a set of negative examples meeting the identification criteria. An alternative example dataset used with the SEQUEST implementation can be found at http://noble.gs.washington.edu/proj/ms-svm *(4)*.

3. Methods

There are four primary steps in developing a SVM peptide scoring model: *(1)* peptide vectorization, *(2)* normalization, *(3)* model development and validation, and *(4)* deployment.

3.1. Peptide Data Vectorization

In order to develop a SVM model, the data must be in a multivariate representation. Peptides in the amino acid string form are not amenable to SVM classification and thus must be transformed into a numeric representation (vectorize). For example, given n properties, the peptide is described as

$$\mathrm{pep}_k = [X_1, X_2, \ldots, X_n].$$

The definition of these properties can be solely based on metrics from the database search routine or include amino acid and spectral properties (*see* **Note 1**).

3.2. Normalization

The data must be normalized to assure equal weight to each variable. Column normalization is performed by simply dividing each column by its total. The normalization constants (sum of each column) must be saved for deployment purposes (**Subheading 3.4**).

3.3. Model Development and Validation

In order to evaluate the model it must be assured that the value attained for each peptide in the benchmark dataset is independent from the data it was trained on. This is achieved through cross-validation (CV). **Figure. 1** gives a pictorial view of **steps 1–4** described here using threefold CV and a dataset of size 2,983 peptides (1,425 true and 1,558 false) as the example.

1. *Generate test and train files.* Testing files are generated by randomly splitting the true and false data into *k* files (*k*-fold CV). The *k*th test file has a corresponding training file that is the union of all test files except for the *k*th test file.
2. Repeat the training and testing phase *k* times
   (a) Train the SVM
   - The SVM training code is run for the kth training file. The kernel choice is selected, common choices are linear, quadratic, or radial basis function (rbf) (see Note 2) and typically normalized. This normalization is an option within the SVM code. Once the parameters of the SVM are chosen, the training routine is run and a predict file is generated.

   (b) Test the SVM
   - Publicly available SVM codes have a classify routine that uses the *predict* file generated in (a) to attain SVM scores for each observation in the test file. The output of this step is a SVM score associated with each peptide in the *k*th test file.
3. *Compile k SVM scores.* Each of the *k* runs in **step 2** will produce *k* sets of SVM scores (S_i), each the same size as the test sets. These *k* sets of SVM scores can be united into a single list of scores of equal size and directly related to the original dataset of size *n*; $S=[S_1 \cup S_2 \cup \ldots \cup S_K]$.
4. *Evaluate the true and false positive structure.* A receiver operating characteristic (ROC) curve yields a measure of the true positive rate (TPR) at any given false positive rate (FPR). The ROC curve is generated by sorting the scores in descending order and calculating the TPR (number of true identifications above a threshold divided by the total number of positives) at each false identification and FPR (number of false identifications above a threshold divided by the total number of negatives). The ROC curve can be used to tailor the classifier to a specific false discovery rate (*see* **Note 3**).

3.4. Deployment: Final Svm Classifier

The entire vectorized and normalized dataset, **Subheadings 3.1 and 3.2**, is used to generate the final classifier. It is trained with the SVM to generate a global *predict* file. Publicly available software will have a classify routine, but the peptides must be transformed into the proper format prior to use. These steps are

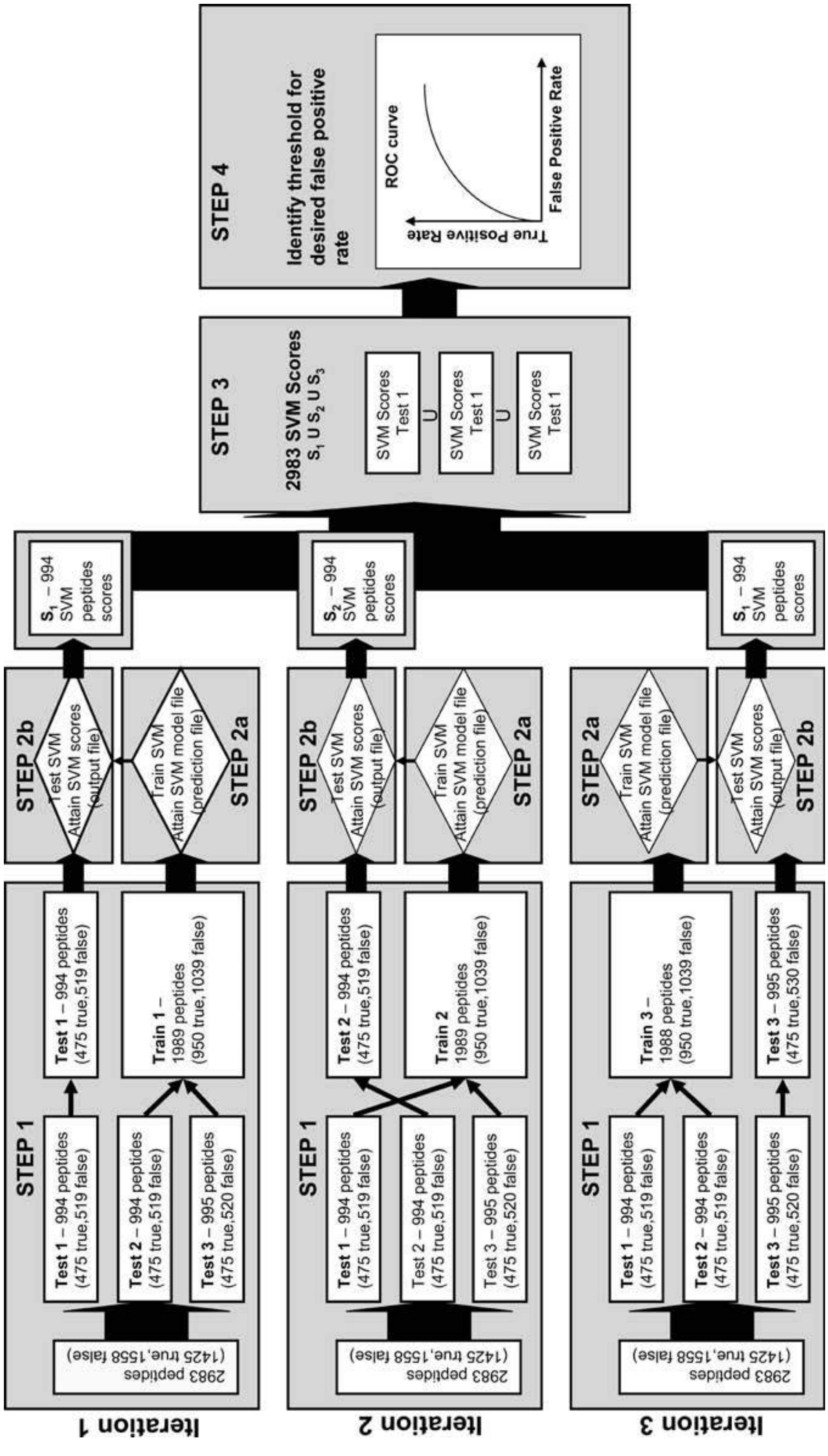

Iteration 1
Iteration 2
Iteration 3
STEP 1
2983 peptides (1425 true,1558 false)
Test 1 – 994 peptides (475 true,519 false)
Test 2 – 994 peptides (475 true,519 false)
Test 3 – 995 peptides (475 true,520 false)
Test 1 – 994 peptides (475 true,519 false)
Train 1 – 1989 peptides (950 true,1039 false)
Test 2 – 994 peptides (475 true,519 false)
Train 2 1989 peptides (950 true,1039 false)
Train 3 – 1988 peptides (950 true,1039 false)
Test 3 – 995 peptides (475 true,530 false)
STEP 2a
Train SVM Attain SVM model file (prediction file)
STEP 2b
Test SVM Attain SVM scores (output file)
S1 – 994 SVM peptides scores
S2 – 994 SVM peptides scores
S1 – 994 SVM peptides scores
STEP 3
2983 SVM Scores S1 U S2 U S3
SVM Scores Test 1
U
SVM Scores Test 1
U
SVM Scores Test 1
STEP 4
Identify threshold for desired false positive rate
ROC curve
True Positive Rate
False Positive Rate

Fig. 1. Overall cross-validation approach to develop and validate the SVM classifier.

serial and can be automated using a scripting language such as PERL.

1. Vectorize the peptide, as in **Subheading 3.1**.
2. Divide each element of the vector by the normalization constants from **Subheading 3.2**.
3. Run the SVM classify routine using the same parameters used to generate the global *predict* file in **Subheading 3.3**.
4. Extract the score from the output file.

4. Notes

1. *The generalizability of peptide vectorization.* The vectorization step can be done for any database search routine. For example, the NWPolygraph *(1)* example used six scores (a) the likelihood ratio, (b) the likelihood ratio difference between the top two scoring peptides, (c) intensity statistic, Ω_I, (d) the rank of Ω_I, (e) total number of histidines, lysines, and arginines, and (f) total number of lysines and arginines. Other implementations have included additional information related to the quality of the spectrum associated with the identification *(4)*. This is why the validation stage in **Subheading 3.3** is essential. If new parameters are being used to describe a peptide, then it must be assured that this is an appropriate representation and the overall sensitivity of the model evaluated. In addition, of equal interest to researchers are the variables (or metrics) that are most discriminating between true and false peptide identifications. Strategies, such as Recursive Feature Elimination *(12)*, can be used to evaluate contribution of each variable to the model. Performing such an analysis is not a common component of SVM analyses but could add value to the overall model by eliminating variables that only contribute noise.
2. *Selection of the Kernel function.* A significant challenge in SVMs is the selection of the transformation function, kernel. The kernel is dependent upon the structure underlying the data. **Figure 2** gives an example in two dimensions of how the kernel function transformations allow separation in a linear space for the rbf and three-degree polynomial. The top left of **Fig. 2** represents data well suited for a rbf kernel: on the bottom left is data separable by a three-degree polynomial. The structures of both kernels are represented in each panel, the appropriate kernel in black and the alternate choice in dashed gray. It is evident from this example that good separation would not be achieved if the wrong kernel was applied. Selection of this kernel, however, is difficult; unlike **Fig. 2** the data typi-

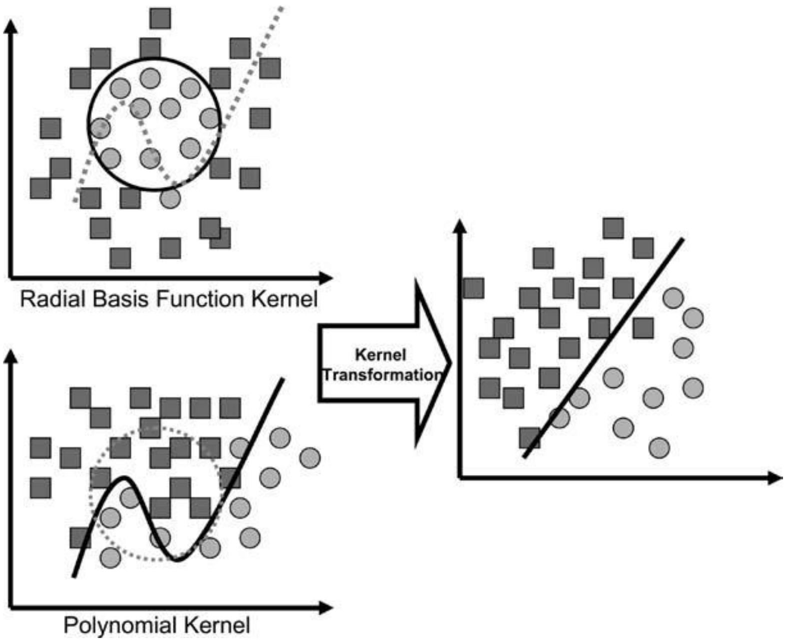


Fig. 2. Pictorial representation of separation via kernel transformation. The rbf kernel does well at separating data captured in circular clusters while the three-degree polynomial kernel separates data that fall into more than one cluster. In both cases, the transformation results in a linear separation.

cally have dimension much greater than 2 and thus a pictorial representation of the structure is not possible. Additionally, methods to directly detect the proper kernel from data are not well developed or easily used. Thus, most often in practice SVM kernels are chosen based on trial and error by repeating the method described in **Subheading 3.3** for multiple kernels, which allows the evaluation of the performance for each kernel and rationale for selection.

3. *Defining the classification threshold*. SVM defines the decision boundary as the function $f(q)=\sum_{i=1}^{N}\alpha_i K(q,x_i)+b,$ where the SVM draws an optimal hyperplane in the feature space (kernel transformation) determined by α and b. The textbook description of classifying a new observation, q, as in the positive or negative class is the sign of this equation. Thus, for peptide identification this is equivalent to if $[f(q)>0]$ then the peptide is a true identification and if $[f(q)<0]$ then it is a false identification. However, any threshold could be used $[f(q)>\mathrm{C}]$, and is more realistic as the selection of C can control the FPR as seen in **Fig. 1**. Additionally, special care must be taken here because if the falses or positives in the training data highly outnumber the other then the classifier scores can end up skewed toward positive or negative values. The SVM

algorithm offers a soft margin that is designed to account for unequal ratios of positive and negative examples. Setting this soft margin is an art that is even less apparent than selecting a kernel function. We have observed cases when the negatives highly outnumber the positives (10–100-fold) where all observations in the test sets always have negative values, but when sorted and plotted have a very good ROC curve. Thus, the defined $(f(q))$is not an appropriate threshold. The ROC curve should be used to select a threshold for deployment.

Acknowledgments

This research was funded by the US Department of Energy (DOE) Office of Advanced Scientific Computing Research under contract No. 47901. The Pacific Northwest National Laboratory is operated by Battelle for U.S. DOE under contract DE-AC06–76RLO 1830.

References

1. Cannon, W. R., Jarman, K. H., Webb-Robertson, B. J., Baxter, D. J., Oehmen, C. S., Jarman, K. D., Heredia-Langner, A., Auberry, K. J., and Anderson, G. A. (2005) Comparison of probability and likelihood models for peptide identification from tandem mass spectrometry data. *J. Proteome Res.* 4, 1687–1698
2. Pappin, D., Rahman, D., Hansen, H., Bartlet-Jones, M., Jeffery, W., and Bleasby, A. (1996) Chemistry, mass spectrometry and peptide-mass databases: Evolution of methods for the rapid identification and mapping of cellular proteins. *Mass Spectrom. Biol. Sci.* 135–150
3. Yates, J. R., III, Eng, J. K., McCormack, A. L., and Schieltz, D. (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal. Chem.* 67, 1426–1436
4. Anderson, D. C., Li, W., Payan, D. G., and Noble, W. S. (2003) A new algorithm for the evaluation of shotgun peptide sequencing in proteomics: support vector machine classification of peptide MS/MS spectra and SEQUEST scores. *J Proteome Res.* 2, 137–146.
5. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 74, 5383–5392
6. Moore, R. E., Young, M. K., and Lee, T. D. (2002) Qscore: an algorithm for evaluating SEQUEST database search results. *J. Am. Soc. Mass Spectrom.* 13, 378–386
7. Strittmatter, E. F., Kangas, L. J., Petritis, K., Mottaz, H. M., Anderson, G. A., Shen, Y., Jacobs, J. M., Camp, D. G., II, and Smith, R. D. (2004) Application of peptide LC retention time information in a discriminant function for peptide identification by tandem mass spectrometry. *J. Proteome Res.* 3, 760–769
8. Cristianini, N., and Shawe-Taylor, J. (2000) *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, Cambridge
9. Vapnik, V. (1995) *The Nature of Statistical Learning Theory*, Springer, New York
10. Scholkopf, B., Tsuda, K., and Vert, J. (ed.) (2004) *Kernel Methods in Computational Biology*, MIT Press, Cambridge
11. Keller, A., Purvine, S., Nesvizhskii, A. I., Stolyar, S., Goodlett, D. R., and Kolker, E. (2002) Experimental protein mixture for validating tandem mass spectral analysis. *Omics.* 6, 207–212
12. Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002) Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422

INDEX

A

AAL. *See* Aleuria aurentia
Accurate mass and time (AMT) tag method........ 40
Acetonitrile........ 325
ACHC. *See* Alpha-cyano-4-hydroxycinnamic acid
ACN. *See* Acetonitrile
Acrylamide isotope labeling. *See also* Cancer patients, plasma profiling in
  materials for........ 311
  methods for........ 312
Affinity chromatography (AC)........ 115
Agilent 1200 Series nanoflow LC system........ 136
Aleuria aurentia........ 325
Alkylation........ 4, 32
  of cysteines........ 59
Alpha-cyano-4-hydroxycinnamic acid........ 78, 110, 370, 371
Amide hydrogen/deuterium (H/D) exchange mass spectrometry........ 256
AMT tag global proteomics strategy........ 40
Anion exchange chromatography, in protein fractionation........ 312–313
Arg-C enzyme........ 34
Atmospheric pressure ion/ion reactions/sampling, 397–399. *See also* Electrospray ionization (ESI)
Atmospheric pressure ionization (API)........ 185

B

Bacillus anthracis........ 374
Bacillus atrophaeus........ 373
Bacillus cereus (Bc)........ 373
Bacillus sphaericus........ 373
Bacillus subtilis (Bs)........ 373
  MALDI-MS of sporulated........ 377
  MALDI-MS spectrum of vegetative........ 376
Bacillus thuringiensis (Bt)........ 373
Bacterial spore mass, estimation of........ 374
BCA assay........ 62
BCA protein assay kit........ 44
Bicinchoninic acid assay........ 44–45
Bioion/ion reactions, tools for, 396–404. *See also* Electrospray ionization (ESI)
Blue-native gel electrophoresis........ 115
Bottom–up proteomic separations
  mobility separations in
    peptide mixture analysis........ 433–436
    proteolytic digestion........ 433
Bovine carbonic anhydrase peptides........ 249–250
Bovine serum albumin........ 50, 292, 325
Bradford assay kit........ 96
Breath-easier tube membranes........ 58
Bromophenol blue........ 111
BSA. *See* Bovine serum albumin

C

Calcium chloride........ 59
Cancer patients, plasma profiling in
  materials for........ 310–311
  methods for........ 312–314
Capillary isoelectric focusing........ 248
Capillary zone electrophoresis. See CZE-MS
Capillary zone electrophoresis, coupling to mass spectrometry........ 202
Carbamidomethylation........ 34
Carbamylation........ 58
Carbohydrate-characteristic oxonium ions, in PID........ 170
CBB G250 staining........ 77, 83
Centrifugal membrane filtration devices, usages of........ 275
Chemical cross-linking, in protein-protein interaction study........ 283–284
  materials for........ 284–286
  in vitro application methods........ 286–288
  in vivo application methods........ 288–290
3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)........ 114
Chromatofocusing (CF) method........ 322, 324–325, 332–333
CIEF. *See* Capillary isoelectric focusing
CIEF/RPLC, yeast cell lysis........ 234
Collisionally induced dissociation (CID)........ 29, 186, 216, 223, 339, 389
Compensation field (CF)........ 418, 420
Computational Proteomics Analysis System........ 314
Concanavalin A (Con A)........ 324
Coomassie blue assay........ 42, 62
Coomassie brilliant blue........ 67
  G250 staining........ 71
CPAS. *See* Computational Proteomics Analysis System
Cross-correlation scores (DeltCn)........ 17
Cross-linking reaction, methods of........ 286–287
Cross-validation (CV)........ 456, 457

Cyanine dyes 94
α-Cyano-4-hydroxycinnamic acid 298
Cystidine deaminase (CDA) enzyme
  ESI-MS spectrum of 222, 223
  FTICR-MS analysis of 224
CZE-MS
  applications of 202
  coupling 204
  protein characterization by
    mass spectrum 208–209
    sample treatment for 204–205
    separation of erythropoietin 208, 210, 211
    separation of fetuin 206–207
    separation procedure 205–206

D

DALT™ gels 108
Data-dependent acquisition 176, 385
Data-dependent MS acquisition (DDA) method. *See* Data-dependent acquisition DDA
DecisionSite 8.0 software 34
Decoy database 27
Decoy proteins 27
Deglycosylated peptides, mass spectrum of 169
2-DE-MALDI-TOFTOF-MS technology 66
De novo peptide sequencing 150–155
2-DE pattern of *Helicobacter pylori* 26695 cellular proteins 67
Deuterium exchange 255
DHB. *See* 2,5-Dihydroxybenzoic acid
Difference in-gel electrophoresis (DIGE) 94
  acetone precipitation 99
  dye reconstitution 99–101
  first dimension, IEF 104–105
  IEF, preparation for 103
    rehydration of immobiline drystrips 103–104
  labeling 34
  materials for 95–98
  MS analysis 110
  peptides, extraction of 110
  protein gel plugs, digestion of 109–110
  protein labeling
    DIGE minimal labeling 101
    DIGE saturation labeling 101–102
    for preparative gels 102–103
  sample labeling 96
  sample lysis 98–99
  scanning gels 107–108
  second dimension, SDS–PAGE 105–107
  spot picking 109
  sypro staining, fixing the gels 108–109
Digestion efficiency 263
Digestor
  cartridge, length selection 246
  materials
    cartridge-based digestion 244
    in LC/MS system 244
    mass spectrometry 245
  methods
    cartridge-based tryptic digestion 245–246
    microfluidics 247
2,5-Dihydroxybenzoic acid 298, 327, 383
Dimethylsulfoxide 284, 327
3D ion traps, in ion/ion reactions 400–402
Dissuccinimidyl tartrate 284
Disulfide exchange-mediated reduction 263
Dithiothreitol 26, 96, 312
DMSO. *See* Dimethylsulfoxide
DST. *See* Dissuccinimidyl tartrate
DTT. *See* Dithiotreitol

E

ECL. *See* Enhanced chemiluminescent (ECL)
Elastase digestion 5
Electrodynamic ion traps, in ion/ion reactions *See also* Electrospray ionization (ESI) ..... 399–404.
Electron capture dissociation (ECD) 216, 223, 389, 413
  for polypeptides structure determination 413–414
    materials for 414
    methods of 414–415
Electron transfer dissociation (ETD) 389
Electron transfer ion/ion reaction 408–409
Electro-osmotic flow (EOF) 210
Electrospray deposition (ESD) 298
Electrospray ionization (ESI) 66, 115, 162, 385
  in peptides and proteins ion reaction 395–396
    bioion/ion reactions, tools for 396–404
Electrospray ionization mass spectrometry
  in protein noncovalent complex study
    materials in 273–274
    procedures for 275–281
Electrospray mass spectrometry 175
Endoproteinase LysC/Trypsin, digestion protocol 4
Enhanced chemiluminescent (ECL) 285
Enolase enzyme, ESI QqTOF mass spectrum of 276, 278
Enzymatic dephosphorylation method 322–323
Erk2 peptide, HCD spectrum of doubly phosphorylated 138
Erythropoietin separation
  acetic acid for 211
  charge-deconvoluted mass spectrum of 210
  glycoforms for 208
Escherichia coli 373
ESD. *See* Electrospray deposition (ESD)
ESI. *See* Electrospray ionization (ESI)

- ESI/corona discharge, 398–399. *See also* Electrospray ionization (ESI)
- ESI emitter 432
- ESI/FAIMS/IMS/ToF MS system
  - diagram of 421
  - safety enclosure 432
- ESI/FAIMS/ToF MS system
  - construction of
    - axial ion focusing in 425–426
    - effective ion injection 430–431
    - ESI emitter 432
    - FAIMS stage and IMS inlet coupling 431–432
    - lossless IMS/MS interfacing 426–429
    - modular IMS drift tube 429–430
    - selection and modification of ToF 423, 425
  - instrument control and data acquisition system 432–433
- ESI-MS. *See* Electrospray ionization mass spectrometry
- ESI QqTOF mass spectra, of α-synuclein 280
- ESIQq-TOF, performance of 207
- ESI-QTOF systems, in noncovalent protein complex detection 275–276
- ESI-TOF, performance of 207
- ESI-TOF systems, in noncovalent protein complex detection 275–276
- ETD. *See* Electron transfer dissociation (ETD)
- ETTAN Spot Picker 107
- Expressed sequence tag (EST) data 447
- Extracted proteolytic peptides, purification of 188

F

- FAIMS/IMS, 424
  - analysis with 50 CF steps 421
  - dispersion of tryptic peptides in 435
  - effective 2D peak capacity of 435
  - palettes 433
  - protein conformer characterization using
    - dynamics of conformational transitions 438, 440
    - highly charged cytochrome *c* ions 438
    - ubiquitin 436–437
  - structural characterization using, issue with 436
- FAIMS/IMS/MS system
  - instrumentation parts 422–423
  - peptide mixture analysis using
    - CV step 436
    - tryptic digests 433–435
- FAIMS/MS systems
  - components of 419
  - dispersion of tryptic peptides in 435
  - ion separation power of 419
- False discovery rate (FDR) 4, 16–17
- Familial amyloid polyneuropathy (FAP) 353
- Ferulic acid (FA) 370
- Fetuin glycopeptides
  - extracted ion chromatograms of 176
  - molecular weights of 168
- Fetuin separation 206–207
- Field asymmetric waveform IMS (FAIMS)
  - basic principle of 418
  - constraints on separation power of
    - peak capacity 419–420
    - resolving power 420
  - coupled to IMS 431–432
  - ion focusing 427
  - limitations of 420
- Filamentous α-synuclein protein 276
- Finnigan LTQ FT mass spectrometer 414
- Flow rate 27
- Fluorescent phosphodyes, usage of 322
- Fourier Transform Ion Cyclotron Resonance (FTICR) 207, 216, 276, 384
- Fragment ions, high MMA of 131
- Free flow electrophoresis (FFE) 216

G

- Gal-GlcNAc oxonium ions 170
- Gas-phase proton transfer reactions 396
- Gel casting apparatus for IEF gels 69
- Gene models 448
- Generic model organism database (GMOD) project 452
- Genome sequence 448
- *Geobacter sulfurreducens* 40
- Glycoform separation 202
- Glycopeptide 162
  - enrichment 175
  - structures 198
- Glycoprotein 162
  - deglycosylation of 164
  - structure analysis of 182
- Glycosylation 182
- Glycosylation sites 162
  - analysis of 170
- Glycosylation sites, assignment approaches
  - materials used for
    - aldehyde-silica, SNA lectin 174
    - bovine fetuin, tryptic digestion of 165
    - enrichment of fetuin glycopeptides 172
    - glycoprotein 165
    - ion extraction chromatograms of 173
    - LC/MSMS analyses of 165–166
    - macroporous silica and SNA lectin 164
    - microheterogeneity of 174
    - product-ion discovery 166–167
    - tryptic digestion 163–164
- GPS-Explorer™ 2.0 software 33

Gradient delay
  method of 248
  reversed-phase column 248
Gradient solutions, and countergradient solutions 247

H

Halobacterium salinarium............ 25
  growth conditions 29–30
  ICPL methodology, for analysis of
      purified membranes 34
HaloLex database 33
H/D Exchange, data analysis 266–268
H/D exchange mass spectra, of peptic peptide 265
H/D exchange, mass spectrometry
  materials
    proteins 258
    solvents and chemicals 257
  methods
    ESI-MS using direct infusion 259–261
    protein fragmentation 262
Helicobacter pylori............ 67
High-field FTICR mass spectrometry 216
High performance liquid chromatography 183
  buffers 24
  fractionation, for MALDI analysis 388–389
High pH proteinase K digestion 5
Histones 216
Hourglass funnels
  effective ion injection into IMS using
    DC and AC power supplies for 430–431
    gas flow 431
  schemes of 427
HPLC. *See* High performance liquid
      chromatography
HPLC-MS, total ion chromatogram 197
Human plasma proteome, difficulties in
      detection of 309–310
Human serum, screening of 322
Hybrid mass spectrometers 132
Hydrofluoric acid (HF) 59
Hydrogen exchange labeling 256
Hydrophobic effects, in noncovalent binding 274

I

IEF. *See* Isoelectric focusing
Imaging mass spectrometry 295–296
Immobiline pH groups (IPG) 103
Immonium ions 223–224
Immunoaffinity chromatography,
      methods of 290
Immunodepletion. *See also* Cancer patients,
      plasma profiling in
  materials used in 311
  methods for 312
Immunoprecipitation, methods of 289
IMS. *See* Imaging mass spectrometry
IMS drift tubes
  electrodes in 430
  scheme and photo of 429
IMS/MS systems, components of 419
IMS/time-of-flight (ToF) MS systems 419
  ToF functionality 423
Infrared multiphoton dissociation
      (IRMPD) 216, 223–224, 389
In-gel proteolytic digestion, of
      deglycosylated proteins 187
Intact protein 233, 246
Intact Protein Analysis System 312
Intact proteins, CIEF/RPLC separations
  materials
    protein denaturation and alkylation 234
    yeast cell lysis 234
  methods
    base peak chromatograms of 237
    integrated platform 235
    online integration, schematic of 236
    protein sample preparation 235
    soluble fraction of 235
Integral membrane proteins (IMP) 41
Iodoacetamide 61
Ion-exchange high-performance liquid
      chromatography (IEX-HPLC) 115
Ion funnels
  capacitance of 428
  ion packets focused in 428
  regular and hourglass 427
Ion/ion reaction methods, in electrodynamic
      ion traps 404–409
Ion mobility spectrometry (IMS)
  axial ion focusing in 425–426
  in axially segmented quadrupole/ion funnel 426
  constraints on separation power of
    peak capacity 419–420
    resolving power 420
  definition of 417
  free thermal diffusion 425
  ion separation 418
  limitations of 420
Ions
  fragmentation of 133
  in uniform electric field 425
Ion trap
  instruments 207
  mass spectrometer 276
IPAS. *See* Intact Protein Analysis System
IRMPD. *See* Infrared multiphoton dissociation
Isoelectric focusing 65, 322
  buffer and solutions for 72–75
  cap gel 73

- cleaning of tubes 83
- gel chamber 69, 70
- preparation of gels 78–79
- running IEF 80
- sample application 79–80
- separation gel 73

Isoform Resolver Programs 36

Isotope coded protein label (ICPL) approach 22

- ICPL workflow 29, 31

Isotope labeling 24, 32

J

JGI Genome Portal 451, 452

K

Kernel, function selection in SVMs 458–459

k sets, of SVM scores 456

L

Label-free proteomic analysis of prokaryotic cells

- bicinchoninic acid assay 54–55
- cell lysis 49
- clustering software OminVis™, use of 47
- Coomassie blue assay 51
- FTICR and ThermoElectron LTQ-FT analysis 57
- FTICR data to PMT, matching 57
- internal reference proteome 51
- to normalize replicate runs 46
- pellet proteins, digestion protocol 52
- *p*-value/ANOVA analyses 49
- reduction and denaturation of proteins 51–52
- sample preparation 41, 46
- SCX clean-up 52–53
- soluble proteins
  - C-18 clean-up of 54
  - reduction and denaturation of 53
  - trypsin digestion and alkylation of 53–54
- strong cation exchange offline fractionation 55–56
- subcellular fractionation 49–50
- *z*-score transformation 47–48

Lab Lemco (LL) 373

LC-MALDI-TOF/TOF analysis 33

LC-MS-based proteomic data 448

LC-MS, sample digestion and preparation 32–33

LC-MS system 136

LDI imaging 296

Lectin affinity chromatography 330

Lectin-bound fraction, data processing 177

Lectin microcolumns, packing of 164

Linear ion trap mass spectrometer 348

Linear ion traps, role of 402–404

Liquid chromatography, phosphoproteomics

- analytical column 136
- bound peptides 137
- HPLC solvent 134
- standard acquisition method settings 138
- Tune acquisition software settings 137

Low-mass cutoff (LMCO) 404–405

Lowry assay kit 96

LTQ. *See* Linear ion trap mass spectrometer

LTQ-Orbitrap hybrid mass spectrometer, schematic of 132

LTQ-Orbitrap "XL" or "Classic" mass spectrometer 137

LTQ series mass spectrometer 10

LTQ Xcalibur™ software 10–11

Luria Berani (LB) 373

M

Maackia amurensis 325

mAb. *See* Monoclonal antibody

Macroporous silica 172

- activation of 164

Magtran software 264

MAL. *See* Maackia amurensis

MALDI. *See* Matrix-assisted laser desorption-ionization

MALDI analysis, HPLC fractionation for 388–389

MALDI-IMS. *See* Matrix-assisted laser desorption and ionization imaging mass spectrometry

MALDI-IMS images, of rat brain tissue 305

Maldi mass spectral data 376–377

MALDI mass spectrum 189

MALDI matrices, structures of 370

MALDI-Matrix for ICPL labeling 24

MALDI MS analysis, of invertebrate tissues .... 386–387. *See also* Neuropeptides, mass spectrometric characterization of

MALDI-MS, sample application 85

MALDI sample preparation 371

MALDI spore sample preparation 375–376

MALDI-TOF MS. *See* Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry

MALDI-TOF-MS/MS analysis 25

MALDI-TOF-TOF-MS analysis

- of spot 550 from *Helicobacter pylori I* 26695 cellular proteins 68

Maleimide reactive dyes 94

MASCOT search algorithms 29, 57, 147, 148, 158

Mascot software 86

Mass measurement accuracy (MMA) 40

Mass spectrometer (MS) 21, 245

- contamination of 210
- for coupling with CZE 207
- proteins detection 248
- settings for analysis 33

Mass spectrometric analysis, of cross-linked protein complex 284

Mass spectrometric characterization, of neuropeptides 381–382
  materials in 382–385
  methods in 386–391
Mass spectrometry-based imaging 295–296
Mass spectrometry-based proteomics experiment 131
Mass spectrometry (MS) 162, 215, 382
  in femtomols detection 310
  mass shifts hydrogen, exchange of 256
  phosphoproteomics
    analytical column 136
    bound peptides 137
    HPLC solvent 134
    standard acquisition method settings 138
    Tune acquisition software settings 137
  proteome-wide protein phosphorylation analysis by 143
Mass spectrometry protein identification. *See also* Cancer patients, plasma profiling in
  materials for 311
  methods for 313–314
Matrix-assisted laser desorption and ionization imaging mass spectrometry 296
Matrix-assisted laser desorption-ionization (MALDI) 66, 115, 162, 370
Matrix assisted laser desorption/ionization mass spectrometry (MALDI MS) 186
Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry 367–368
  in microorganisms characterization 367–371
    bacterial spore mass, estimation of 374
    bacterial spore preparation 373–374
    MALDI mass spectral data 376–377
    MALDI spore sample preparation 375–376
    MALDI vegetative cell preparation 374
    materials in 371–373
    matrix-to-analyte ratio 375
    vegetative bacterial sample preparation 373
Matrix-enhanced SIMS 296
Matrix-to-analyte ratio 375
MCP. *See* Microchannel plate
Measurement mass accuracy (MMA) 131
  high 132, 133
Membrane proteins 113
β-Mercaptoethanol (β-ME) 341
ME-SIMS. *See* Matrix-enhanced SIMS
Microchannel plate 300
Microheterogeneity, of glycosylation site 163
Microorganisms characterization, MALDI-TOF MS in, 367–371. *See also* Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry
Mitochondria, isolation 26
Mitochondrial extraction buffer 23
Molecular ions, generation methods of 295–296
Molecular weight cut-off tube 384
Monoclonal antibody 163
  base peak intensity chromatogram of tryptically digested 171
Monoisotopic masses, of amino acids 145
MS. *See* Mass spectrometry (MS)
MS/MS
  ion mobility analog of 438
  phosphopeptide enrichment 147–150
MS/MS Ions Search form 86
MS/MS spectra 144
  acquisition of 147
  for fragmentation of parent ion 988.51 (+2). ... 156–158
  identifying and annotating 147–150
  searching programs for 158
MudPIT analyses, of protein mixture 2
  data analysis 16–17
  microcapillary column construction 5–8
  multidimensional chromatography and tandem mass spectrometry 8
    chromatographic profile of mediator sample 9
    complex mixture analysis, gradient profiles for 11–15
    instrument method design description 10–11
    nanoelectrospray stage 8
  sample loading 5–8
  sample preparation 2–5
  spectral count normalization 17–18
MudPIT-based proteomics 22
Multidimensional protein identification technology (MudPIT) 2, 22
Multistage activation 133, 138, 141
MWCO. *See* Molecular weight cut-off tube

N

NaCNBH$_3$. *See* Sodium cyanoborohydride
Nanoelectrospray ionization source 385
NanoESI CID spectrum 189
NanoESI Q-TOF fragment ion spectrum 196
  nomenclature of 197
NanoESI Q-TOF mass spectrum 193, 194
Nano LC-MS/MS system 164
NanoLC-trypsin-MS analysis 249
NanoLC-trypsin-MS system 246
Nanoscale liquid chromatography (nano-LC) 162
National Institutes of Health (NIH) 178
NCBI BLAST 448
n-Dodecyl b-D-maltoside (b-DM) 114
nESI. *See* Nanoelectrospray ionization source
Neuropeptides
  extraction 387–388

  - mass spectrometric characterization of 381–382
    - materials in 382–385
    - methods in 386–391
- N-Glycoproteins
  - analysis, via in-gel procedures 186
  - materials used for
    - HPLC-MS analysis of 185
    - in-gel procedures for 184
    - in-solution procedures for 185
  - methods
    - HPLC-MS analysis of 195–197
    - in-capillary proteolytic digestion 190, 192–193
    - in-solution proteolytic digest 193–195
    - nanoelectrospray ionization
      - mass spectrometry 185–186
    - NanoESI Q-TOF mass spectrum 193
    - separate analysis of 186–190
    - size exclusion chromatography 191
    - structure elucidation of 186–190
    - ultrafiltration, 191–192
- N-glycosylated proteins, structure elucidation of 182
- NHS-ester dye 101
- *N*-(2-hydroxyethyl) piperazine-*N* ¢-(2-ethanesulfonic acid) (HEPES) 22
- *N*-Hydroxysuccinimide esters 283, 286
- $^{12}C_6$-and $^{13}C_6$-nicotinoylation 34
- Nonporous-reverse phase (NPS-RP)-HPLC 325
- Nonporous silica 333
- Normalized spectral abundance factors (NSAF) for proteins 13, 17
- NPS. *See* Nonporous silica
- Nuclear extraction buffer 23
- Nutrient sporulation medium, in microorganisms characterization, 372. *See also* Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry
- NWPolygraph (algorithm) 453–454, 458

O

- *O*-glycosylation, of proteins 183
- Oligosaccarides 182
- Online trypsin digestions 245
- Orbitrap mass spectrometer 276
- Organelle fractionation 25

P

- *Pantoea agglomerans* 373
- PBS. *See* Phosphate buffered saline (PBS)
- PDCH. *See* Perfluoro-1,3-dimethylcyclohexane
- Peakpicker software 33, 34
- Peanut agglutinin 325
- Peptic digestion
  - LC-ESI-MS 264
  - reproducibility of 264
- Peptic peptides
  - deuterium contents of 267
  - isotope exchange 268
  - time course of deuterium incorporation for 266
- Peptide fragment fingerprinting (PFF) 115
- Peptide mass fingerprinting (PMF) 115
- Peptide-*N*-glycosidase F (*PNGase* F) 183
- Peptide ratio, calculation of 34
- Peptides
  - identification after database search stage 454
  - mapping using tool 449
  - quality assessment 449
  - sequences, global proteomic analysis of 448
  - vectorization step 458
- Perfluoro-1,3-dimethylcyclohexane 401
- Perfluoro methyl decalin 401
- Permethylation, of oligosaccharides 344–345
- *Phanerochaete chrysosporium*
  - peptides
    - distribution of mapped 451
    - peptide length distribution for 450
- Phosphate buffered saline (PBS) 95, 284, 355
- Phosphopeptide enrichment 133
  - sample preparation 144–145
  - by selective sample purification 144
  - spectral acquisition
    - eluted peptides 147
    - protocols for 144
- Phosphoprotein analysis 322
  - antibodies in 325–326
  - ProQ Diamond in 326
- Phosphoproteins, enrichment of 133
- Phosphoproteomics workflow
  - challenges in using mass spectrometry for 144
  - in-solution protein digestion
    - materials for 133–134
    - procedure 135
  - liquid chromatography-mass spectrometry
    - analytical column 136
    - bound peptides 137
    - HPLC solvent 134
    - standard acquisition method settings 138
    - Tune acquisition software settings 137
  - strong cation exchange chromatography
    - materials for 134
    - procedure 135
  - TiO$_2$ chromatography
    - procedure 136
    - solutions 134
- Phosphorylation site identification
  - problems in 143–144
  - on specific residue 156
  - tyrosine phosphorylation 155

- Phosphospecific antibodies, usage of 322
- Photomultiplier tube 339
- PMD. *See* Perfluoro methyl decalin
- PMF MS spectra 85
- PMT. *See* Photomultiplier tube
- PNA. *See* Peanut agglutinin
- PNgase F glycans, release of 187
- Polypeptides structure determination, ECD for 413–414
  - materials for 414
  - methods of 414–415
- Posphopeptides, enrichment of. *See* Phosphopeptide enrichment
- Postsource decay 389
- Posttranslational modifications (PTMs) 215, 321
- Prognosis, in peptides and proteins ion reaction 409–410
- Protease inhibitors 116
- Protein characterization
  - by CZE-MS
    - mass spectrum 208–209
    - sample treatment for 204–205
    - separation of fetuin 206–207
    - separation procedure 205–206
- Protein conformers, characterization
  - using FAIMS/IMS
    - dynamics of conformational transitions 438, 440
    - highly charged cytochrome *c* ions 438
    - ubiquitin 436–437
- Protein denaturation 4
- Protein digestion and preparation, for MudPIT analysis 23–24
- Protein fractionation. *See also* Cancer patients, plasma profiling in
  - materials for 311
  - methods for 312–313
- Protein glycosylation 162
  - analysis, mass spectrometry in 326–327
- Protein glycosylation and phosphorylation, detection of 321–322
  - materials for 323–327
  - methods for
    - high abundance proteins removal, IgY antibody column in 328–329
    - lectin affinity glycoprotein extraction 330
    - peptide sequencing in protein identification 339–340
    - protein fractionation based on hydrophobicity 333–335
    - protein glycosylation analysis, mass spectrometry in 340–346
    - protein microarrays methods 335–339
    - protein phosphorylation, mass spectrometry in 346–348
    - sample preparation for 327–328
- Protein identification 243
  - by MASCOT Search 86–87
  - by peptide sequencing 326
- Protein inference problem 36
- Protein–ligand complex detection, ESI-MS in 274
- Protein microarrays, usages of 322, 335–339
- Protein noncovalent complex study ESI-MS, 273–274
  - materials for 274–275
  - procedures for 275–276
    - measurement of 277–279
    - sample preparation 276–277
    - solution binding constant measurement for 279–281
- Protein phosphorylation 133
  - analysis, mass spectrometry in 327
- *ProteinProphet* Programs 36
- Protein-protein interaction study, chemical cross-linking in 283–284
  - materials for 284–286
  - *in vitro* application methods 286–288
  - *in vivo* application methods 288–290
- Proteins/glycoproteins, analysis of 175
- Proteomics 21, 233, 241
- 4700 proteomics analyzer 72
- PSD. *See* Postsource decay
- *Pseudomonas putida* 373
- *Pseudomonas stutzeri* 373
- PTM. *See* Posttranslational modifications

Q

- Q-Star operating software 432
- Quadrupole-Fourier Transform Ion Cyclotron Resonance (Qq-FTICR) 224
- Quadrupole mass spectrometer 276
- Quadrupole time-of-flight (Q-TOF) 162, 185, 275, 385
  - components of 419
  - and ToF–ToF, comparison between 440
  - transit time of ions of 425
- Quadrupole type instruments 207
- Quantitative proteomics
  - based on SpC 29–30
  - by ICPL 31
- 2D Quant kit 95

R

- Rainin Dynamax HPLC system 385
- Random sampling effect 35
- Receiver operating characteristic (ROC) curve 456, 460
- Reversed-phase chromatography (RP) 2
  - in protein fractionation 313
- Reversed-phase high-performance liquid chromatography (RP-HPLC) 114
- Reversed-phase liquid chromatography (RPLC) 233

Reversed-phase protein separation 243
Rotofor, usages of 331
RPLC. *See* Reversed phase liquid chromatography (RPLC)
RXR, global exchange profile for.................................. 263

S

SA. *See* Sinapinic acid
Sambucus nigra..164, 325
SASPs. *See* Small acid-soluble proteins
Scanning mass spectrometer... 163
SCX solvent... 134
SCX wash buffer 43
SDS. *See* Sodium dodecyl sulfate
SDS–PAGE
    application of IEF gels and run of....................... 82–83
    buffer and solutions for....................................... 75–77
    chamber DESAPHOR VA 300 71
    gels, preparation of... 81–82
    incubation of IEF gels 81
    isolation, of cross-linked products 287
SDS–polyacrylamide gel electrophoresis 65
Secondary ion imaging mass spectrometry.................... 296
Selected-ion-count 304
SEQUEST 450
    algorithm 339, 340, 453–454
    percentile statistics of peptide matches from 12
4000 Series Explorer Software Version 3.6............... 85–86
Serine, phosphorylation of... 155
Serratia marcescens 373
Sialylated glycopeptides, MS spectrum of 172
SIC. *See* Selected-ion-count
Silver staining...71, 77–78, 84
SIMS. *See* Secondary ion imaging mass spectrometry
Sinapinic acid 371
Single proton transfer ion/ion reaction
        experiments....................................... 405–408
Size exclusion chromatography 191
Small acid-soluble proteins.. 371
SNA. *See Sambucus nigra*
SNA lectin, immobilization of............................. 164, 174
SOD. *See* Superoxide dismutase
Sodium cyanoborohydride... 355
Sodium dodecyl sulfate.. 325
Sodium dodecylsulfate polyacrylamide gel electrophoresis
        (SDSPAGE) 186
Sonication 60, 128
Spectral abundance factor (SAF).................................... 17
Spectral counting (SpC) 22
Spectral counts 17
Split triple-phase column 5
Split triple-phase fused-silica microcapillary column 7–8
Spraying solvents, in top-down experiments 224–225
STATQUEST algorithm
    Sequest searches and validation by 27–28
        elution profiles 28
        for LTQ data 28
        Swiss-Protein/TrEMBL and IPI databases,
            use of.. 27
Stem-loop binding protein (SLBP)...................... 221–222
Strong cation exchange chromatography,
        phosphoproteomics
    materials for.. 134
    procedure 135
Strong cation exchange chromatography (SCX) 2
Subtilisin A digestion 5
Sucrose cushion solution 23
Superoxide dismutase 251
SVM peptide scoring model
    code for generating and validating.................. 454–455
    primary steps in developing
        cross-validation approach 456, 457
        peptide data vectorization and normalization 455
        predict file 456, 458
SVMs
    classification threshold................................... 459–460
    classifier
        approach to develop and validate 457
        predict file 456, 458
    selection of kernel function in......................... 458–459
    supervised learning approach 454
    training code.. 456

T

Tandem mass spectra... 85
Tandem mass spectrometer 2
Tandem mass spectrometry, in glycan
        structure 345–346
TBS-T. *See* Tris-buffered saline with Tween
TCA precipitation 5, 18
TCEP. *See* Tris[2-carboxyethyl] phosphine
Testing files 456
TFA. *See* Trifluoroacetic acid (TFA)
ThermoFinnigan Orbitrap Ms 46
ThermoFinnigan LTQFT.. 46
Threonine, phosphorylation of 155
Thylakoid membrane proteins, characterization............ 114
    chlorophyll concentration, estimation..................... 118
    digitonin, for fractionation of PSII and PSI 118–119
    grana and stroma protein fractions,
        analysis of... 124–126
    IP-RP-HPLC–ESI-MS, setup for................. 120–122
    materials for... 115–117
    retention times, molecular masses, and
        identification of proteins............................. 126
    sample preparation and separation of PSI
        and PSII... 119
    separation of PSII from PSI membrane proteins..... 120
    spinach plants, hydroponic culture of 117
    thylakoid membranes, isolation of.................. 117–118
    total thylakoid protein extract, analysis of........ 122–124

TIC. *See* Total-ion-count
Time-of-flight 296
Tissue analysis, IMS in 295–296
  materials used in 297–299
  methods 299–305
Tissue homogenization buffer 23
Titanium oxide chromatography, phosphoproteomics
  procedure 136
  solutions 134
ToF. *See* Time-of-flight
TOFTOF machines 66
Top-down proteomics experiments 216
  collision-induced dissociation 228
  cytidine deaminase isoelectric focusing, by IEF-FFE 220
  data analysis 229
  electron capture dissociation 228
  by high-field FTICR-MS 221–224
  IEF-FFE fractions, C_4 reversed phase purification of 221
  infrared multiphoton dissociation 228–229
  materials for 217–218
  MS experiments 227–228
  protein separation, by FFE 219–220
  proteins preparation, for direct infusion into FTICR 221
  source mode 225–226
  spraying solvents for 224–225
Total-ion-count 296
  of rat brain tissue 297
Transthyretin 353
  interpreting and reporting results 361–362
  materials required in 354–359
  MS quality control 360
  patient samples in 360–361
  sample processing 359
  valves and MS configuration 359–360
Trifluoroacetic acid (TFA) 23, 298, 325, 371, 384
2′, 4′, 6′,–Trihydroxyacetophenone monohydrate 327
Triple-phase fused-silica microcapillary column 6–7
Tris-buffered saline with Tween 285
Tris[2-carboxyethyl] phosphine 323
Trypsin 41
Trypsin digestion 43
Tryptic digestion 72, 78
Tryptic digest, SCX chromatogram of 135
Tryptic in-gel digestion 84
TTR. *See* Transthyretin
Tune acquisition software settings 137
Two-dimensional IMS, effective peak capacity of 420–421
Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) 93
Tyrosine phosphorylation 155

U

Ubiquitylation detection, using LC-trypsin-MS 251
UltraTrol 210
Urea 58, 59

V

Varian OMIX cartridges 27
Varitemp heat gun 18

W

Water-insoluble cross-linkers, usages of 286
Wheat Germ Agglutinin (WGA) 324

X

Xcalibur Instrument Setup 137
Xcalibur software 13
Xcorr 450–451

Y

Yersinia pestis 40
Yersinia proteome analysis 41
Y-Tube, in atmospheric pressure ion/ion reactions/ sampling, ... 397–398. *See also* Electrospray ionization (ESI)

Z

ZipTip cleaning 85