Dynamic Combinatorial Chemistry

Joost N. H. Reek
and Sijbren Otto, Editors

WILEY-VCH Verlag GmbH & Co.

Edited by Joost N. H. Reek and Sijbren Otto

Dynamic Combinatorial Chemistry

WILEY-VCH Verlag GmbH & Co. KGaA

Edited by
Joost N. H. Reek and
Sijbren Otto

Dynamic Combinatorial Chemistry

Further Reading

Diederich, F., Stang, P. J., Tykwinski, R. R. (eds.)

Modern Supramolecular Chemistry

Strategies for Macrocycle Synthesis

2008
ISBN: 978-3-527-31826-1

Van Leeuwen, P. W. N. M. (ed.)

Supramolecular Catalysis

Strategies for Macrocycle Synthesis

2008
ISBN: 978-3-527-32191-9

Schalley, C. A. (ed.)

Analytical Methods in Supramolecular Chemistry

2007
ISBN: 978-3-527-31505-5

Vögtle, F., Richardt, G., Werner, N.

Dendrimer Chemistry

Concepts, Syntheses, Properties, Applications

2009
ISBN: 978-3-527-32066-0

Sliwa, W., Kozlowski, C.

Calixarenes and Resorcinarenes

Synthesis, Properties and Applications

2009
ISBN: 978-3-527-32263-3

Hecht, S., Huc, I. (eds.)

Foldamers

Structure, Properties, and Applications

2007
ISBN: 978-3-527-31563-5

Schneider, G., Baringhaus, K.-H.

Molecular Design

Concepts and Applications

2008
ISBN: 978-3-527-31432-4

Höltje, H. D., Sippl, W., Rognan, D., Folkers, G.

Molecular Modeling

Basic Principles and Applications

2008
ISBN: 978-3-527-31568-0

Bannwarth, W., Hinzen, B. (eds.)

Combinatorial Chemistry

From Theory to Application

2006
ISBN: 978-3-527-30693-0

Edited by Joost N. H. Reek and Sijbren Otto

Dynamic Combinatorial Chemistry

WILEY-VCH Verlag GmbH & Co. KGaA

The Editors

Prof. Dr. Joost N. H. Reek
Van t Hoff Institute for
Molecular Sciences–HIMS
Nieuwe Achtergracht 166
1018 WV Amsterdam
The Netherlands

Dr. Sijbren Otto
University of Groningen
Centre for Systems Chemistry
Stratingh Institute
Nijenborgh 4
9747 AG Groningen
The Netherlands

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Typesetting Toppan Best-set Premedia Limited
Printing and Binding Strauss GmbH, Mörlenbach
Cover Design Grafik-Design Schulz, Fußgönheim

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN: 978-3-527-32122-3

Contents

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

List of Contributors

Luis Amorim
KTH – Royal Institute of Technology
Department of Chemistry
Teknikringen 30
10044 Stockholm
Sweden

Sophie R. Beeren
University of Cambridge
Department of Chemistry
Lensfield Road
Cambridge CB2 1EW
UK

Nicole E. Botterhuis
Eindhoven University of Technology
Chemical Engineering and Chemistry
Supramolecular Polymer Chemistry
PO Box 513, Helix STO 4.43
5600 MB Eindhoven
The Netherlands

Pierre-Alain R. Breuil
University of Amsterdam
Van't Hoff Institute for Molecular Sciences
Nieuwe Achtergracht 166
1018 WV Amsterdam
The Netherlands

Rémi Caraballo
KTH – Royal Institute of Technology
Department of Chemistry
Teknikringen 30
10044 Stockholm
Sweden

Andrea M. Escalante
Universidad Nacional de Rosario – Consejo Nacional de Investigaciones Científicas y Técnicas
Facultad de Ciencias Bioquímicas y Farmacéuticas
Siupacha 531
2000 Rosario
Argentina

Ricardo L.E. Furlan
Universidad Nacional de Rosario – Consejo Nacional de Investigaciones Científicas y Técnicas
Facultad de Ciencias Bioquímicas y Farmacéuticas
Siupacha 531
2000 Rosario
Argentina

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

Oscar Norberg
KTH – Royal Institute of Technology
Department of Chemistry
Teknikringen 30
10044 Stockholm
Sweden

Sijbren Otto
University of Groningen
Centre for Systems Chemistry
Stratingh Institute
Nijenborgh 4
9747 AG Groningen
The Netherlands

Jeremy K.M. Sanders
University of Cambridge
Department of Chemistry
Lensfield Road
Cambridge CB2 1EW
UK

Kay Severin
École Polytechnique Fédérale de Lausanne
Institut des Sciences et Ingénierie Chimiques
1015 Lausanne
Switzerland

Rint P. Sijbesma
Eindhoven University of Technology
Chemical Engineering and Chemistry
Supramolecular Polymer Chemistry
PO Box 513, Helix STO 4.44
5600 MB Eindhoven
The Netherlands

Olof Ramström
KTH – Royal Institute of Technology
Department of Chemistry
Teknikringen 30
10044 Stockholm
Sweden

Joost N.H. Reek
University of Amsterdam
Van't Hoff Institute for Molecular Sciences
Nieuwe Achtergracht 166
1018 WV Amsterdam
The Netherlands

Jean-Luc Wietor
Eindhoven University of Technology
Chemical Engineering and Chemistry
Supramolecular Polymer Chemistry
PO Box 513
5600 MB Eindhoven
The Netherlands

1
History and Principles of Dynamic Combinatorial Chemistry

Sophie R. Beeren and Jeremy K.M. Sanders

1.1
Introduction

Dynamic combinatorial chemistry (DCC) [1]—combinatorial chemistry under thermodynamic control—is a tool for the efficient synthesis of libraries of complex structures whose individual properties may be explored through the library's response to the stabilizing influences of external stimuli. A dynamic combinatorial library (DCL) is generated by combining building blocks, functionalized such that they can react with one another either through reversible covalent reactions or specific noncovalent interactions, to form a mixture of interconverting library members. As the exchange of building blocks between library members takes place, the product distribution moves towards equilibrium—the thermodynamic minimum of the system.

The concentrations of the different species in a library will be dependent upon the intrinsic stability of the various library members. However, the library composition is not fixed and the introduction of any external stimulus that can alter the relative stability of a library member will influence the product distribution. In particular, stabilization of a particular library member through noncovalent interactions with an added template will alter the positions of the equilibria governing the system (Figure 1.1). Such a shift in equilibrium position will ideally lead to an increased production—an "overexpression" or "amplification"—of the stabilized library member at the expense of the other species in the mixture. In this way, a library may be probed for species with affinity for a given target molecule.

While amplification of a library member in the presence of a template is usually a good indication of a favorable interaction between that species and the template, one has always to be alert to the possibility of counterintuitive "systems" behavior by the library [2]. The equilibrium distribution is determined by the sum of the total thermodynamic stabilities of all species in the library. Since all the components of a library are linked through a set of equilibria, the stabilization of one library member will be felt by the others and in some circumstances the most amplified species may not be the one that in isolation binds most strongly to the

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

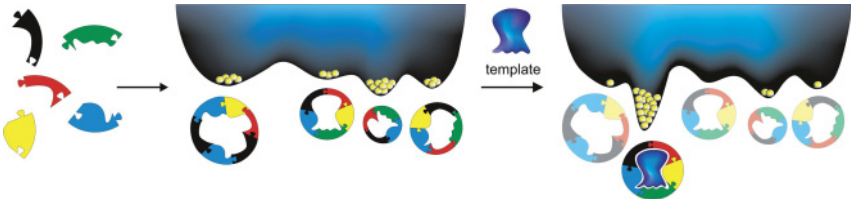


Figure 1.1 A DCL is formed from the different combinations of several building blocks. Ideally, addition to the library of a template will alter the library distribution to amplify the receptor that forms the most stable complex with the template.

target. With careful design of the experiment, however, such misleading amplifications may be minimized [2] (see Chapter 2).

Identification and synthesis of receptors for small molecules has been one of the most popular early uses for DCC (Chapter 3). However, its application now extends beyond this. DCC has been used to generate effective ligands for biomacromolecules (Chapter 5), to identify foldamers stabilized by self-recognition through intramolecular noncovalent interactions (Chapter 6), and to find stable aggregates formed due to interactions between combinations of library members [3]. It has found application in the synthesis of catalysts (Chapter 4), sensors (Chapter 7), and dynamic materials (Chapter 6). Furthermore, stabilization of components in a library need not necessarily be through interactions with a chemical template. Variation of environmental conditions, such as temperature, pH [4], light [5], and electric fields [6], or removal of solvent to induce a phase change [7–9], can be used to influence library distribution and to probe the properties of the species formed in the library.

In conventional synthesis, chemists typically aim to prepare a single target species in each reaction. Reversible reactions are frequently avoided to ensure that the intended products, once synthesized, do not revert to the starting materials or convert to different products. The formation of multiple products is viewed as undesirable, given that the target compound would inevitably be produced in lower yield and would require separation from the various side products. The formation of unexpected products is often viewed not as fortuitous, but indicative of a lack of understanding or foresight in the design of the synthesis.

In conventional combinatorial chemistry, large libraries of related compounds can be generated quickly, and then efficiently screened and tested for desirable properties [10]. In theory, if not always in practice, the larger the library generated, the higher the probability of discovering a useful compound from among the many products. DCC embraces diversity and complexity as an efficient means to discover new molecules or supramolecular assemblies with unanticipated recognition properties. While many chemists would consider a mixture of multiple constantly interconverting compounds an overwhelming mess, in DCC these "libraries" are rather seen as a complex system of compounds whose potential properties and uses are awaiting discovery. There is the prospect of discovering not only new

receptors, catalysts, inhibitors, sensors, or materials, but also hitherto unsuspected (or unexplored) interactions and systems behaviors.

1.2 History

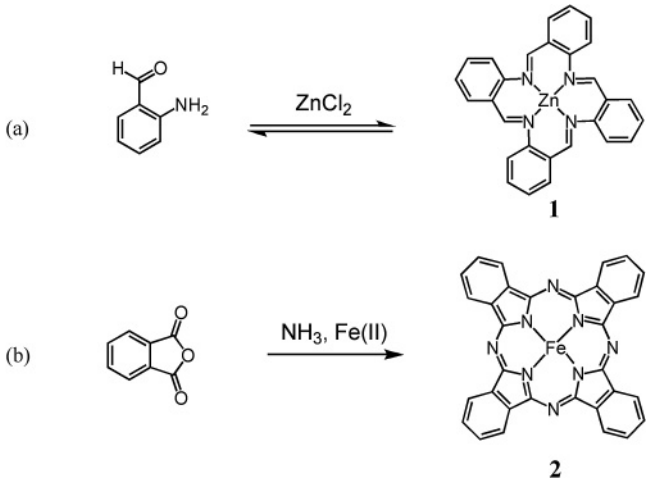
The basic principles of DCC, developed in the 1990s, stemmed from the realization that the task of constructing effective receptors capable of specific molecular recognition – a task that biological systems have accomplished only over millions of years of evolution – is very difficult to address using a straightforward synthetic design approach. A new, more efficient, and general approach was devised that captured the combinatorial, selection, and amplification elements exhibited by the mammalian immune system. The guest molecule would be allowed to select its own, most effective host from among a mixture of possible hosts. Selection from an equilibrium mixture of potential hosts would lead to a shift in the equilibrium position and amplification of the "best binder". Another attraction of this approach would be that, in principle, it minimizes synthetic effort, in that a small number of building blocks can lead to a wide range of large, complex products.

Although this concept was only recently articulated as a general approach to the synthesis of chemical species exhibiting molecular recognition capabilities, many of the principles and practices that characterize DCC had been in place for several decades. DCC may be viewed as the intersection of two pre-existing approaches to synthesis: thermodynamically controlled templated synthesis and combinatorial chemistry.

1.2.1 Thermodynamically Controlled Templated Synthesis

Thermodynamic control in organic synthesis is very familiar, for example, in the synthesis of esters, acetals, or imines. Specific product formation can be favored by shifting the equilibrium position, through the removal or addition of water, the choice of solvent, the use of excess reagents, or by controlling the temperature and pressure of the reaction. Templates may be used in thermodynamically controlled synthesis to guide the system towards the production of a desired product that "fits" the template.

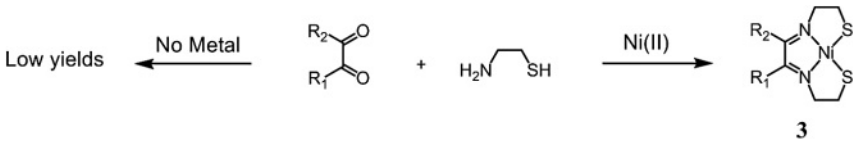
This approach can arguably be traced back to the nineteenth century, and the studies of Emil Fischer on carbohydrates and of Werner on coordination complexes. When Watson and Crick discovered the DNA double helix in 1953 they realized immediately that its replication involved a templated synthesis [11]. In retrospect, it is clear that the metal-ion-templated synthesis had been achieved as early as 1926 when Seidel reacted 2-aminobenzaldehyde with $ZnCl_2$ [12]. An imine-based macrocycle bound to zinc was formed, but was not identified until much later (Scheme 1.1a) [13]. A few years later, Fe(II) phthalocyanine was formed



Scheme 1.1 (a) Imine-based macrocycle synthesis templated by $ZnCl_2$ [13]. (b) Fe(II)-templated synthesis of a phthalocyanine [14].

through an unexpected iron-templated reaction between phthalic anhydride and ammonia (Scheme 1.1b) [14].

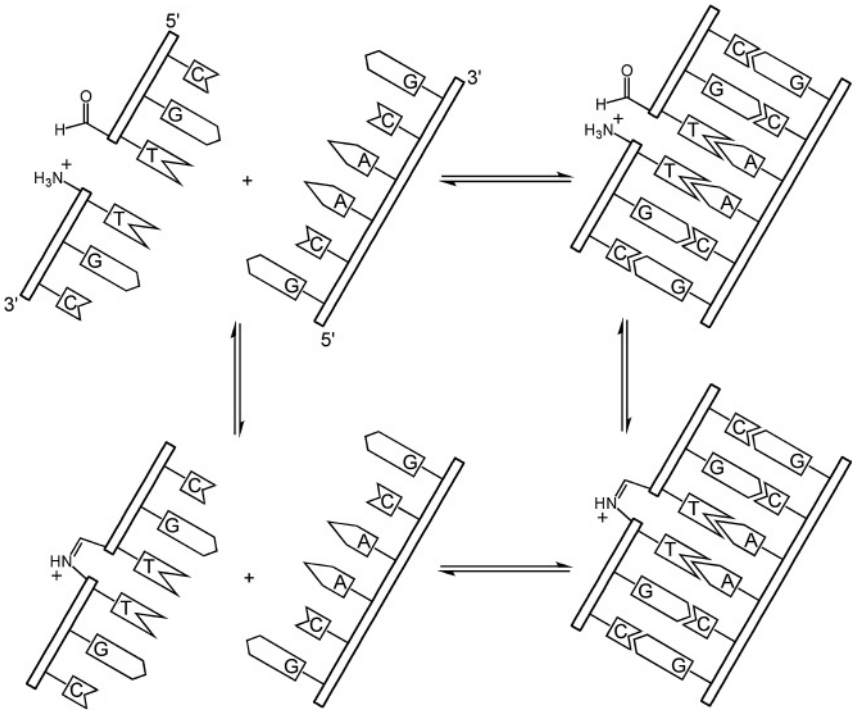
It is through the pioneering work of Busch in the 1960s that the role of templates in synthesis, both kinetically and thermodynamically controlled, was clarified and developed [15]. Busch wrote that "a chemical template organizes an assembly of atoms with respect to one or more geometric loci, in order to achieve a particular linking of atoms" [16]. Kinetic templates operate on irreversible reactions by stabilizing the transition states leading to the desired product [17]. In thermodynamic templating, the template is added to a reaction mixture under equilibrium conditions where it binds to a particular desired product and then shifts the equilibrium to favor the formation of that species. Busch's Ni(II)-templated bis-imine macrocycle **3** synthesis is probably the first published example to clearly articulate the role of a template in stabilizing a desired product from a complex equilibrating mixture (Scheme 1.2) [18].



Scheme 1.2 Synthesis of a bis-imine macrocycle **3** directed by Ni(II) templation [18].

Subsequently, thermodynamic templated syntheses developed along parallel and largely noncommunicating organic, inorganic, and biochemical tracks. A notable contribution by Goodwin and Lynn features reversible imine-mediated

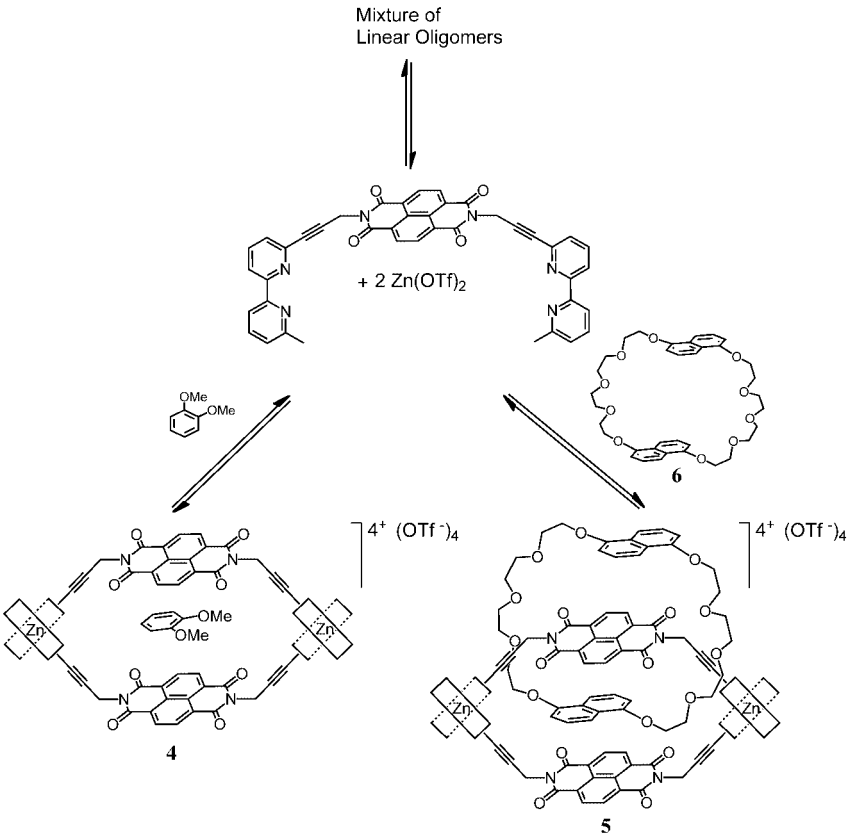
synthesis on a DNA template (Scheme 1.3) [19]. Prior attempts to exploit template-directed synthesis in polymerization reactions sought kinetic differentiation of product distribution. DNA and RNA polymerases employ reaction reversibility and sophisticated proofreading mechanisms to ensure high fidelity in template translations. In Goodwin and Lynn's approach the role of the template is to shift the equilibrium position. By controlling this equilibrium, they both demonstrated unique chemistry on a DNA template, and achieved the first chain-length and sequence-specific template-directed polymerizations.



Scheme 1.3 Thermodynamic cycle showing Goodwin and Lynn's reversible imine-mediated synthesis on a DNA template [19].

1.2.2
Early DCLs

In 1995, Hamilton *et al*. described one of the first examples of a combinatorial approach to synthetic receptors using reversible coordination around a metal ion [20]. In the same year Harding *et al*. described the guest-induced amplification of a metallo-macrocycle **4** and metallo-[2]catenane **5** from a mixture (Scheme 1.4) [20, 21]. Although not described as such, these examples employed the basic principles of DCC – template-stabilized selection of effective synthetic receptors from among an equilibrating library of potential receptors.



Scheme 1.4 An electron-rich dimethoxybenzene template was used to direct the synthesis of a naphthalenediimide-based macrocycle **4** [21]. Addition of dinaphtho-crown ether **6** led to the formation of [2]catenane **5** [22].

The first papers clearly articulating the concept of DCC appeared in the mid-1990s, the idea being conceived and developed independently in both the Sanders and Lehn groups. In the Sanders group, early experiments employed base-catalyzed transesterification to generate libraries of macrocycles formed from steroid-based building blocks [23]. A proof-of-principle study demonstrated modest amplifications of specific macrocycles in the presence of alkali metal ions [24].

Lehn developed the dynamic combinatorial approach as a result of his work on metal helicates, observing that the major product in a dynamic mixture of helicates was determined by the nature of the counterion that binds in the center of the helicates [25]. Huc and Lehn then extended their work to include templating of a ligand by a protein, describing the inhibition of carbonic anhydrase by a library of imines created *in situ* [26]. This work was preceded by a publication by Venton *et al.* who used nonspecific proteases to prepare and degrade a set of peptides

reversibly, with a view to amplifying the sequences that bound most strongly to fibrinogen [27]. In 1997, Miller *et al*. described the first dynamic combinatorial approach to DNA-binding compounds [28]. In the same year, Sasaki *et al*. published an elegant "self-adjusting" metal-centered ligand for lectins [29], and Eliseev and Nelen used light-induced alkene isomerization as a reaction to drive chemical evolution in an equilibrating mixture of simple arginine receptors [30]. The first disulfide exchange work to articulate a version of the dynamic combinatorial idea was probably that of Hioki and Still in 1998 [31], although the reversibility of thiol–disulfide exchange had been known and exploited for many years.

1.3
Exercising Control over a DCL to Influence Species Distribution

1.3.1
Selection through Molecular Recognition of an External Template

One of the most appealing applications of DCC is the possibility to determine from a complex mixture the "best binder" of a given target molecule. Upon addition of the template molecule to the library, the means of identifying an efficient binder is through its overexpression in the library. Therein the advantage of the approach is clear – not only is the desired molecule identified, but it is synthesized in preference to library members with a lesser affinity for the target molecule. In a particularly effective library, the addition of a template to the mixture may induce all the library material to convert to just one favored species.

Since the first examples of DCC appeared in the literature, external templating has been by far the most extensively developed means of directing a library's distribution. The template may be either the guest or the host molecule. In the first instance (see Chapter 3), building blocks may combine to form a receptor, often a macrocycle, that may be stabilized by binding to a small guest molecule (Figure 1.2a). Alternatively, the building blocks may come together to form a species that binds within a cavity or a ligand that stabilizes a macromolecule (Figure 1.2b). A detailed discussion of ligands for biomolecules identified using DCC can be found in Chapter 5.

1.3.2
Selection through Self-Templating

Internal templating refers to the self-selection of library members through intramolecular or intermolecular stabilizing noncovalent interactions. Intramolecular self-templating is observed when the species formed in a DCL are capable of folding upon themselves (Figure 1.2c). The library members that are best able to form favorable noncovalent interactions within themselves will be amplified in the library. DCC has therefore been used to study the folding of peptides, nucleotides, and synthetic polymers (Chapter 6). It can be used to direct the reversible

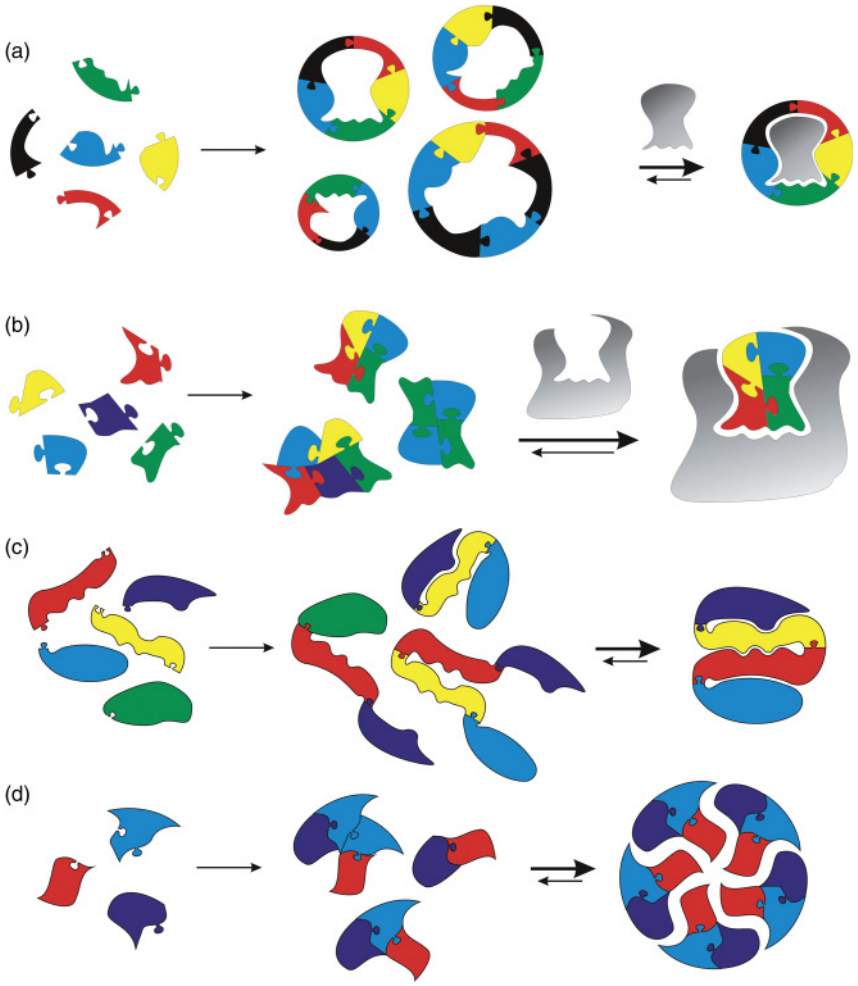
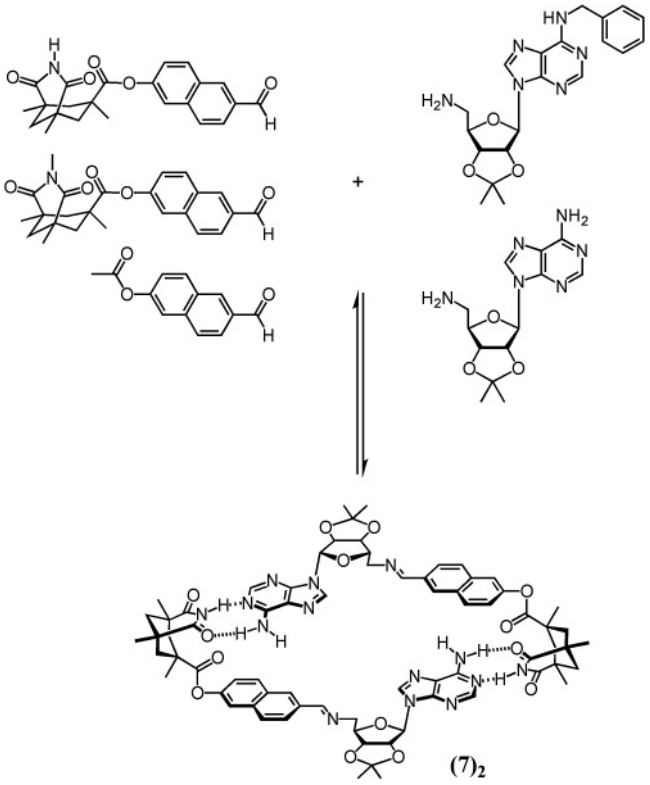


Figure 1.2 The different ways in which molecular recognition can exert control over library distribution: selection may be through external templating by an added (a) guest or (b) host molecule, or though self-templating that is either (c) intramolecular or (d) intermolecular.

synthesis of foldamers by amplifying the formation of the most stable from among a library of oligomers [32].

Selection and amplification of a molecule in a library is also possible where it is stabilized through intermolecular noncovalent interactions with one or more of the same or different library members (Figure 1.2d). Such interactions would necessarily have to be strong enough to outweigh the significant entropy cost associated with such an aggregation process. The first example of this type of system was recently published [3] by Xu and Giuseppone – a library was formed



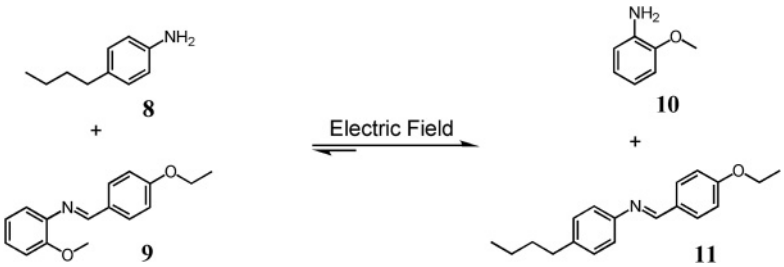
Scheme 1.5 The combination of these three aldehydes and two amines leads to the formation of a library of six imines. Imine **7**, the dominant species formed, is selected through intermolecular self-templation [3].

from the condensation of Kemp's imide-based aldehydes and adenosine-based amines (Scheme 1.5). Imine **7** was amplified from this mixture because it was capable of forming stable dimers through a complementary hydrogen bonding motif.

1.3.3
Selection Directed by External Physical Stimuli

Dynamic materials are being developed that can undergo constituent exchange, reorganization, and selection in response to external physical stimuli. In a DCL where the distribution changes upon variation of external conditions, amplification of a particular species within the library is indicative of its greater stability, relative to the other library members, when acted upon by an external physical stimulus.

Giuseppone and Lehn have recently investigated how the distribution in a DCL of imines can be directed by variation of the temperature and pH [4]. They have also shown that electric field modulation can be used to dictate component evolution in a DCL containing liquid crystals [6]. A simple library was formed from the interconversion of two imines (**9** and **11**) and two amines (**8** and **10**) (Scheme 1.6). When exposed to an electric field, imine **11**, which exhibits liquid crystalline behavior, and which coupled most to and was thus best stabilized by the electric field, was amplified. Ingerman and Waters recently reported a DCL in which irradiation with appropriate wavelengths of light allowed photochemical control over library distribution [5]. An azobenzene building block was incorporated into hydrazone-based libraries with the goal of developing macrocyclic hosts whose binding properties could be modulated by irradiation.



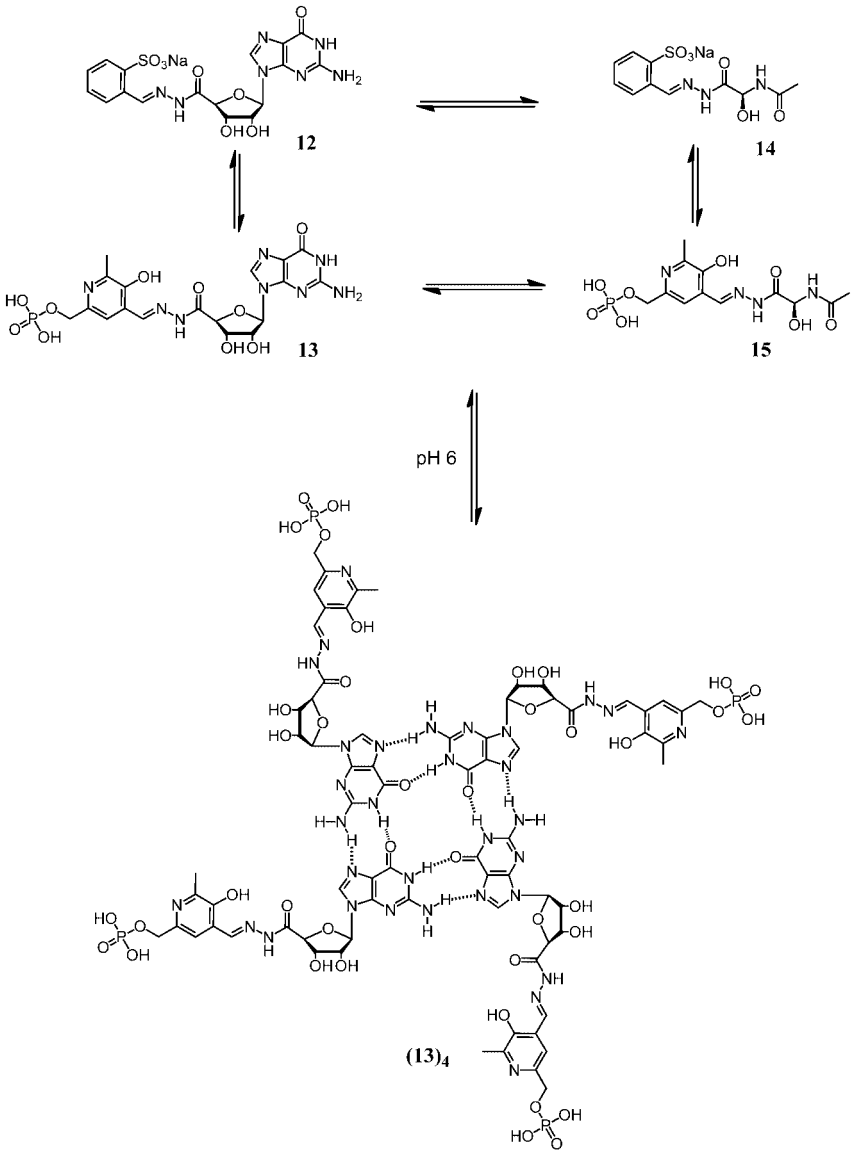
Scheme 1.6 In the presence of an electric field, the equilibrium in this small DCL shifts to increase the production of the liquid crystal **11** [6].

1.3.4
Selection Through a Stabilizing Phase Change

Component evolution in a DCL may be directed by the self-selection of library members that lead to the formation of the most highly organized and stable phase. This concept has been applied to both gelation and crystallization.

Sreenivasachary and Lehn have described a hydrazone DCL in which selection is driven by the formation of a stable gel (Scheme 1.7) [7]. Hydrazide-functionalized guanosine and serine building blocks were reacted with two different aldehyde building blocks in sodium acetate buffer at pH 6 to form a library of four interconverting hydrazones (**12–15**). Hydrazone **13**, which is capable of forming a stable gel based on the G-quartet motif, was amplified in the mixture; here, the distribution of components in the library was determined not by the relative stabilities of the individual hydrazones, but rather by the formation of the insoluble fibers that create a stable gel.

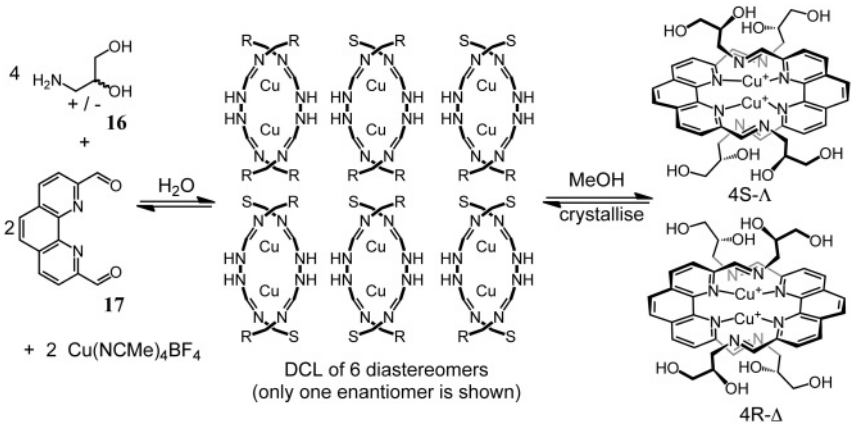
Dynamic polymers ("dynamers"), formed by linking monomers through reversible reactions, have emerged as a means to generate adaptive materials [33]. It has been possible to direct the distribution of different dynamers from within a DCL



Scheme 1.7 In this DCL, the formation of a stable gel gives rise to the selection and amplification of hydrazone **13** [7].

by altering the environmental conditions to stabilize or destabilize the mesoscopic states of the dynamers. Removal of solvent leads to preferential formation of the dynamer that gives the most ordered and stable crystalline phase [8].

In a related example, sorting of a DCL during crystallization has led to the selection of a pair of enantiomers from a complex library of diastereomers



Scheme 1.8 Selection via crystallization of a single pair of enantiomeric Cu(I) helicates from among a complex library of diastereomers [9].

(Scheme 1.8) [9]. A library of six pairs of enantiomers of different diastereomers of metal helicates was generated from the reaction of a racemic mixture of amine **16** with dialdehyde **17** in the presence of $Cu(NCMe)_4BF_4$. Imine exchange and metal–ligand exchange ensured that thermodynamic control of the library was maintained. During crystallization, a single pair of enantiomers was selected as a result of its lower solubility. The consequent continuous removal of this racemate from the solution drives the exchange of all other helicates in the solution toward this one "selected" product.

1.4 Designing a Dynamic Combinatorial System

In DCC, an element of design is sacrificed in order to efficiently obtain a diverse range of products, thus allowing for the possibility of generating unexpected molecular structures with unanticipated properties. However, this does not mean that building blocks are chosen at random. There is a broad spectrum of approaches to the synthesis of supramolecular entities spanning from very carefully designed templated thermodynamic synthesis, such as the construction of three interlocking molecular Borromean rings published by Stoddart *et al.* [34], to the mere mixing of arbitrarily chosen building blocks. DCC lies somewhere in the center of this continuum.

To set up a DCL, a reversible chemistry must be chosen and then suitable building blocks selected or synthesized. To generate a useful DCL the design of the building blocks may be very important. Although an increase in the proportion of the building blocks that have been specially designed may make the outcome of equilibration more predictable, it may also increase the probability that one or more of the library members formed will respond to external influences upon the library.

1.4.1 Building Block Design

The design of building blocks will necessarily require the incorporation of suitable functional groups at one or multiple positions that are capable of reacting reversibly when combined with other building blocks. The remainder of the molecule may then be designed to fit the purpose for which the library is intended, making sure that it does not contain functional groups that will interfere with the exchange reaction.

One of the most common design features in building blocks is the inclusion of functionalities likely to aid in molecular recognition. For building blocks designed to be used in libraries in organic solvents, the incorporation of hydrogen bond donors and acceptors could be important. In aqueous solution, motifs allowing for the possibility of forming hydrophobic pockets when combined with other building blocks might be useful. Charged building blocks may be chosen for recognition of anions or cations. Electron-deficient and electron-rich aromatics may be included where donor–acceptor interactions with an added template or between building blocks are anticipated. Where other external stimuli, such as light, temperature, and pH, are intended to direct the evolution of the library, building blocks may need to be designed such that the species they form under equilibrium conditions will respond to such stimuli.

The overall shape of the building blocks may also be chosen to suit a particular purpose. Where bifunctional building blocks capable of forming oligomers are reacted together in DCLs, it may be desirable to design them either with curved or linear structures, to encourage macrocyclization or polymerization, respectively. In the design of DCLs for the generation of receptors, the relative rigidity or flexibility of the building blocks should be considered. If the building blocks are excessively rigid, then the range of macrocycles that can form may be limited. One might expect to form a library with a narrow distribution, dominated by homo-species resulting from the self-sorting of the building blocks [35]. Rigidity may be useful in designed thermodynamic synthesis, where a single product is sought, but the same does not apply to DCC, where the aim is to form a diverse mixture of products. On the other hand, excessive flexibility is equally undesirable – too great a flexibility may allow the building blocks to collapse on themselves and form stable cyclic monomers rather than a range of higher oligomers. It appears that a balance, or a combination of rigidity and flexibility, is required to produce libraries with a broad distribution of oligomers and maximum chance of diversity.

Other considerations in the design of building block scaffolds might include the incorporation of solubilizing groups, chromophores, or other reporter groups. Analyzing the composition of a library is frequently the most challenging aspect of DCC, and therefore all building blocks should ideally have unique masses and/or spectroscopic signatures. In general, building blocks should be simple, easy, and inexpensive to synthesize, and straightforward to analyze. The idea is that

DCC should be an efficient approach to the generation of sophisticated molecular and supramolecular systems – diversity and complexity are obtained not through complicated synthesis, but through different combinations and permutation of simple building blocks linked by reversible reactions.

1.4.2 Exchange Reactions

The reactions responsible for the formation of DCLs from simple building blocks must necessarily be reversible, allowing for the exchange of building blocks between different library members. To establish a useful and efficient DCL the reaction chosen should meet several criteria.

First, the reaction must be reversible on a reasonable timescale, which implies that the forward and reverse reactions should ideally be fast. It is preferable that a large number of library members be relatively isoenergetic to avoid library mixtures in which there is a strong bias towards certain species, and a high energy cost and long equilibration time required to alter the library distribution to favor a different stabilized species. Where a library is dominated by one cyclic species, it is often a good indication that the intramolecular ring-closing reaction is exceptionally fast relative to any intermolecular reactions. Exchange, therefore, will be slow and the time required to reach equilibrium will be very long, even though the reversible reaction might be intrinsically fast.

In a true DCL, equilibration and selection need to take place simultaneously so the reaction conditions required for reversibility must also be compatible with the noncovalent interactions employed in the selection process. If basic conditions, for example, are required for exchange to take place, then certain recognition groups may be deprotonated and the building blocks negatively charged under library conditions. The reaction conditions should ideally be mild, so as not to disturb the delicate noncovalent interactions involved in molecular recognition.

Exchange reactions used in DCLs include reversible covalent reactions, metal–ligand coordination, and noncovalent interactions (in particular, hydrogen bonds). Of these three exchangeable linkage types, reversible covalent reactions have been by far the most extensively used in DCC. While the typically weaker, and more labile, hydrogen and coordinative bonds allow for fast exchange and short equilibration times, the supramolecular structures formed are inherently less stable, and more difficult to analyze and isolate.

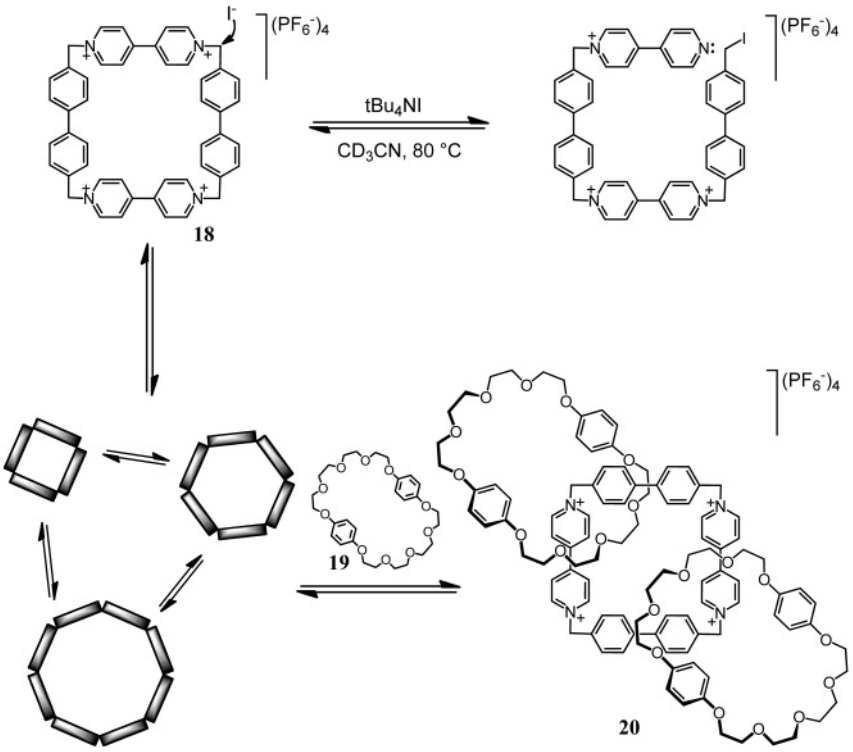
Reversible covalent reactions, although slower, give rise to more robust products. Often requiring a catalyst to ensure reversibility, such reactions are likely to be significantly slowed by removing the catalyst so that exchange may be effectively "switched off," allowing for isolation of selected library members without the complication of further exchange. However, the approach towards equilibrium requires an extremely large number of turnovers. This reduces the utility of reactions where the catalyst has a limited lifetime (e.g., alkene and alkyne metathesis) or where side-reactions occur [36].

1.4.3
Exchange Reactions Currently in Use

The exchange reactions used to date are listed below in Figure 1.3. A comprehensive review of exchange reactions used in DCC was published only recently [1(a)]. Therein the particulars of the reaction conditions required are discussed and relevant examples are provided. An overview of the most commonly used reversible chemistries will be provided in the first part of Chapter 3, discussed in the context of the development of synthetic receptors. We will therefore describe here specifically only the most recent additions to the growing repertoire of reversible covalent reactions that can be used in DCC or dynamic covalent synthesis.

1.4.3.1 Reversible Benzylic Nucleophilic Substitution

Stoddart *et al.* have recently used iodide-catalyzed reversible nucleophilic substitution in the thermodynamically controlled assembly of a donor–acceptor [3]catenane **20** (Figure 1.3l) [37]. Exposure of cyclobis-(paraquat-4,4′-biphenylene) tetrakis-hexafluorophosphate (**18**) to 2 equiv. of bis-*para*-phenylene [34]crown-10 (**19**) in the presence of tetrabutylammonium iodide in acetonitrile at 80 °C led to the formation of [3]catenane **20** in 80% yield (Scheme 1.9). The S_N2 reaction



Scheme 1.9 [3]Catenane **20** has been generated using thermodynamically controlled templated synthesis [37].

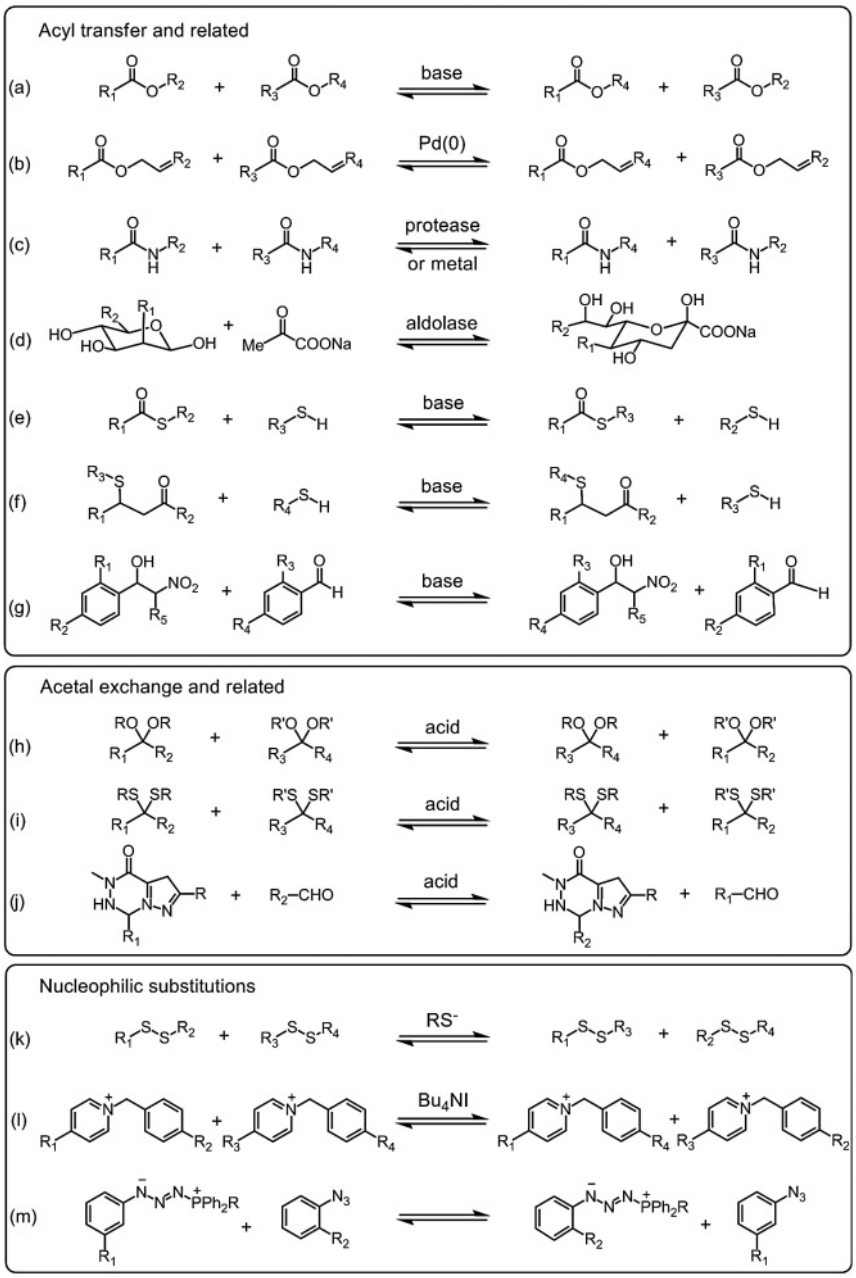


Figure 1.3 Reversible reactions used in DCLs to date: (a) transesterification, (b) transallylesterification, (c) transamidation, (d) aldol exchange, (e) transthioesterification, (f) Michael/retro-Michael reactions, (g) nitroaldol exchange, (h) acetal exchange, (i) thioacetal exchange, (j) pyrazolotriazone metathesis, (k) disulfide exchange, (l) reversible benzylic nucleophilic substitution, (m) phosphazide exchange,

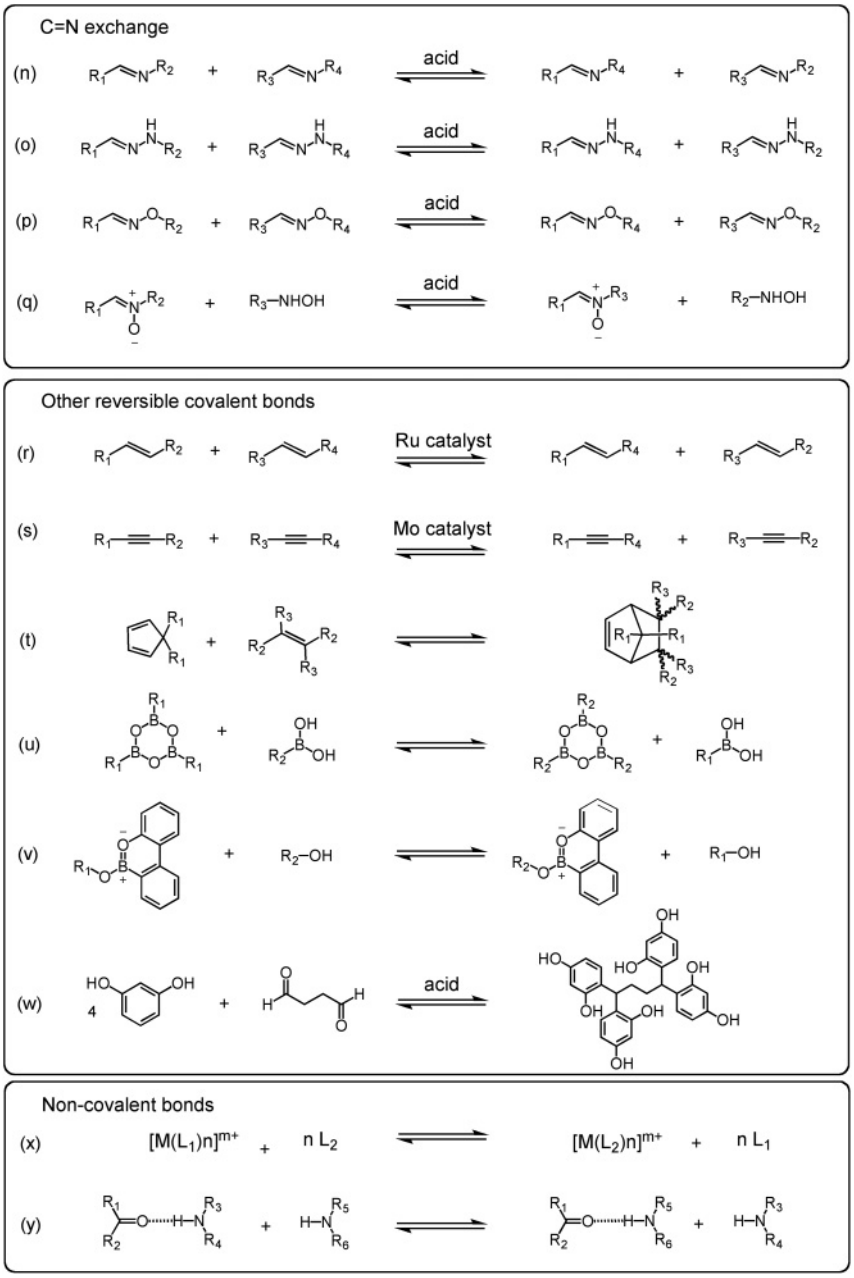
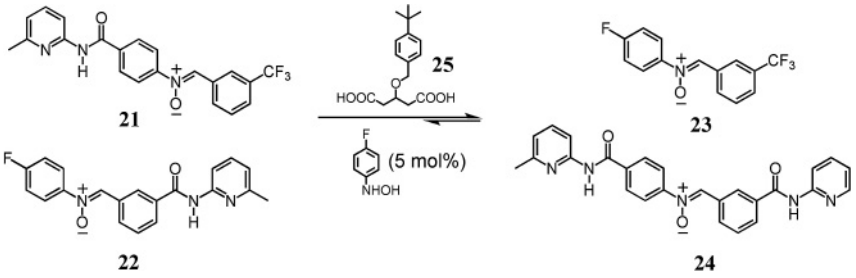


Figure 1.3 (*Continued*) (n) imine exchange, (o) hydrazone exchange, (p) oxime exchange, (q) nitrone exchange, (r) alkene metathesis, (s) alkyne metathesis, (t) Diels–Alder/ retro-Diels–Alder reactions, (u) reversible boroxine formation, (v) transboroxoaromatic esterification, (w) reversible resorcinol and 1,4-butanedial condensation, (x) metal–ligand exchange, and (y) hydrogen bond exchange.

at benzylic carbons has traditionally been viewed as reversible, with the nucleophiles alternatively acting as leaving groups and vice versa. However, it is only very recently, employing tetrabutylammonium iodide as a catalyst, that the reaction reversibility has been optimized for use in dynamic covalent synthesis.

1.4.3.2 Nitrone Exchange

Diaryl nitrones undergo exchange in chloroform under acidic conditions in the presence of a catalytic amount of a hydroxylamine (Figure 1.3q) [38]. Philp *et al.* first utilized this reaction in DCC to generate a small DCL of nitrones at equilibrium after 48 h (Scheme 1.10). Selection and amplification of nitrone **24** was observed upon addition to the library of a dicarboxylic acid capable of binding to two amidopyridine recognition groups. Philp's group have further used nitrone exchange as the basis for their studies of self-templating, extrapolating upon the basic concepts of DCC to investigate replication processes [39].



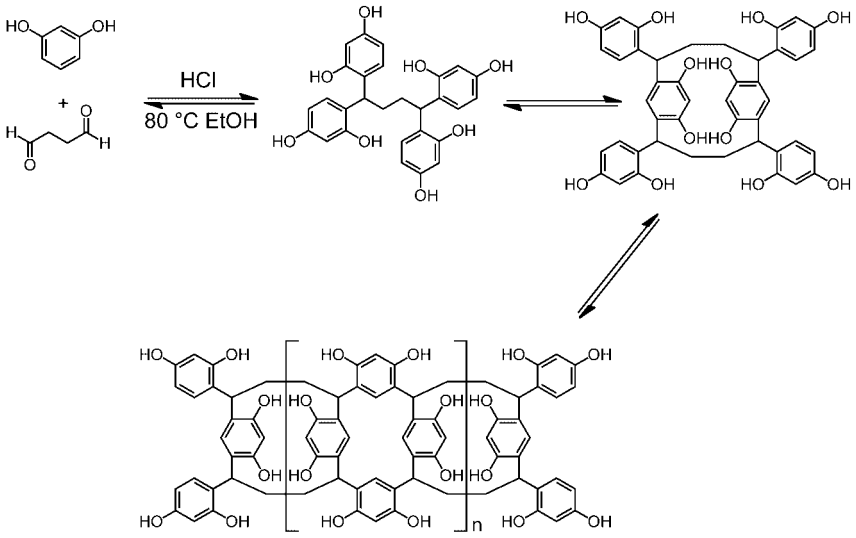
Scheme 1.10 Addition of dicarboxylic acid **25** to this small library of nitrones shifts the equilibrium to favor formation of nitrone **24** that has two amidopyridine recognition groups [38].

1.4.3.3 Reversible Nitroaldol Reaction

The reaction between nitroethane and aromatic aldehydes in the presence of triethylamine in chloroform reaches equilibrium within hours (Figure 1.3g). Selection of library members through a subsequent tandem irreversible reaction (i.e., Henry-iminolactone rearrangement) has been demonstrated [40].

1.4.3.4 Reversible Resorcinol and Alkanedial Condensation

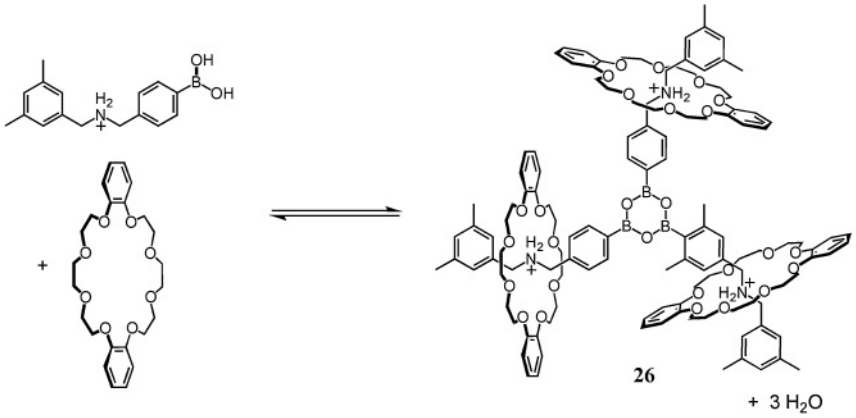
A DCL of short polymers has been generated from the condensation of resorcinol and 1,4-butanedial (Scheme 1.11) [41]. Resorcinol and 2,5-dimethoxytetrahydrofuran (3.3 equiv.) were combined in ethanol at 80 °C in the presence of HCl. Acid-catalyzed decomposition of the 2,5-dimethoxytetrahydrofuran led to the formation of 1,4-butanedial. Each aldehyde then underwent two condensation reactions with resorcinol to form ladder-like polymers containing calixarene moieties under thermodynamic control.



Scheme 1.11 A DCL of ladder-like calixarene containing polymers was generated from the reversible condensation of resorcinol and 1,4-butandial [41].

1.4.3.5 **Reversible Boroxine Formation**

Boroxine formation, through the cyclotrimerization of three boronic acid units, is another reversible reaction recently applied in dynamic covalent synthesis (Figure 1.3u). The forward reaction is entropically driven by the release of water molecules upon condensation and is favored where electron-donating groups in the *para* position are used [42]. To date, this reaction has only been used in the designed thermodynamic synthesis of a C_3-symmetric [4]rotaxane **26**, apparently under thermodynamic control (Scheme 1.12) [43]. However, it could be envisaged that

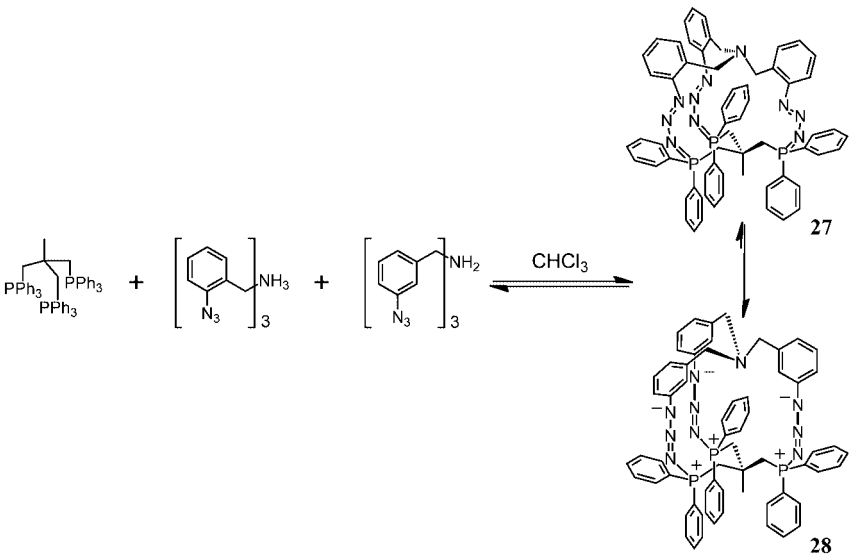


Scheme 1.12 Templated synthesis of a C_3-symmetric [4]rotaxane **26** generated through reversible boroxine formation [43].

diverse DCLs could be generated using building blocks functionalized with two or more boronic acid functional groups.

1.4.3.6 **Phosphazide Exchange**

A new reversible reaction for DCC has emerged from the synthesis of macrobicyclic tri-λ^5-phosphazides (Scheme 1.13) [44]. When carried out in chloroform, the reaction between tris-(*ortho*- or *meta*-azobenzyl)amines and triphos is dynamic and reversible. Selection of the *meta*-substituted triphosphazide from a dynamic mixture was observed as a result of greater stability of this less sterically congested species. In fact, complete triazide exchange was seen when tris-(*meta*-azobenzyl) amine was added to the *ortho*-triphosphazide **27**.



Scheme 1.13 Phosphazide exchange is rapid in chloroform: here the equilibrium favors the more stable *meta*-substituted triphosphazide **28** [44].

1.4.3.7 **Transboroxoaromatic Esterification**

The transesterification of boroxoaromatic esters is currently being developed for DCC. Thus far, Philp *et al.* have shown that small libraries of diesters can be generated rapidly under thermodynamic control from the reaction of boroxoaromatics with bis-alcohols (Figure 1.3v) [45]. The formation of more complex cyclic superstructures has not yet been achieved due to the low reactivity of the bifunctional boroxoaromatics tested.

1.4.3.8 **Future Reactions**

The search for other reversible reactions that would be suitable for use in DCLs is ongoing. Using computational methods, Houk *et al.* have investigated substitu-

ents effects in thermal azide 1,3-dipolar cycloadditions with a view to optimizing the reversibility of the reaction [46]. These theoretical studies have led to the preliminary prediction that the reaction between methanesulfonylazide and *N*,*N*′-dimethylvinylamine (Scheme 1.14) would be rapid and reversible at low concentrations required for DCLs. However, the instability of the enamine in water is expected to cause practical difficulties.

Scheme 1.14 The reaction between methanesulfonazide and *N*,*N*′-dimethylvinylamine is predicted to proceed rapidly and be reversible at micromolar concentrations [46].

1.5 Conclusions

In just over a decade, DCC has emerged as an efficient and powerful approach for generating and exploring systems reliant upon molecular recognition. It has been shown to offer an attractive route to the synthesis of complex molecules, with useful and unanticipated recognition properties, not easily accessible by other means. Additionally DCLs, viewed as complex networks of molecules, have provided a platform to study the emergent properties of systems. During this short period, the field has expanded to applications far beyond those originally conceived. These are explored in the following chapters.

References

1 (a) Corbett, P.T., Leclaire, J., Vial, L., West, K.R., Wietor, J.-L., Sanders, J.K.M., and Otto, S. (2006) *Chem. Rev.*, **106**, 3652–3711.
(b) Ladame, S. (2008) *Org. Biomol. Chem.*, **6**, 219–226.
(c) Lehn, J.-M. (2007) *Chem. Soc. Rev.*, **36**, 151–160.
(d) Rowan, S.J., Cantrill, S.J., Cousins, G.R.L., Sanders, J.K.M., and Stoddart, J.F. (2002) *Angew. Chem. Int. Ed.*, **41**, 898–952.

2 Ludlow, R.F. and Otto, S. (2008) *Chem. Soc. Rev.*, **37**, 101–108.

3 Xu, S. and Giuseppone, N. (2008) *J. Am. Chem. Soc.*, **130**, 1826–1827.

4 Giuseppone, N. and Lehn, J.-M. (2006) *Chem Eur. J.*, **12**, 1715–1722.

5 Ingerman, L.A. and Waters, M.L. (2009) *J. Org. Chem.*, **74**, 111–117.

6 Giuseppone, N. and Lehn, J.-M. (2006) *Angew. Chem. Int. Ed.*, **45**, 4619–4624.

7 Sreenivasachary, N. and Lehn, J.-M. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 5938–5943.

8 Chow, C.F., Fujii, S., and Lehn, J.-M. (2007) *Chem. Commun.*, 4363–4365.

9 Hutin, M., Cramer, C.J., Gagliardi, L., Shahi, A.R.M., Bernardinelli, G., Cerny, R., and Nitschke, J.R. (2007) *J. Am. Chem. Soc.*, **129**, 8774–8780.

10 Lowe, G. (1995) *Chem. Soc. Rev.*, **24**, 309–317.

11 Watson, J.D. and Crick, F.H.C. (1953) *Nature*, **171**, 737–738.

12 Siedel, F. (1926) *Chem. Ber.*, **59B**, 1894–1908.

13 Melson, G.A. and Busch, D.H. (1965) *J. Am. Chem. Soc.*, **87**, 1706–1710.

14 (a) Linstead, R.P. (1934) *J. Chem. Soc.*, 1016–1017.
(b) Byrne, G.T., Linstead, R.P., and Lowe, A.R. (1934) *J. Chem. Soc.*, 1017–1022.

15 Thompson, M.C. and Busch, D.H. (1964) *J. Am. Chem. Soc.*, **86**, 213–217.

16 Busch, D.H. (1992) *J. Inclusion Phenom.*, **12**, 389–395.

17 Anderson, S., Anderson, H.L., and Sanders, J.K.M. (1993) *Acc. Chem. Res.*, **26**, 469–475.

18 Thompson, M.C. and Busch, D.H. (1962) *J. Am. Chem. Soc.*, **84**, 1762–1763.

19 Goodwin, J.T. and Lynn, D.G. (1992) *J. Am. Chem. Soc.*, **114**, 9197–9198.

20 Goodman, M.S., Jubian, V., Linton, B., and Hamilton, A.D. (1995) *J. Am. Chem. Soc.*, **117**, 11610–11611.

21 Bilyk, A. and Harding, M.M. (1995) *J. Chem. Soc., Chem. Commun.*, 1697–1698.

22 Try, A.C., Harding, M.M., Hamilton, D.G., and Sanders, J.K.M. (1998) *Chem. Commun.*, 723–724.

23 Brady, P.A., Bonar-Law, R.P., Rowan, S.J., Suckling, C.J., and Sanders, J.K.M. (1996) *Chem. Commun.*, 319–320.

24 Brady, P.A. and Sanders, J.K.M. (1997) *J. Chem. Soc., Perkin Trans. 1*, 3237–3253.

25 Hasenknopf, B., Lehn, J.-M., Boumediene, N., Dupont-Gervais, A., Van Dorsselaer, A., Kneisel, B., and Fenske, D. (1997) *J. Am. Chem. Soc.*, **119**, 10956–10962.

26 Huc, I. and Lehn, J.-M. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 2106–2110.

27 Swann, P.G., Casanova, R.A., Desai, A., Frauenhoff, M.M., Urbancic, M., Slomczynska, U., Hopfinger, A.J., Le Breton, G.C., and Venton, D.L. (1996) *Pept. Sci.*, **40**, 617–625.

28 Klekota, B., Hammond, M.H., and Miller, B.L. (1997) *Tetrahedron Lett.*, **38**, 8639–8642.

29 Sakai, S., Shigemasa, Y., and Sasaki, T. (1997) *Tetrahedron Lett.*, **38**, 8145–8148.

30 Eliseev, A.V. and Nelen, M.I. (1997) *J. Am. Chem. Soc.*, **119**, 1147–1148.

31 Hioki, H. and Still, W.C. (1998) *J. Org. Chem.*, **63**, 904–905.

32 Oh, K., Jeong, K.S., and Moore, J.S. (2001) *Nature*, **414**, 889–893.

33 Lehn, J.-M. (2005) *Prog. Polym. Sci.*, **30**, 814–831.

34 Chichak, K.S., Cantrill, S.J., Pease, A.R., Chiu, S.H., Cave, G.W.V., Atwood, J.L., and Stoddart, J.F. (2004) *Science*, **304**, 1308–1312.

35 Rowan, S.J., Hamilton, D.G., Brady, P.A., and Sanders, J.K.M. (1997) *J. Am. Chem. Soc.*, **119**, 2578–2579.

36 Cacciapaglia, R., Di Stefano, S., and Mandolini, L. (2008) *J. Phys. Org. Chem.*, **21**, 688–693.

37 Patel, K., Miljani, O.Š., and Stoddart, J.F. (2008) *Chem. Commun.*, 1853–1855.

38 Turega, S.M., Lorenz, C., Sadownik, J.W., and Philp, D. (2008) *Chem. Commun.*, 4076–4078.

39 Sadownik, J.W. and Philp, D. (2008) *Angew. Chem. Int. Ed.*, **47**, 9965–9970.

40 (a) Vongvilai, P., Angelin, M., Larsson, R., and Ramström, O. (2007) *Angew. Chem. Int. Ed.*, **46**, 948–950.
(b) Angelin, M., Vongvilai, P., Fischer, A., and Ramström, O. (2008) *Chem. Commun.*, 768–770.
(c) Angelin, M., Fischer, A., and Ramström, O. (2008) *J. Org. Chem.*, **73**, 3593–3595.

41 Kudo, H., Shigematsu, K., Mitani, K., Nishikubo, T., Kasuga, N.C., Uekusa, H., and Ohashi, Y. (2008) *Macromolecules*, **41**, 2030–2036.

42 Tokunaga, Y., Ueno, H., Shimomura, Y., and Seo, T. (2002) *Heterocycles*, **57**, 787–790.

43 Tokunaga, Y., Ito, T., Sugawara, H., and Nakata, R. (2008) *Tetrahedron Lett.*, **49**, 3449–3452.

44 Alajarín, M., Berná, J., López-Leonardo, C., and Steed, J.W. (2008) *Chem. Commun.*, 2337–2339.

45 Greig, L.M., Slawin, A.M.Z., Smith, M.H., and Philp, D. (2007) *Tetrahedron*, **63**, 2391–2403.

46 Jones, G.O. and Houk, K.N. (2008) *J. Org. Chem.*, **73**, 1333–1342.

2
The Practice of Dynamic Combinatorial Libraries: Analytical Chemistry, Experimental Design, and Data Analysis

Sijbren Otto

2.1
Introduction

Dynamic combinatorial chemistry (DCC) [1–3] was originally developed to solve the problem of obtaining compounds with distinctive molecular recognition properties. These could be either ligands for biomolecules or synthetic receptors for small-molecule guests. Obtaining such molecules by a design-and-synthesis approach remains a daunting task; we simply do not sufficiently understand issues like conformation and cooperation of the many individually weak noncovalent interactions that together determine efficiency in molecular recognition. In DCC this problem is circumvented by limiting the design to only small fragments, which are assembled into the desired structures under the direction of the molecules themselves. A description of the basic principles of DCC is provided in Chapter 1. Chapters 3 and 5 in this book provide ample proof that this approach indeed facilitates efficient access to synthetic receptors and small-molecule ligands alike. The concept of DCC has also been explored in the context of catalyst development (Chapter 4) and polymer chemistry (Chapter 6).

In many of these applications of DCC the purpose is to identify a specific molecule with exceptional properties. The dynamic combinatorial library (DCL) is then merely a vehicle through which this molecule may be generated and discovered. In a typical experiment a library is prepared by mixing building blocks, and its composition is analyzed in the absence and presence of an added template molecule or some other stimulus. Any compound in the mixture that is significantly amplified upon addition of the template should be a strong binder. This principle still drives much of the work in DCC, but is arguably rather simplistic. DCLs are complex mixtures of species that are interconnected through a set of equilibria and mass balances, and can behave in ways that are not always immediately intuitive, particularly when the focus is on the amplification of individual molecules. In order to fully appreciate the behavior of a DCL it should be regarded in its entirety as a molecular network that strives to attain the lowest overall Gibbs energy. One of the implications is that the extent to which template binding and amplification are correlated can vary, depending in part on how the experiments

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

are set up (i.e., the idea that amplification is selective for the fittest species [4] may not always hold). The complex behavior of DCLs has been investigated in some detail, initially with a view to identifying the experimental conditions that are optimal for discovering new synthetic receptors and small-molecule ligands. Interestingly, this new perspective on dynamic libraries has sparked the development of new applications of DCC, such as in sensing (described in detail in Chapter 7), which rely on the unique characteristics of the entire DCL. Such applications go significantly beyond the purpose for which the approach was first developed.

This chapter focuses on aspects of DCC common to nearly all its applications. The analytical tools with which DCLs can be analyzed are discussed in Section 2.2. Section 2.3 explores the behavior of DCLs as networks, including network topology and the computational strategies that have been developed to guide experimental design. Finally, Section 2.4 outlines some of the new opportunities provided by DCLs, including data mining, information processing, and the potential for use of DCLs to explore molecular recognition within and between library members.

2.2 Analytical Methods

Despite research efforts in DCC spanning well over a decade, the development of the associated analytical chemistry is still limited. This largely reflects the fact that most studies have aimed at establishing proof-of-principle and relied on relatively small libraries, often containing only a single building block, producing merely a handful of library members. Yet, as also pointed out in a critical review by Ladame [3], before DCC will fulfill its promise, much larger libraries will need to be analyzed. With recent developments in analytical equipment and automated data analysis such studies are well within current analytical capabilities. Indeed, it is possible to obtain useful results from libraries with close to 10 000 constituents [5] and theory suggests that larger libraries are likely to produce even better results.

Screening DCLs requires quantification of large numbers of compounds in a complex mixture. More often than not the different components vary widely in concentration. Thus, an ideal analytical technique will combine high sensitivity with a large dynamic range (i.e., allowing small quantities of a compound A to be quantified in the presence of a large excess of B). While good results have been obtained by mass spectrometric analysis, the most promising techniques are multidimensional, allowing library members to be separated on the basis of two or more orthogonal characteristics. The most commonly used and probably most powerful technique is liquid chromatography-mass spectrometry (LC-MS), but also other techniques such a time-resolved heteronuclear single quantum correlation nuclear magnetic resonance (HSQC NMR) [6] or diffusion-ordered spectroscopy (DOSY) NMR [7] have found application, albeit as yet only in smaller libraries. These techniques are discussed in more detail in Sections 2.2.2 and 2.2.3. Alternatively, screening can also be

performed by using immobilization strategies similar to those developed in traditional combinatorial chemistry. This approach has been pioneered by Miller [8] and is discussed in Section 2.2.4.

2.2.1 MS Analysis

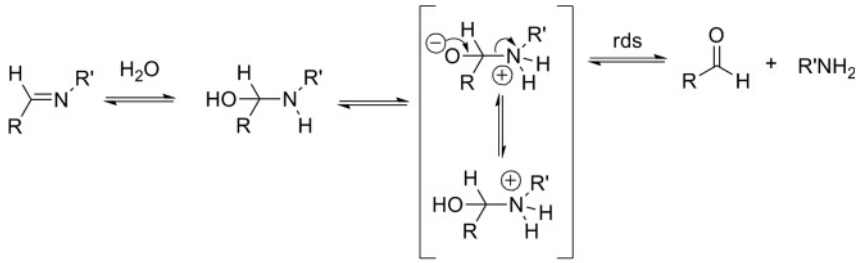
MS is an attractive tool to analyze DCLs, owing to its high resolution and the fact that it provides the masses of the library members, which is often essential in order to identify the molecular structures. A disadvantage of the technique is that its use for quantification is not straightforward.

Poulsen has used Fourier transform ion cyclotron MS to identify compounds with affinity for bovine carbonic anhydrase II from a DCL of hydrazones [9]. Direct injection of the DCL revealed a cluster of signals resulting from noncovalent complexes of the protein with selected library members. The large size of the protein (29 kDa) causes the mass signals to have extensive and overlapping isotopic envelopes, which prevented direct identification of the bound library members. However, upon collision-induced gas-phase fragmentation it was possible to confirm the identity of the selected hydrazone ligands.

2.2.2 LC-MS Analysis

Analysis of DCLs by a combination of high-performance liquid chromatography (HPLC) and MS (LC-MS) has proven very powerful, provided the library members are stable enough to allow chromatographic separation. LC-MS analysis can yield both the quantity and the identity of individual library members. Care should be taken that the exchange between library members is slow on the timescale of the experiment. This is usually the case for hydrazone and disulfide libraries, without requiring special precautions. It has recently been reported that even the composition of DCLs of hydrolytically labile imines can be analyzed directly by HPLC, provided an acidic mobile phase is used. Beau *et al.* successfully used a common reversed-phase HPLC solvent system based on acetonitrile/water mixtures containing 0.1% trifluoroacetic acid [10]. Imine hydrolysis is slow under these conditions because at low pH the zwitterionic hemi-aminal intermediate that is believed to be important in the hydrolysis process is more difficult to attain (Scheme 2.1).

In principle, LC-MS analysis can be used for libraries that are substantially larger than those typically used to date. For larger libraries, baseline separation of most library members can no longer be expected. In fact, in DCLs of more than a few tens of compounds overlapping peaks are more likely than separated ones. However, this need not be problematic as it should be possible to identify all compounds using the mass spectrum at a given retention time. In fact, a comparison between untemplated and templated libraries may be best made on the basis of the two corresponding mass chromatograms.



Scheme 2.1 The mechanism for imine hydrolysis under acidic conditions is believed to involve the decomposition of the zwitterionic hemi-aminal intermediate as the rate-determining step [10]. The rate of decomposition slows down at lower pH due to the concentration of the zwitterionic intermediate being reduced by protonation of the oxygen atom.

As the masses of all potential library members can be predicted from the nature of the building blocks, it should be possible to automate the procedure of comparing mass chromatograms for all potential library members between templated and untemplated libraries. This may then lead to a computer-generated shortlist of amplifications, which can then be verified, within a reasonable time, by hand and by some selected additional experiments using only the building blocks incorporated into the amplified compounds.

While I am not aware of reports where such automated analysis has been carried out, we have recently reported an analysis of a large DCL (approximately 9000 compounds) using LC-MS in which amplifications were detected by simply comparing the HPLC ultraviolet (UV) traces [5]. The specific library was made by mixing eight different thiol building blocks (**1**–**8** in Figure 2.1) and screened for receptors for ephedrine (**9**). A comparison of the HPLC chromatograms of the untemplated and templated libraries is shown in Figure 2.1(a and b, respectively), showing some clear amplifications in a poorly resolved mixture. The nature of the compounds corresponding to the amplified peaks was established by an LC-MS study, which revealed some complications that are likely to be common when using large libraries. The mass spectrum at the retention time where amplification was detected revealed the presence of several library members.

Fortunately, comparing this spectrum of the templated with that of the untemplated library allowed establishing the mass of the amplified species. Figure 2.1(c and d) illustrates this for one of the amplified signals. However, it turned out that there were several potential library members with a similar mass. For the example of Figure 2.1(d), compounds $(\mathbf{6})_2(\mathbf{8})_2$, $(\mathbf{4})_2(\mathbf{7})(\mathbf{8})_2$, $(\mathbf{2})_2(\mathbf{6})_3$, and $(\mathbf{2})_4(\mathbf{3})$ are all within 0.5 amu of the observed mass of 922.7. Distinguishing between these could theoretically be possible if the mass were determined at high resolution or from the isotopic substitution pattern. In our case we were able to resolve the issue by analyzing the MS-MS fragmentation pattern. Thus, using LC-MS-MS analysis we were finally able to identify the amplified species as cyclic oligomers of building blocks **6** and **8**. Subsequent binding studies on isolated material revealed that these

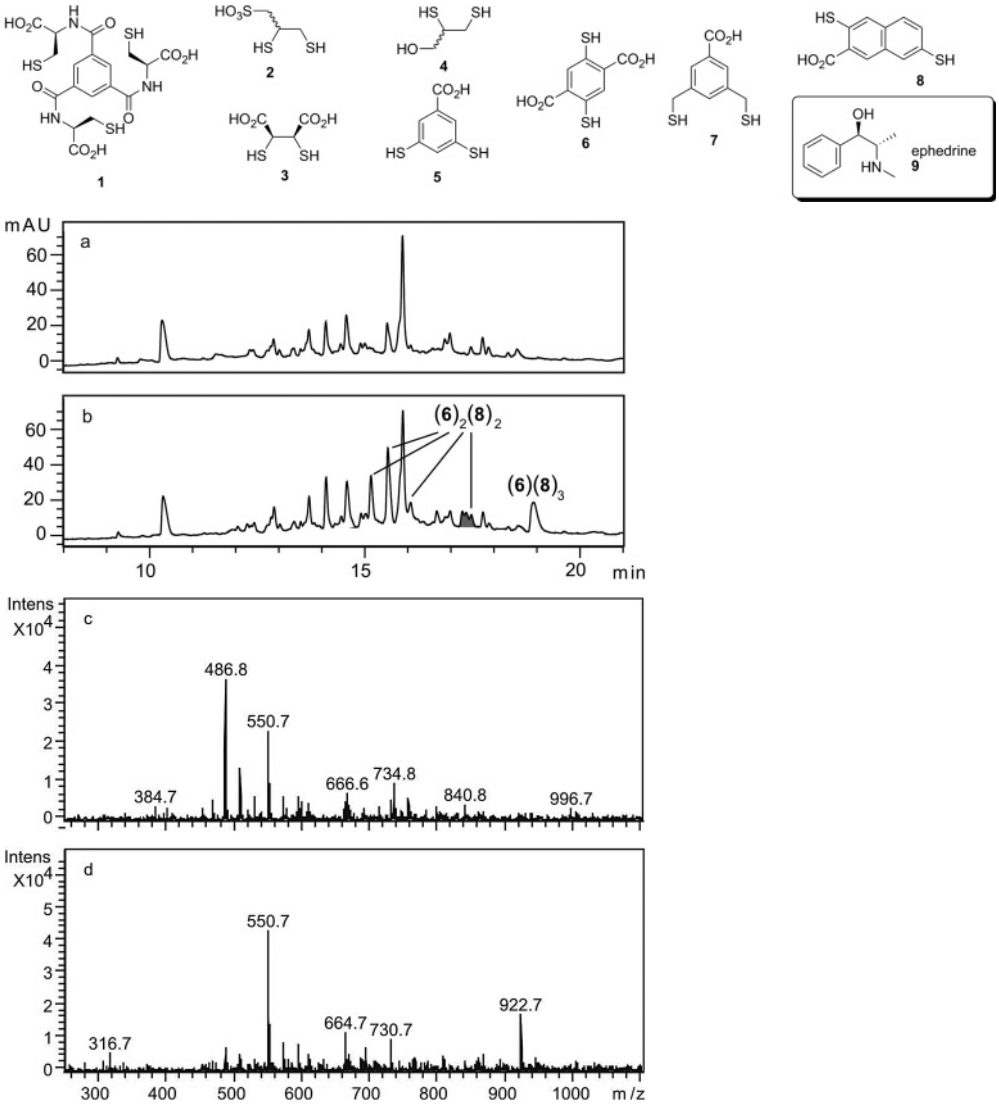


Figure 2.1 HPLC chromatograms of DCLs made from building blocks **1**–**8** (5.0 mM total concentration in 50 mM pH 8.0 borate buffer) (a) in the absence of template; and (b) in the presence of 5.0 mM ephedrine **9**. Mass spectrum of the HPLC peak at 17.5 min of the DCLs in the absence (c) and presence (d) of ephedrine **9**. Data reproduced from [5].

macrocycles have affinities for the ephedrine template in the order of −24 kJ/mol. Note that even these relatively modest affinities can induce clearly detectable template effects in large DCLs.

LC-MS has been used by Gagné *et al.* in an elegant study showcasing some clever analytical chemistry for investigating chiral recognition in DCLs [11, 12]. The

authors screened a library made from a racemic hydrazide/aldehyde building block for binding to a chiral template using an optical rotation detector. In this way only chiral recognition events are detected. Using nucleosides as templates they indeed observed enantioselective amplification of specific macrocyclic receptors. In order to assess which stereoisomer was selected the authors prepared a pseudo-racemic mixture of the building block in which one of the enantiomers was labeled with deuterium. Analysis by LC-MS now allowed all HPLC signals to be assigned to the various stereoisomeric products, as these products now have different masses.

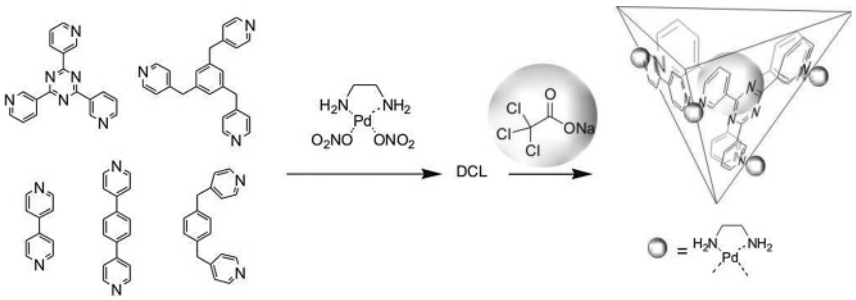
Work by the same group has recently also shown that caution is required in the analysis of DCLs by LC-MS [13]. The labile nature of the covalent bonds inherent to DCC (in this case the hydrazone linkages) can cause unwanted reactions to occur between coeluting species during desolvation in the mass spectrometer. For example, tetrameric species were detected that could only be explained by a reaction between two coeluting dimeric molecules. This, together with the always present risk of fragmentation during desolvation, serves as a clear warning to be critical and keep an open mind when interpreting MS data, and to perform the necessary control experiments to minimize the chances of misinterpretation of the data.

2.2.3 Analysis by NMR

Analysis of DCLs by NMR has the advantage that it can also be applied in instances where library members are highly labile, as may be the case for noncovalent reversible chemistries. Unlike analyses involving physical separation of the constituents (chromatography, MS), NMR is normally performed on the complete mixture. However, this is at the same time a disadvantage as the spectra of the mixtures quickly become very difficult to interpret, particularly when compounds give rise to multiple signals.

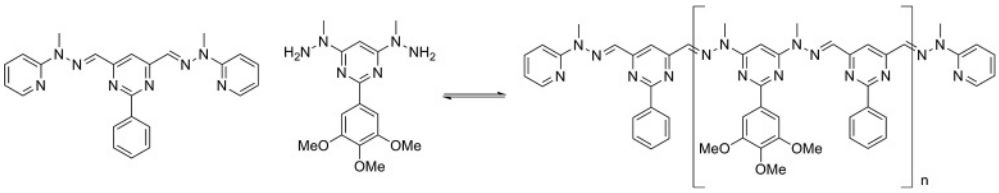
One solution to this problem is to simplify the interpretation by making use of difference spectra. Fujita *et al.* have applied this approach to the analysis of a small DCLs of capsules by mixing two tripodal pyridine ligands with three dipodal ones (Scheme 2.2), using palladium centers to link the ligands together [14]. The authors were able to identify a new receptor for sodium trichloroacetate from this DCL. Subtracting the spectrum of the library in the absence of guest from that in the presence of the trichloroacetate guest allowed them to identify building blocks from which the capsule was constructed. A second round of similar NMR experiments was necessary to also obtain the building block stoichiometry.

Another approach to improve the tractability of the spectra is to introduce additional dimensions complementary to the chemical shift. Two such approaches have been explored: Lehn *et al.* have used DOSY NMR, where the differences in diffusion times of the library members are used as the second dimension [7], and Jeannerat and Prins *et al.* have used ^{1}H–^{13}C-HSQC NMR and also introduced a temporal dimension, by adding different building blocks at different times and monitoring the evolution of the two-dimensional (2-D) NMR spectrum [6].



Scheme 2.2 Mixing di- and tripyridyl ligands with Pd(en) produces a diverse DCL of cyclic and cage coordination compounds. Exposure of this library to sodium trichloroacetate induces the amplification of a new receptor. Adapted from [14].

The DOSY technique [15] relies on a spin-echo pulse sequence that involves gradient pulses that lead to a *Z*-gradient in the magnetic field experienced by the sample. Two such pulses are given that effectively cancel each other in the absence of any diffusion, but not when diffusion takes place. From the reduction in signal intensity the diffusion coefficient of the molecules in the sample can be estimated, which is related to their hydrodynamic radii. The result of a DOSY experiment is a 2-D NMR spectrum with the chemical shift on one axis and the diffusion coefficient on the other. This technique was used by Giuseppone *et al.* in a proof-of-principle study on a DCL based on scandium triflate-mediated hydrazone exchange. After 12 h of data collection they were able to distinguish between five different levels of diffusion rates that were assigned to linear hydrazones of different lengths (n = 0–4 in Scheme 2.3) [7]. It is encouraging that the relatively small changes in size of the library members were resolvable on the diffusion time axis.



Scheme 2.3 Building blocks used to generate a small DCL of linear hydrazides that can adopt helical conformations and that was analyzed by DOSY NMR to reveal five different compounds (n = 0–4) [7].

The temporally resolved HSQC approach by Jeannerat and Prins *et al.* has uncovered a wealth of kinetic and thermodynamic data on their small hydrazone DCL, reproduced in Figure 2.2. The authors started by adding hydrazide **B** to a solution of hydrazone **10A** and analyzed the mixture using HSQC, which is a 2-D NMR technique in which the ^{13}C spectrum is plotted on one axis and correlated to the ^{1}H spectrum on the other axis. Use of this 2-D technique was essential to

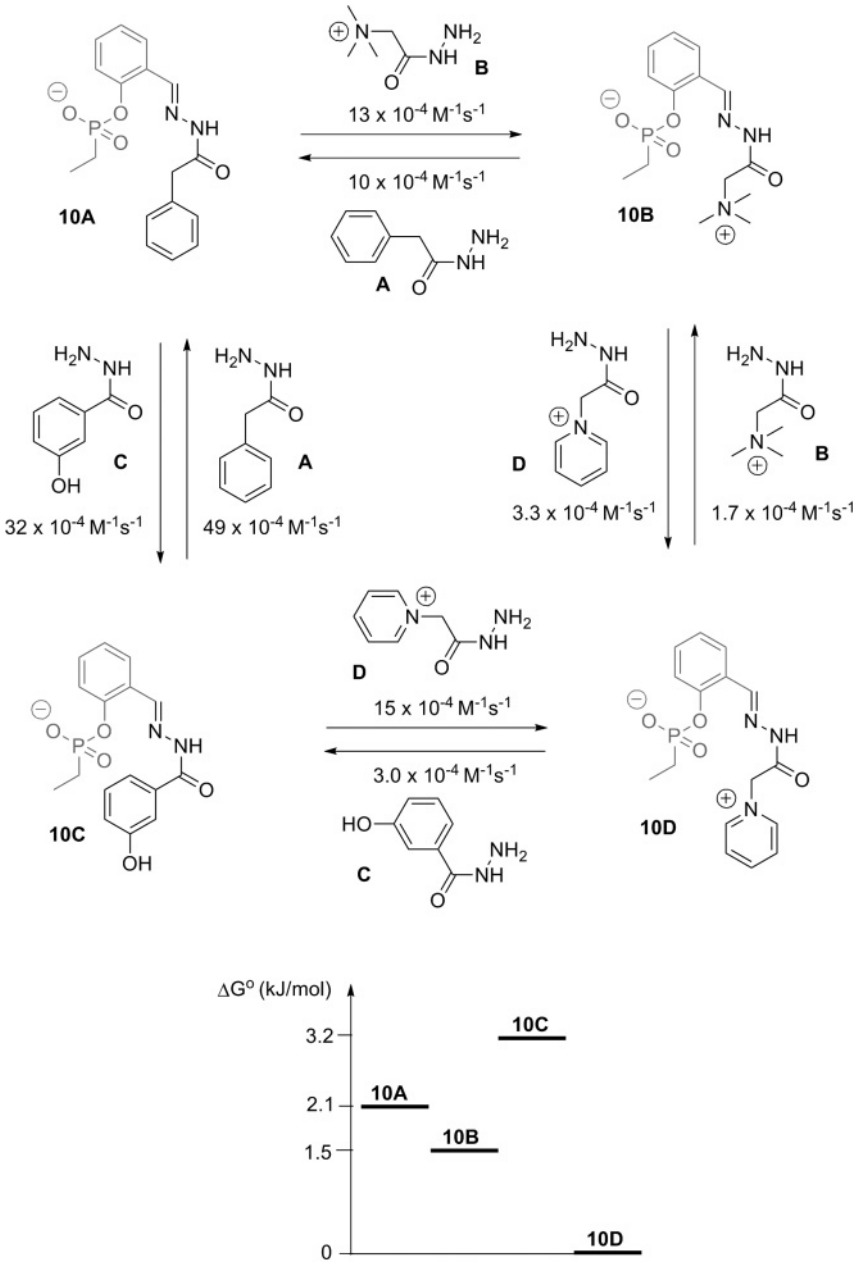


Figure 2.2 A small library of hydrazones that was analyzed by time-resolved ^{1}H–^{13}C HSQC NMR to give all kinetic and thermodynamic parameters [6].

quantify the products in the DCL as these gave mostly overlapping signals in the one-dimensional ^{1}H NMR. They then monitored the exchange process in time, from which they were able to determine the equilibrium constants and also the rate constants for the two hydrazone exchange reactions. They then proceeded by adding hydrazide **C** to the solution and after equilibrium had been reached, they finally added hydrazide **D**. On the basis of the resulting data, they were able to fully characterize the thermodynamics and kinetics of the system [6].

2.2.4
Resin-Bound DCC

A drastically different method for identification of library members is possible when individual building blocks are bound to a solid-phase resin. This approach has been developed by Miller *et al.* who set up multiphase DCLs by mixing solution-phase building blocks with identical resin-bound ones and a fluorescent target [8, 16].

Chemical tagging or spatial segregation informs on the nature of the resin-bound building blocks, while screening of the resin for binding of the fluorescent target enables identification of the building blocks that are part of strongly binding library members.

This approach was successfully implemented for the development of a ligand that selectively binds to an RNA sequence implicated in the replication of the human immunodeficiency virus (HIV). A library of disulfides was made of 150 resin-bound thiol-functionalized building blocks that were allowed to equilibrate with their solution-phase counterparts in the presence of the fluorescently labeled RNA target. From the library with a theoretical diversity of 11 325 members, a small number of beads turned fluorescent and were isolated. After cleavage of the bound compounds, the identity of the resin-bound building block was identified by MS, suggesting three different building blocks were involved in generating strong binders (**11**–**13** in Figure 2.3). A new experiment was then set up in which the nine different binary combinations of resin-bound and solution-phase building blocks were screened for target binding. This revealed an interesting feature of this two-phase approach: the experiment with **11** bound to the resin and **13** in solution resulted in highly fluorescent beads, while in the reverse experiment, where **11** is in solution and **13** resides on the resin, no fluorescence was detected.

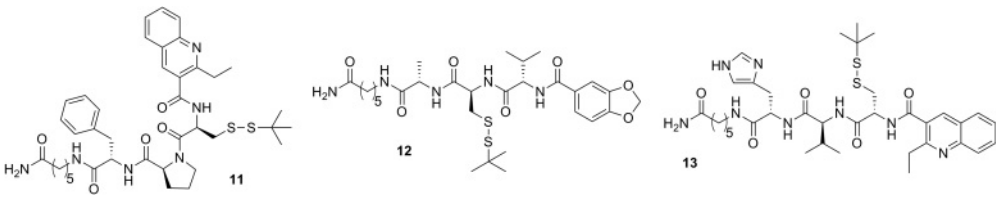


Figure 2.3 Thiol building blocks that were selected from a resin-bound disulfide DCL when exposed to a HIV RNA target [16].

This is a consequence of competition between resin-bound and solution-phase disulfides for the limiting amount of fluorescent target. With **13** on the resin, immobilized **13–11** competes with solution-phase **11–11**, while in the reverse experiment no **11–11** can be formed. Thus, the outcome suggested that **11–11** is a stronger binder than **13–11**, which was subsequently confirmed by bioassays on isolated material.

The resin-bound DCL strategy lends itself to further increases in library size, in particular when diversity is generated from binary combinations of monofunctionalized building blocks. Usage of multifunctionalized building blocks may also be possible, but will require more elaborate deconvolution experiments.

2.3 Experimental Design

DCLs are complex molecular networks [17] and their behavior is not always immediately intuitive. Moreover, how a particular DCL will behave depends on the molecular recognition events taking part between the various species present in solution. Many of these will be unknowns at the outset of a dynamic combinatorial experiment; indeed, learning more about these is often the goal of the experiment. As will become apparent in this section, when competing binding events take place this may complicate the interpretation of the results. A careful design of the experiment is required to minimize such complications. Also, some insight in the, at first sight, counterintuitive behavior that DCLs may exhibit is indispensable in order to be able to correctly interpret the data. The remainder of this chapter describes some of the idiosyncrasies of DCLs and provides some guidelines on how to properly design dynamic combinatorial experiments.

2.3.1 Importance of Template Concentration

Much of the early development of DCC was based on the assumption that exposing a DCL to a target will result in selective amplification of the best binding library member. As first pointed out by Severin *et al.* [18, 19], this assumption is not necessarily valid. DCLs respond to external stimuli, such as the introduction of a target, in a way that minimizes the Gibbs energy of the entire system. In the presence of excess template the system may have a preference for amplifying a large number of smaller macrocycles over a small number of larger ones. Similarly, in libraries containing several different building blocks, the amplification of mixed macrocycles is preferred over macrocycles containing only one type of building block (Figure 2.4).

This behavior can be rationalized by considering the total template binding energy gained by the system. In the presence of excess template the building blocks are limiting. For example, where a system has a choice between producing two copies of a dimeric library member versus one copy of a tetramer, it may prefer

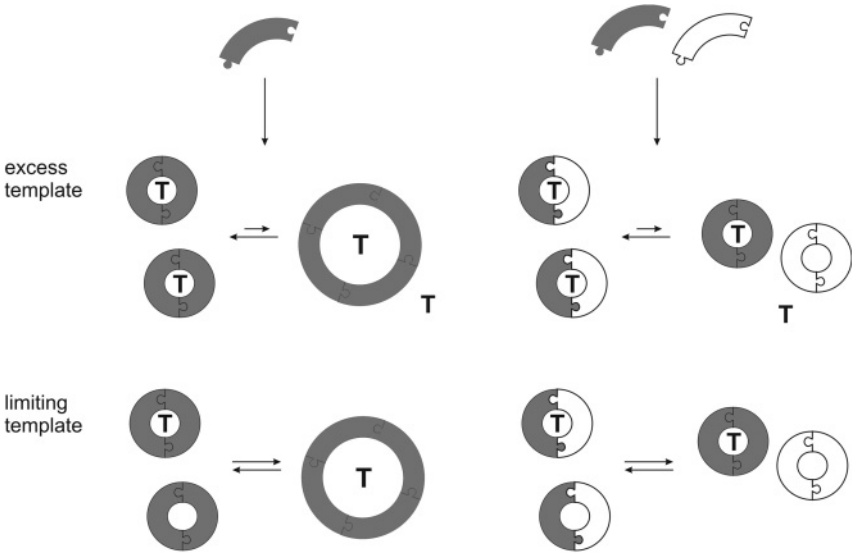


Figure 2.4 Competition between small and large library members is biased towards the former only when excess template is available. Similarly, only in the presence of excess template is the competition between hetero- and homo-oligomers biased towards the hetero-compounds.

to amplify the dimer as it can gain twice the template–dimer binding energy versus only once the template–tetramer binding energy. In libraries made from several different building blocks the system can normally produce more of a hetero-oligomer than of a homo-oligomer of the same size. In this scenario the system is able to harvest a larger number of template–hetero-oligomer binding energies than template–homo-oligomer energies. Thus, in the presence of excess template amplification is biased towards small oligomers and towards hetero-oligomers; these compounds may be amplified in preference over larger and/or homo-oligomers, even in cases where the latter are the stronger binders. This leads to an undesirable situation where the magnitude of the amplifications of the various library members may not correlate very well with their affinities for the template. Fortunately, there is a solution to this problem: simply reduce the template concentration. When the amount of template is limiting the competition is no longer between, for example, two dimers and one tetramer (or several hetero-oligomers versus one homo-oligomer), but between one dimer and one tetramer (or one hetero-oligomer and one homo-oligomer). Under such conditions the strongest binder will be amplified selectively.

The importance of the template concentration is illustrated by the behavior of the small disulfide DCL shown in Figure 2.5, which features two different macrocyclic receptors **17** and **18** that both bind ammonium guest **16**, with host **18** binding the strongest [20]. Thus, if amplification would be selective for the fittest

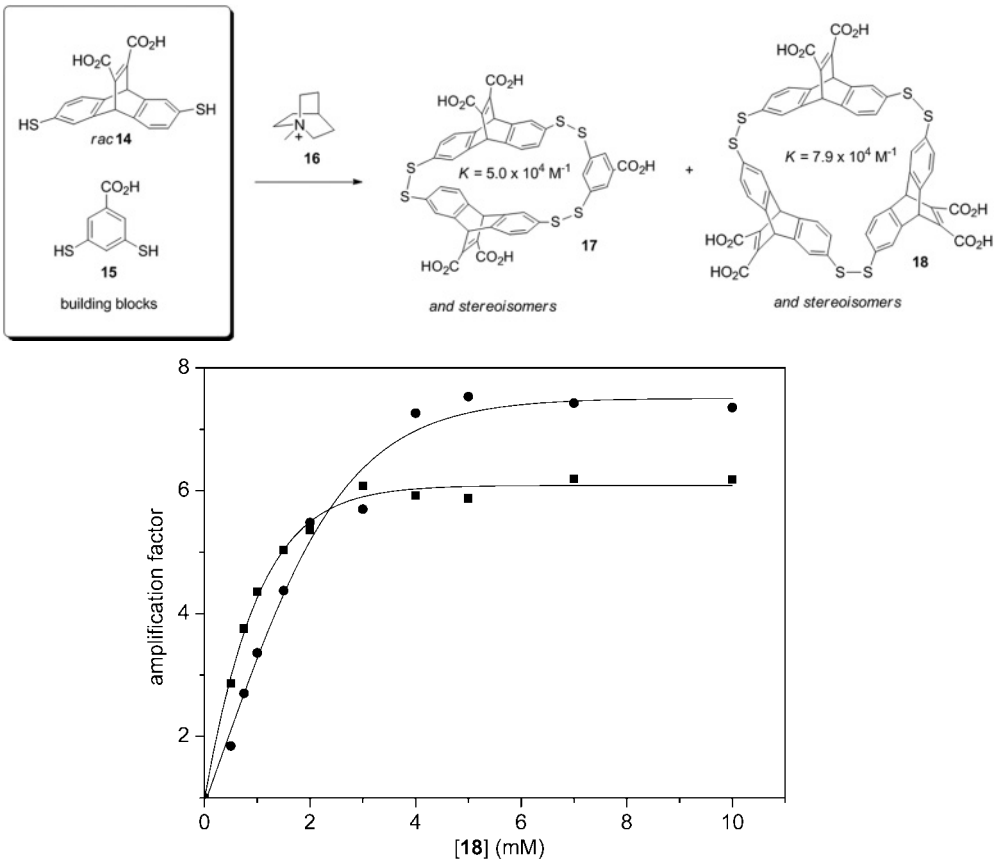


Figure 2.5 A small DCL made from thiol building blocks **14** (3.33 mM) and **15** (1.67 mM) produces a mixture of receptors **17** and **18** for guest **16**. Amplification factors for hosts **17** (●) and **18** (■) depend on the concentration of template **16** [20].

receptors then host **18** should be amplified more than **17**. This is indeed observed when low concentrations of guest are present. However, when the guest concentration is increased, the reverse is observed. The system is able to gain more binding energy by producing **17** as compared to **18**, simply because it can make more of it. Receptor **17** is built up from two units of building block **14** and one of **15**, while host **18** contains three units of **14**. Thus, with a limiting amount of **14** the system can make three copies of heterotrimer **17** against two of homotrimer **18**. If there is sufficient guest present to bind to all receptors it is more advantageous to produce more of the weaker binder as opposed to less of the stronger one. However, at low guest concentrations, where the amount of guest is limiting, amplification is selective for the best binder.

The above example makes it clear that the amount of template in the system is an important and easily controllable experimental variable. There is a trade-off

between two desirables: on the one hand, it is advantageous to have strong amplifications, for which it is beneficial to have a large concentration of template present; on the other hand, a large excess of template may result in amplification becoming less selective for the best binder. This raises the question which experimental conditions are optimal. We have addressed this issue computationally using in-house developed DCLSim software (see Box 2.1).

Using DCLSim, we simulated libraries that were based on an arbitrarily chosen number of seven building blocks, which were allowed to form all possible cyclic library members up to tetramers, resulting in a DCL of 322 library members. Each of these was assigned a binding affinity for the template that was drawn randomly from a normal distribution. The mean value for $\log(K)$ was set at 2 (i.e., the average binding constant was $100\,M^{-1}$) and the standard deviation for $\log(K)$ was 1. We simulated DCLs using a range of building block and template concentrations, and assessed the correlation between host–guest binding energy and the amplification factor. For each set of experimental conditions a graph comparable to those shown in Figure 2.6 was obtained.

The results showed considerable variability from one run to the next, even when the only difference was the random assignment of binding affinities to library members. A particularly clear-cut example is shown in Figure 2.6(a and b), which shows the behavior of DCLs run at 10 mM template and 10 mM building block concentrations. Where the DCL in Figure 2.6(a) gives a good correlation between binding affinity and amplification factor, that in Figure 2.6(b) gives a particularly poor correlation, with no amplification at all for the best binder. When performing dynamic combinatorial experiments it is not know *a priori* how the binding constants are distributed over the various library members, so it is a matter of chance how good the binding affinities will correlate with amplification factors for a given set of experimental conditions. In order to obtain more clarity we repeated the simulations of the type of Figure 2.6 a total of 50 times each for a variety of experimental conditions. This allowed the quality of the correlation between amplification factor and binding affinity (as quantified by the mean correlation coefficient R^2) to be assessed as a function of template and building block concentrations. The resulting 2-D graph is shown in Figure 2.7. It shows that in general running DCLs of the type simulated here, using a template to building block ratio of 1 : 10 ensures satisfactory correlations.

Going back to the problem-case of Figure 2.6(b): when we resimulated this library at a 10-fold reduced template concentration of 1.0 mM the correlation has improved markedly. However, the magnitudes of most of the amplification factors are reduced (Figure 2.6c). This may be undesirable, so particularly when screening new DCLs it may be better to use a modest excess of template to increase the likelihood of detecting any template effects. A second round of screening can then be performed using a smaller amount of template to assess whether the observed template effects are indeed pointing to the better binders in the system.

Apart from a careful choice of template and building block concentrations there are also other strategies that can minimize the occurrence of troublesome library behavior such as that shown in Figure 2.6(b). As this behavior is a result

Box 2.1 Simulating DCLs

Computer simulations are an important tool for improving our understanding of the often capricious behavior of DCLs. Simulating thermodynamic equilibria (without considering molecular structures) is relatively straightforward and it is very simple to subsequently modify *in silico* the position of specific equilibria in order to investigate their influence on the system. Any comparable experimental study would require collecting data for an extremely large number of equilibria and is therefore not a viable alternative. Moreover, in an experimental system it is not normally possible to adjust the value of selected equilibrium constants to probe how the system responds, while *in silico* this is simply a matter of typing in another number.

For simulating *small* DCLs of the order of 10–20 library members commercial packages are available. The general mathematic package MathCAD has been used and also off-the-shelf kinetics software [18, 19] has been used to model equilibrium systems. However, when equilibria involve a significantly larger number of compounds (hundred to thousands) such packages are no longer adequate. This has prompted us to develop DCLSim software [21, 22] that is able to cope with DCLs of this size. At the core of DCLSim is a modified version of the COGS algorithm [23] that finds equilibrium distributions using a series of equilibrium constants and mass balances as input. More specifically, DCLSim requires as input the equilibrium constants describing the formation of library members from their constituent building blocks and (where appropriate) the Gibbs energies of binding of the library members to a template molecule. After the concentrations of the various building blocks and the template have also been specified DCLSim then simulates the resulting library distribution, giving the concentrations of all library members in the absence and presence of the template as output.

Note that the equilibrium constants describing the formation of library members from their building blocks is not always chemically realistic. For example, in disulfide chemistry the conversion of thiols into disulfides is an irreversible process. However, treating this process as a fictitious equilibrium is equivalent to going through a series of more chemically realistic exchange steps and mathematically much simpler to deal with.

In addition to manually assigning template binding affinities to library members DCLSim contains functionalities that enable these affinities to be assigned randomly. This is a useful tool for the exploration of libraries where it is not known beforehand how binding affinities will be distributed. In situations like these multiple simulations are necessary in order to arrive at statistically significant conclusions. Carrying out multiple simulations that only differ in the random assignment of equilibrium constants also gives important insights into the inherent variability in library behavior. DCLSim allows simple statistical analyses to be carried out on the relatively large datasets that result.

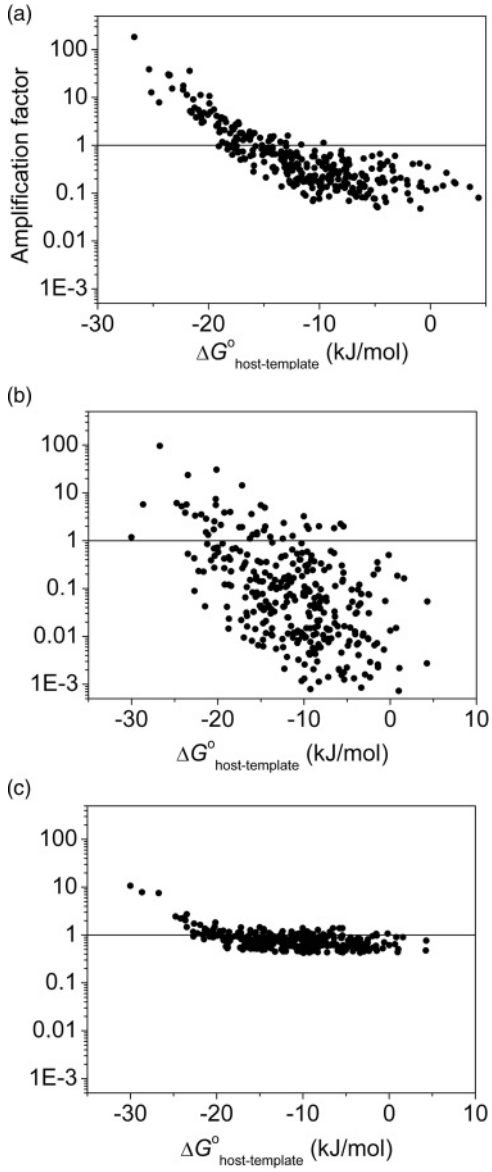


Figure 2.6 (a and b) The relationship between amplification and free energy of binding for all binders in two randomly generated DCLs that differ only in the way the binding constants are distributed over the various library members. In both DCLs the total concentration of the building blocks and the concentration of the template is 10 mM. (c) The library of (b) simulated at a reduced template concentration of 1.0 mM [22].

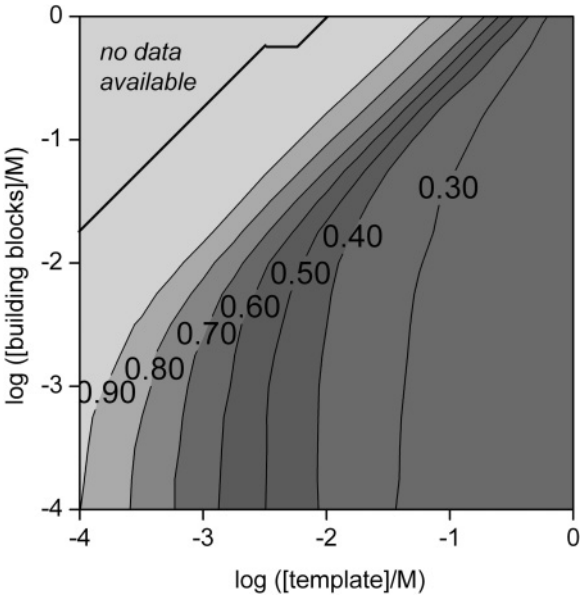


Figure 2.7 Correlation between binding affinity and amplification in simulated DCLs, as a function of template and total building block concentration. The numbers indicate the correlation coefficient (averaged over 50 simulations) [22].

of a situation in which there is competition between library members for scarce building blocks, poor correlations can be avoided by ensuring that building blocks do not become scarce. For example, if the reversible chemistry is such that most of the building blocks remain free in solution (e.g., in imine chemistry in water) then any amplification tends to be selective for the best binder. An alternative strategy that achieves essentially the same situation is to use a limiting amount of a linker unit. An example of such a system is shown in Figure 2.8 in which dihydroxypyridine ligands X and Y serve as linkers that hold together a macrocycle of three half-sandwich metal-ligand complexes A–D [24]. When the system is set up with a limiting amount of X or Y, then a DCL of different macrocycles is generated that can draw on an excess of building blocks A–D. Any template-induced amplifications in such a system will more directly reflect the binding affinity. This was shown for the addition of lithium salt for which the strongest binder is $(BY)_3$ while $(BY)_2DY$ also binds. In the presence of a reservoir of A–D the best binder is faithfully amplified, while in the absence of such reservoir (using stoichiometric amounts of Y) the inferior binder $(BY)_2DY$ is the dominant species.

In summary, careful design of the building blocks and the experiments, and a critical and prepared mind when interpreting the results, are essential before the correct conclusions can be drawn from the often complex response of DCLs to molecular recognition events.

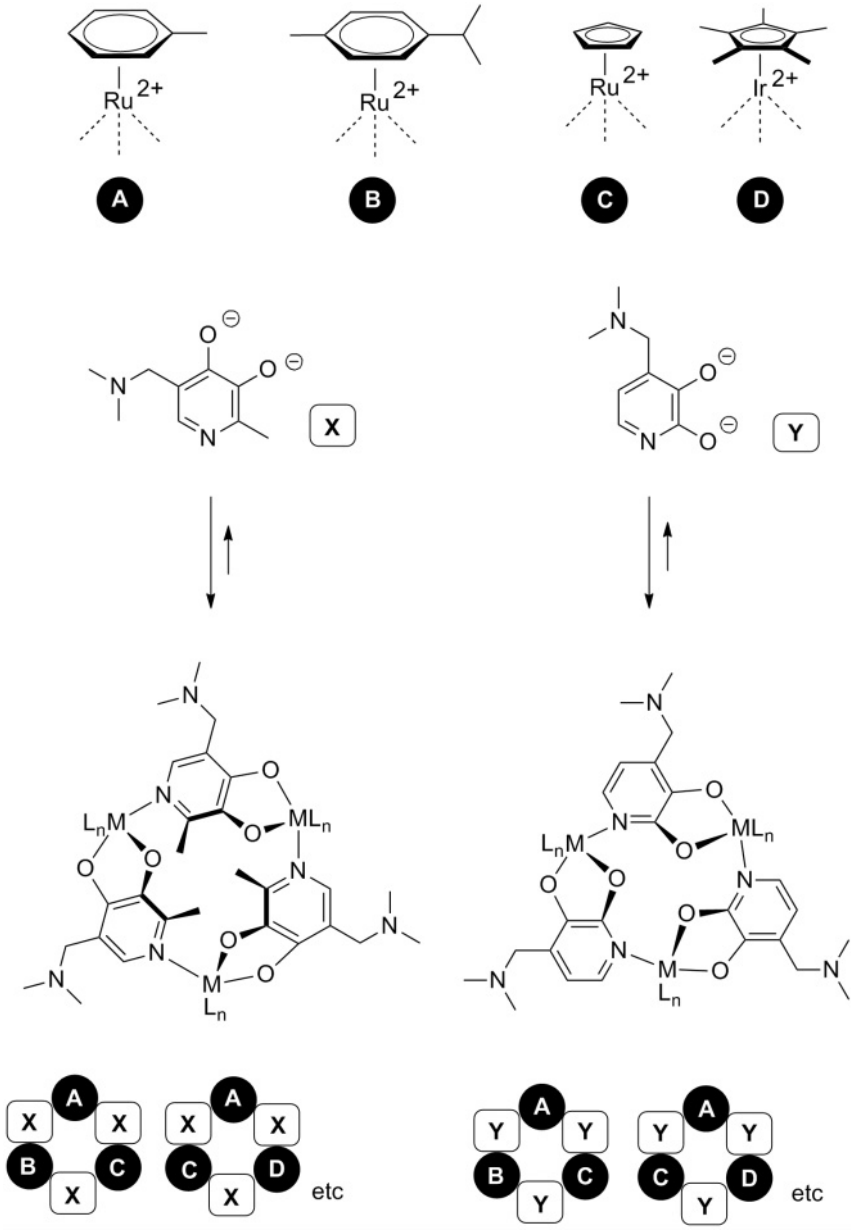


Figure 2.8 DCL made from a metal-half-sandwich complexes (A–D) held together by dihydroxypyridine linkers X and Y. Using a limiting amount of linker and an excess of each of the metal-half-sandwich complexes generates a DCL in which a reservoir of the building blocks A–D is present, increasing the chances of selective amplification of the best binder [24].

2.3.2 Library Size

In Section 2.2 we have already shown that it is possible to analyze DCLs of the order of 10 000 compounds. The future will tell whether libraries that are significantly larger than these are also tractable – thus far there are no indications that they are not. This raises the question of how large a DCL should be in order to have the highest probability of identifying a truly outstanding binder? Will template-induced shifts in product distributions still be useful in very large libraries? These are questions that have been asked from the very outset of the development of DCC [21, 25]. At this stage experimental studies on large libraries are not sufficiently numerous to allow statistically significant conclusions to be drawn. Theoretical approaches to these questions have led to different answers. In early work, Moore and Zimmerman concluded that template-induced shifts will only be of limited use in very large libraries [25]. In this work the product distribution of the DCL was approximated by a continuum. This is a fair approximation for the bulk of the material. However, in every combinatorial approach the interest is usually not in the bulk of the material but in the outliers – the best binders that are found at the tail end of the distribution. The tail ends of any distribution are not treated adequately by a continuum model. Repeating the simulations by Moore and Zimmerman, but using discrete species showed that useful amplifications were still possible even in libraries as large as 1 000 000 species [21]. In this work a highly simplified model for a DCL was used in which every species could convert into every other library member. In reality, libraries are not that adaptive (e.g., building block A cannot usually turn itself into building block B).

In order to get a more realistic sense of amplification effects in very large DCLs we recently performed a number of more detailed simulations in which we varied the size of the DCL from 65 (made from four building blocks) to 4828 compounds (made from 16 building blocks) [26]. Template binding affinities were once again assigned randomly from a normal distribution (mean binding constant $100\,M^{-1}$ and standard deviation $10\,M^{-1}$). Based on our experience with the LC-MS detection in large libraries, we assumed that we would be able to detect any compound that was amplified at least 2-fold and for which the difference between its concentration in the untemplated relative to the templated library represented at least 1% of the concentration of the most abundant species in the mixture. This detection cut-off means that as the library gets larger the probability that we are able to detect amplification events diminishes. This is illustrated by Figure 2.9, which shows the fraction of the compounds that are in the top 10% by affinity for the template that were above this detection cut-off as a function of the experimental conditions (template and total building block concentrations). For the larger libraries only a few percent of the best binders are in fact detectable, while most of these compounds will go unnoticed.

This raises the question to what extent the library size influences the affinity of the most amplified compounds that are within detection limits. In order to quantify this we looked at the average affinity of the best binder from among the three

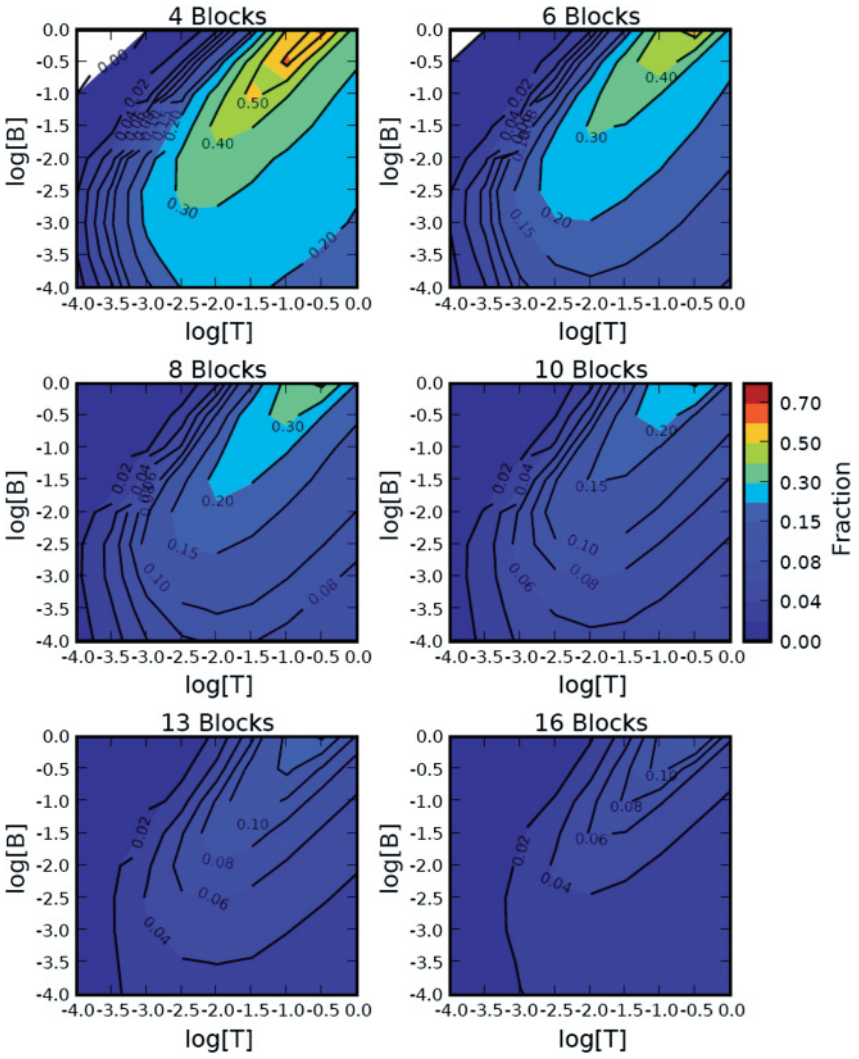


Figure 2.9 Fraction of the library members constituting the top 10% by affinity that are within detection limits for various library sizes (i.e., number of building blocks). Each graph surveys different combinations of template and building block concentrations [26].

highest *detectable* amplifications. The results of this analysis are shown in Figure 2.10 as a function of the experimental conditions.

The first thing to notice is the trend that with increasing library size the affinity of the detected library members increases as well. For every library size analyzed the highest affinity library members are obtained in the top left corner of the graphs (i.e., using a high concentration of building blocks and a low concentration of template). However, this combination very often fails to produce any template

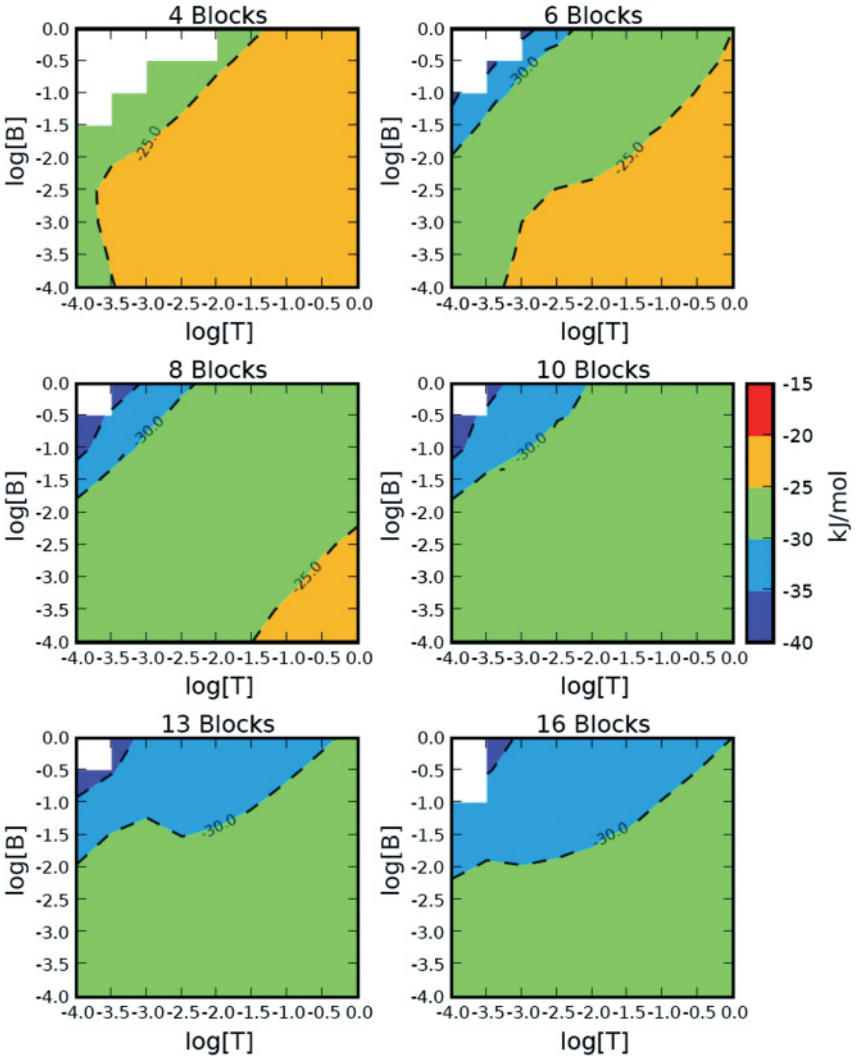


Figure 2.10 Binding energy of the best binder of the three most amplified detectable library members for various library sizes (i.e., number of building blocks). Each graph surveys different combinations of template and building block concentrations [26].

effects at all – the graphs only show the results for the rare occasions in which a detectable amplification effect occurs. The probability of detecting amplification effects rises as the ratio of template to building block increases (i.e., from the top left to the bottom right of the graphs) as shown in Figure 2.11. It is interesting to note that larger libraries are characterized by a higher probability of detecting amplifications, presumably because of the increased probability that they contain strong binders.

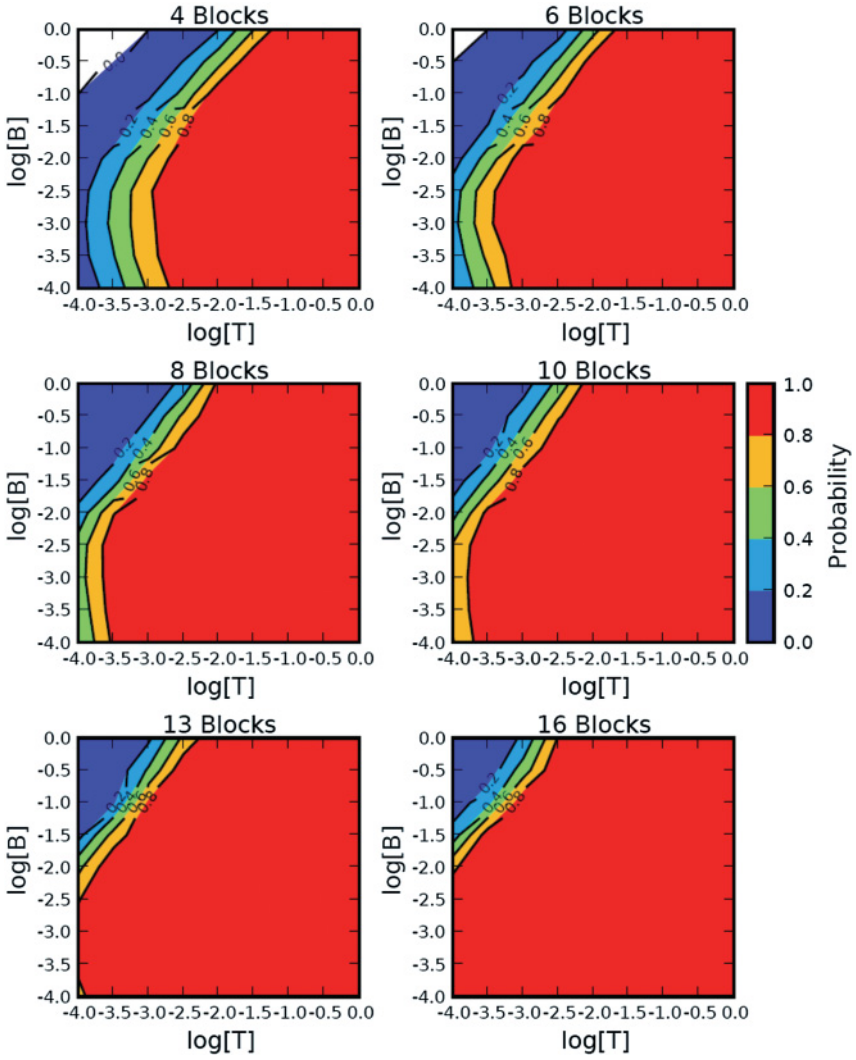


Figure 2.11 Probability of detecting any amplification for various library sizes (i.e., number of building blocks). Each graph surveys different combinations of template and building block concentrations [26].

Comparing Figures 2.9 and 2.10 indicates that in selecting experimental conditions for the DCLs a trade-off has to be made between the probability of detecting amplifications and the average affinity of the library members causing the amplifications. It appears that a total building block to template ratio of 10:1 gives a good initial compromise. If experiments under these conditions fail to produce any amplification it is advisable to increase the concentration of template.

In conclusion, within the constraints of the chosen model (normal distribution of binding affinities) and within the range of library sizes investigated, it appears that bigger is better – bigger libraries are more likely to produce strong binders, despite the fact that the inevitable limitations of the analytical techniques mean that most library members will remain undetected.

2.4 Data Analysis

2.4.1 Quantifying Equilibrium Constants

The purpose of many DCL experiments is the identification of compounds that are efficient at molecular recognition. As pointed out in the previous sections, with careful experimental design DCLs may be used to identify such compounds by comparing the amplification factors of the various library members. However, given the complex nature of library behavior some means of verifying the molecular recognition behavior of the amplified compounds is desirable. Amplification factors of the selected hits alone are not always reliable when it comes to comparing the binding behavior of the amplified library members and are certainly inadequate for establishing absolute affinities. Traditionally, binding affinities are determined by isolating the compounds in question and subsequently studying their behavior using any one of a variety of different experimental tools (i.e., host–guest titrations by NMR, isothermal titration calorimetry, fluorescence, etc.). This is often a painful and time-consuming process. Isolating selected library members from the other compounds in the DCL can be hard and binding studies are often time-consuming.

DCLs provide a promising alternative approach to establishing the binding properties of the compounds of interest. We have recently shown that it is possible to obtain host–guest binding affinities directly from library distributions, without the need for isolation of any of the library members. Instead, library distributions were determined for a set of different experimental conditions (template and building block concentrations). Once analytical methods have been established it is relatively straightforward to collect such datasets. Equilibrium constants for host–guest interactions can then be fitted to this data. We have developed a dedicated DCLFit software tool for this purpose [27]. An impression of how well the fitted values of the equilibrium constants for host–guest interactions approach the real values was obtained by using a series of simulated DCL compositions generated using DCLSim and based on known binding constants as the input. DCL compositions for a set of 12 different experimental conditions (different ratios of the three building blocks and different template concentrations) were simulated. After introducing random errors into this data (similar to those expected for a true experimental dataset) it was used as input for DCLFit, which produced the data shown in the black bars in Figure 2.12 as output. The predicted binding energies for the stronger

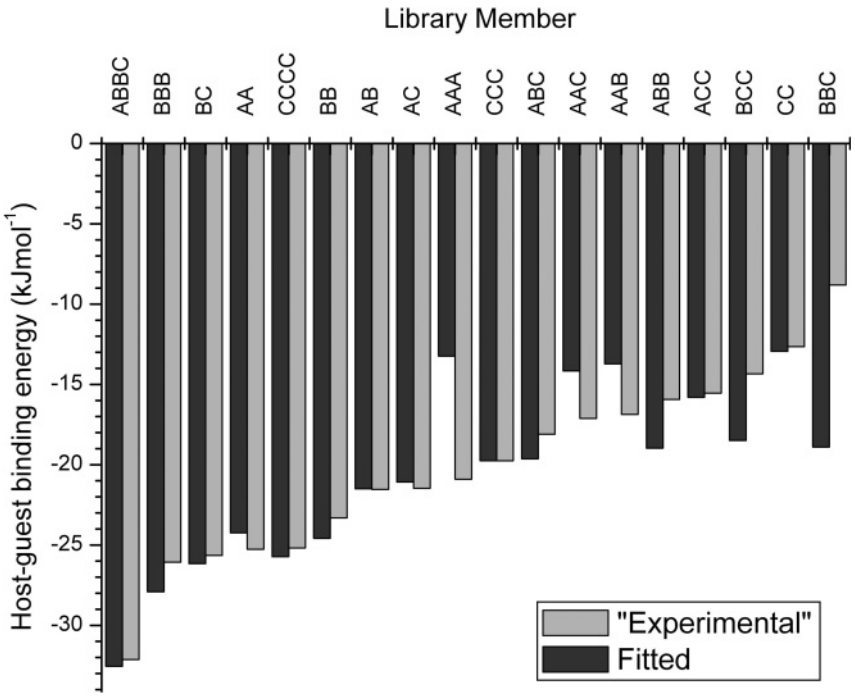


Figure 2.12 Comparison of "experimental" and fitted values for the host–guest binding energies in a simulated 31-component DCL [27].

binders are in good agreement with the real values. The error for the weaker binders is substantially larger, as the concentrations of these species tend not to respond very much to the changes in template concentration. Application of DCLFit in experimental systems are now starting to appear [27, 28]. For a smaller DCL a similar approach to determining binding constants has been reported making use of off-the-shelf software [12].

These fitting approaches allow multiple host–guest binding constants to be estimated in parallel, illustrating that taking a global "systems" view of molecular networks has important advantages over the reductionist approach of focusing on individual host–guest interactions.

2.4.2 DCLs as Sensors

Another example that uses the global response of dynamic libraries to a template is the application of DCLs as sensors [29–32]. The idea is that the pattern of change in the concentration of all network members carries the signature of the template (i.e., analyte) to which the DCL is exposed. If the library members have different optical properties (i.e., different UV-visible absorption spectra) then the change in

library composition will be reflected by a change in the UV spectrum of the entire mixture. The challenge is now to correlate this change to the nature and/or concentration of the analyte. For this purpose several mathematical pattern recognition protocols are available. Using such multivariate analysis methods, Severin *et al.* have succeeded in using DCLs as sensors for peptides [31, 32] and nucleotides [30]. This subject is treated in detail in Chapter 7.

2.5 Conclusions

DCC has primarily developed as a tool for discovering new molecules with special properties. The approach shifts the challenge away from the design and synthesis of complex structures to the analysis of complex mixtures. Given the number of successful new molecules identified by the still relatively small community currently working in this area, it looks as if the analytical challenges are more easily overcome than the challenges posed by design and synthesis. Add to this the fact that analytical capabilities are not yet used to their full potential and the conclusion is that we can expect many more exciting molecules to be discovered through dynamic combinatorial strategies.

It is also interesting to note that the complex behavior of dynamic libraries that complicate their use for compound discovery at the same time provide new opportunities in the form of parallel data collection and sensing as outlined in Section 2.4.

Acknowledgments

I am grateful to the students and postdocs I have had the pleasure to work with over the last few years who contributed with many ideas and much enthusiasm, most notably Peter Corbett (who wrote DCLSim) and Fred Ludlow (who wrote DCLFit) on the theoretical side, and Ana Belenguer, who has been the driving force behind the analytical chemistry.

References

1 Corbett, P.T., Leclaire, J., Vial, L., West, K.R., Wietor, J.L., Sanders, J.K.M., and Otto, S. (2006) *Chem. Rev.*, **106**, 3652–3711.

2 Lehn, J.M. (2007) *Chem. Soc. Rev.*, **36**, 151–160.

3 Ladame, S. (2008) *Org. Biomol. Chem.*, **6**, 219–226.

4 Freemantle, M. (2002) *Chem. Eng. News*, **80** (35), 31–33.

5 Ludlow, R.F. and Otto, S. (2008) *J. Am. Chem. Soc.*, **130**, 12218–12219.

6 Gasparini, G., Vitorge, B., Scrimin, P., Jeannerat, D., and Prins, L.J. (2008) *Chem. Commun.*, 3034–3036.

7 Giuseppone, N., Schmitt, J.L., Allouche, L., and Lehn, J.M. (2008) *Angew. Chem. Int. Ed.*, **47**, 2235–2239.

8 McNaughton, B.R. and Miller, B.L. (2006) *Org. Lett.*, **8**, 1803–1806.

9 Poulsen, S.A. (2006) *J. Am. Soc. Mass Spectrom.*, **17**, 1074–1080.

10 Zameo, S., Vauzeilles, B., and Beau, J.M. (2006) *Eur. J. Org. Chem.*, 5441–5444.

11 Voshell, S.M., Lee, S.J., and Gagne, M.R. (2006) *J. Am. Chem. Soc.*, **128**, 12422–12423.

12 Chung, M.K., Hebling, C.M., Jorgenson, J.W., Severin, K., Lee, S.J., and Gagne, M.R. (2008) *J. Am. Chem. Soc.*, **130**, 11819–11827.

13 Schiltz, H., Chung, M.K., Lee, S.J., and Gagne, M.R. (2008) *Org. Biomol. Chem.*, **6**, 3597–3600.

14 Kubota, Y., Sakamoto, S., Yamaguchi, K., and Fujita, M. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 4854–4856.

15 Cohen, Y., Avram, L., and Frish, L. (2005) *Angew. Chem. Int. Ed.*, **44**, 520–554.

16 McNaughton, B.R., Gareiss, P.C., and Miller, B.L. (2007) *J. Am. Chem. Soc.*, **129**, 11306–11307.

17 Ludlow, R.F. and Otto, S. (2008) *Chem. Soc. Rev.*, **37**, 101–108.

18 Grote, Z., Scopelliti, R., and Severin, K. (2003) *Angew. Chem. Int. Ed.*, **42**, 3821–3825.

19 Severin, K. (2004) *Chem. Eur. J.*, **10**, 2565–2580.

20 Corbett, P.T., Sanders, J.K.M., and Otto, S. (2005) *J. Am. Chem. Soc.*, **127**, 9390–9392.

21 Corbett, P.T., Otto, S., and Sanders, J.K.M. (2004) *Org. Lett.*, **6**, 1825–1827.

22 Corbett, P.T., Otto, S., and Sanders, J.K.M. (2004) *Chem. Eur. J.*, **10**, 3139–3143.

23 Perrin, D.D. and Sayce, I.G. (1967) *Talanta*, **14**, 833–842.

24 Saur, B., Scopelliti, R., and Severin, K. (2006) *Chem. Eur. J.*, **12**, 1058–1066.

25 Moore, J.S. and Zimmerman, N.W. (2000) *Org. Lett.*, **2**, 915–918.

26 Ludlow, R.F. and Otto, S., submitted.

27 Ludlow, R.F., Liu, J., Li, H., Roberts, S.L., Sanders, J.K.M., and Otto, S. (2007) *Angew. Chem. Int. Ed.*, **46**, 5762–5764.

28 Au-Yueng, H.Y., Pengo, P., Pantos, G.D., Otto, S., and Sanders, J.K.M. (2009) *Chem. Commun.*, 419–421.

29 Otto, S. and Severin, K. (2007) *Top. Curr. Chem.*, **277**, 267–288.

30 Buryak, A., Pozdnoukhov, A., and Severin, K. (2007) *Chem. Commun.*, 2366–2368.

31 Buryak, A. and Severin, K. (2005) *Angew. Chem. Int. Ed.*, **44**, 7935–7938.

32 Buryak, A. and Severin, K. (2006) *J. Comb. Chem.*, **8**, 540–543.

3
Development of Synthetic Receptors using Dynamic Combinatorial Chemistry

Andrea M. Escalante and Ricardo L.E. Furlan

3.1
Introduction

In order to bind a given guest, a receptor must be complementary – it must have recognition groups which are of the appropriate electronic character to complement the binding sites of the guest. Those recognition groups must be positioned on the receptor in a way that they can interact with recognition groups from the guest when both receptor and guest are in the binding conformation. Upon binding, the receptor and guest reorganize their interactions with solvent molecules (disrupting some and probably creating new ones), and change their conformation to achieve a suitable binding conformation. In the process, noncovalent interactions within each molecule may be broken or formed.

The requirement to balance all the variables involved in this recognition process makes the design of successful synthetic receptors a significant challenge.

A frequently explored shortcut that facilitates receptor design is to incorporate rigidity in order to avoid as much as possible changes in conformation that are difficult to predict. In addition, a rigid preorganized receptor will decrease the cost of unfavorable reorganization energy. However, high-affinity perfectly rigid receptors are not easy to achieve in practice since minor errors in design that demand readjustment from the receptor can be energetically very expensive [1]. Indeed, nature does not use rigidity to improve affinity, but it still produces receptors with affinities that are on average much better than the best synthetic receptors prepared to date. A recent survey of the binding efficiency of synthetic and biological hosts by Houk *et al.* indicates that synthetic systems are typically several orders of magnitude less efficient in binding small molecules than their biological counterparts [2].

Guest binding by natural receptors such as proteins is reinforced by intrareceptor interactions, which require conformational rearrangements of parts of the receptor that are very similar to the conformational rearrangement necessary to interact with the guest. Consequently the interactions that contribute to guest binding extend beyond the immediate binding site well into the protein structure [3]. Synthetic receptors that operate by the same principle will need to be able to

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

fold into conformations in which intrareceptor interactions are formed. Such characteristics increase even further the demands on predictions to be made during the design of the receptor.

Despite these difficulties, numerous receptors have been prepared to date that are able to bind guests in water and in organic solvents with affinities that range between 1 and $10^7\,M^{-1}$ [2]. However, as most practitioners in the field have found, the rational design and synthesis of novel receptors – and a stepwise, iterative approach to producing increasingly selective receptors – can be both time-consuming and frustrating [4].

In the 1990s, an alternative to the iterative approach for the identification of selective peptide receptors was pioneered by Still [5]. Mirroring developments in medicinal chemistry, a combinatorial approach was used to prepare libraries of possible receptor structures, which could then be screened to identify a receptor for a given substrate. Twenty years later, combinatorial methods are clearly regarded as useful tools to optimize designs, facilitate access to new hosts, and ultimately aid in the understanding of host–guest interactions.

Developing synthetic receptors using a combinatorial strategy is attractive as it reduces the level of detail required in the design of the receptors. Instead of having to design and synthesize complete receptors, the chemist designs and synthesizes a series of potential fragments of receptors (building blocks) that will be combined in different assemblies.

Moving from receptor design to library design permits the incorporation of different recognition groups and shapes as well as levels of rigidity in the set of building blocks to be combined. It then becomes unnecessary to know in detail the exact combination of those variables that will produce a good receptor, as long as such combination is present within the library.

The dynamic combinatorial strategy goes one step further – ideally the preferred receptor is not only selected by the guest, but also amplified at the expense of the unselected compounds within the library [6]. This is achieved through the use of reversible bonds to attach building blocks to each other so that when the stability of the whole system requires it, some of the building blocks that are part of unselected library members may be recycled to produce the selected receptor. As a consequence, the amplified receptor can be identified by comparing the library distribution in the absence of the guest with that in the presence of the guest [7]. The whole process can be regarded as the adaptation of the dynamic system to the presence of the guest. Ideally, this adaptation includes corrections in the original composition that leads to an increased concentration of the best receptors.

Under such assumption, the dynamic combinatorial approach represents a one-pot process that allows for preparation of a mixture of potential receptors and the screening of such mixture to spot good binders within the library as revealed by their amplification. The desired receptors may then be obtained by isolation of the amplified compound(s) or by resynthesis.

The latter may employ the same templated reversible chemistry used for screening of the library, but now restricting the number and relative concentration of building blocks to those required to prepare the desired product. These biased

dynamic combinatorial libraries (dynamic combinatorial library DCLs) have proved to be a useful synthetic tool for the preparation of complex macrocyclic receptors.

The value of the dynamic combinatorial chemistry (DCC) approach has been demonstrated by some interesting examples that reveal receptors with exceptionally strong affinities for the target [8].

A frequent feature of receptors discovered from DCLs to date is their unexpected structures. This is true even in cases where the library had been designed to include a member that resembles a known host for a given guest and such guest was used as template. The DCC approach has also led to the discovery of rather flexible receptors that are not preorganized. It is very unlikely that these receptors would have been discovered by more traditional design and synthesis strategies [9].

3.2 Experimental Considerations

The general experimental setup for a DCC approach to discover new synthetic receptors includes:

- Design and preparation of the DCL and analysis of its composition at equilibrium.
- Addition of a template molecule and analysis of the library composition at equilibrium.
- Identification of any amplified library members.
- Isolation of the amplified receptor(s).
- Study of recognition properties of the receptor and reuse.

Although each of these steps includes experimental operations that are common to other disciplines, there are some experimental considerations that are particularly important when developing synthetic receptors using a DCC strategy.

3.2.1 Preparation of DCLs

3.2.1.1 Reversible Chemistry

Key to the success of a DCC approach towards the development of synthetic receptors is the reversible chemistry involved in the preparation of the libraries. There is a long list of requirements for the reversible reaction to be used [6]. In order to develop new receptors from DCLs two requirements are crucial. (i) The exchange should be active in a reasonable timescale under appropriate reaction conditions. This determines the time that is necessary for the DCL to respond to the addition of a template molecule. Appropriate reaction conditions are those that are compatible with the structure of building blocks and template, and in particular with the recognition groups involved in the intermolecular interactions that drive

molecular amplification. (ii) It should be possible to stop the exchange so that the amplified library members can be isolated and characterized. Ideally, the changes in the receptor and/or template structures, produced by the experimental operations necessary to stop the exchange, should not affect the binding between the amplified host and the guest used as template.

Noncovalent exchange processes involving hydrogen bonds [10–14] and metal–ligand coordination [15–21] have been used extensively in the preparation of DCLs of potential receptors. The exchange of these noncovalent connections between building blocks proceeds under mild conditions and is usually fast, and several examples of receptor amplification have been reported [22–29]. However, isolation and reuse of the amplified receptors is problematic because of the labile connections between building blocks. To overcome this problem the covalent connection of building blocks through a separate irreversible covalent reaction has been explored by the group of Reinhoudt [30]. This approach should be used carefully since there is always the possibility that covalent capture will affect the binding properties of library members.

It is important to note that, rather than isolating the individual new receptors, DCLs can also be used in their entirety as complex sensors, where a signal is derived from the collective response of all potential binders to the introduction of the guest. In this approach fast exchange is important while robustness of the connections between building blocks is not crucial since isolation is not necessary. Noncovalent assembles will probably play a key role in such systems, which are discussed in detail in Chapter 7. In this chapter we focus on robust covalent reversible connections.

Esters and Related Connections One of the first reactions used to generate DCLs of potential receptors was the alkali catalyzed transesterification [31]. Although the robustness is a good feature of the ester bond when it comes to product stability, it necessitates the use of rather harsh exchange conditions. This reduces the variety of functional groups that can be present in the library members and templates as well as the intermolecular interactions that can be used to drive amplification in those systems.

Some DCLs of potential oligoester receptors have been synthesized under thermodynamic control with 5 mol% of the complex of potassium methoxide and dicyclohexyl [18]crown-6 in refluxing toluene (Figure 3.1) [32–35]. These conditions allow the formation of alkoxide intermediates that remain soluble in the anhydrous solvent mixture at millimolar concentrations yielding the desired thermodynamic mixture of oligoester products within 10 min. Adding alkali metal iodide salts as templates induces modest changes in the product distribution of DCLs of oligocholates. The most significant shift in library composition was induced by sodium iodide [36].

Milder reaction conditions are sufficient for palladium-catalyzed allyl ester exchange. The equilibrium of allyl ester palladium complexes can be reached within hours in the presence of a mild base, palladium catalyst, and moderate heating [37, 38]. These conditions are compatible with the binding between a zinc

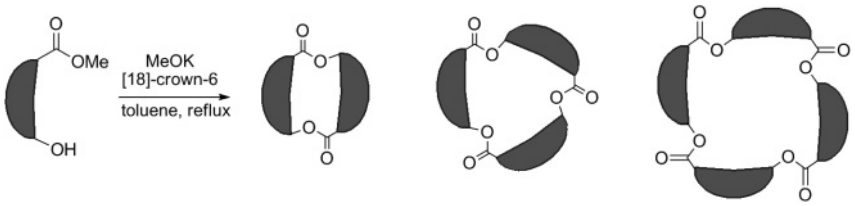
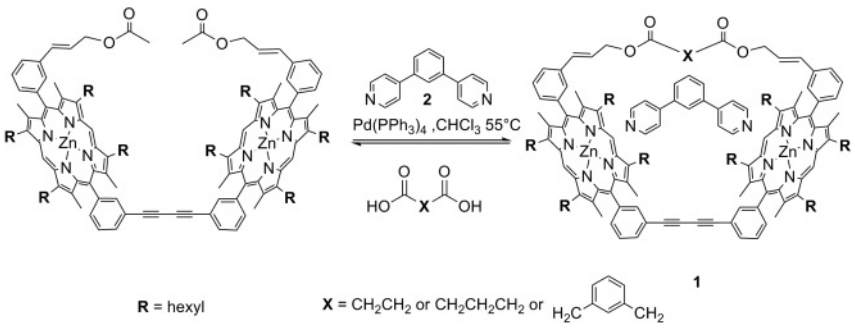


Figure 3.1 Transesterification as a reversible reaction to prepare macrocyclic oligoester DCLs.

porphyrin to pyridine-based templates. Such recognition event was used to induce the formation of a porphyrin receptor **1** by bis-pyridine guest **2** in chloroform (Scheme 3.1) [38].

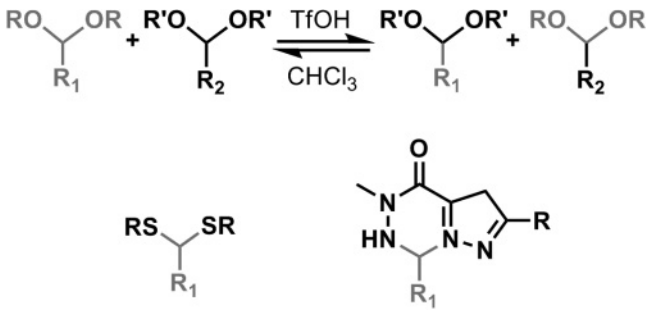


Scheme 3.1 Synthesis under reversible conditions of cyclic porphyrin dimers using palladium-catalyzed allyl transesterification.

Related exchange processes that could be alternatives to the transesterification reaction to produce similar DCLs of receptors under different reactions conditions are the exchange of thioesters and the exchange of amides. Thioesters have been used recently as the reversible bond in the formation of DCLs in water [39–41]. The exchange reaction is usually performed by mixing stoichiometric amounts of thiols and thioesters in aqueous solution (pH 7–9), without the need for any activation procedure.

The thermodynamically controlled amide exchange may be catalyzed by amidoaluminum complexes in organic solvents; however, the relatively high temperatures required (90–120 °C) represent a limitation for its use in DCLs of receptors [42, 43]. Alternatively, enzyme-catalyzed transimination has been reported. Thioester exchange and amide exchange have not yet been explored for the preparation of DCLs of potential receptors.

Acetals and Related Connections Acetal exchange proceeds smoothly in organic solvents in the presence of a strong acid catalyst (Scheme 3.2). The groups of Fuchs



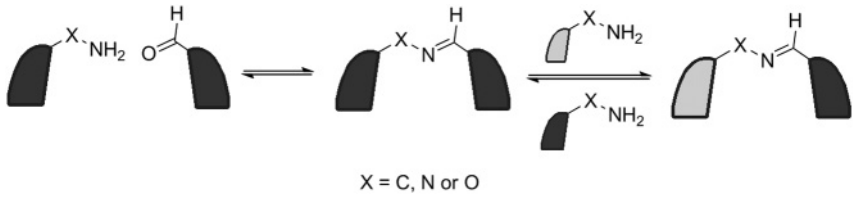
Scheme 3.2 Acetal exchange and related exchange processes explored for the preparation of DCLs.

and Mandolini have used acetal exchange as the reversible reaction for the construction of synthetic receptors [44, 45], and their results demonstrate that the reaction conditions required for the exchange (typically triflic acid (TfOH) in chloroform) are compatible with the recognition of cation templates such as Ag^+ and Cs^+, as well as secondary ammonium cations.

Two exchange processes that are related to acetal exchange have been studied under reversible conditions (Scheme 3.2). The reversible formation of thioacetals has been explored by Hirsch *et al.* to generate equilibrium mixtures of bis- and tris-(thioacetals) [46]. The condensation of benzene hexathiol with *p*-tolualdehyde went to completion in 3 days in refluxing $CHCl_3$ in the presence of the Lewis acid $Zn(OTf)_2$. The products could be isolated and are stable in the absence of catalyst.

Wipf *et al.* reported the reversible cyclocondensation reaction between a pyrazolotriazone and aldehydes [47]. The exchange reaction proceeds in water at pH 4 and equilibrium was reached in 3 days at 40 °C. Exchange could be stopped by raising the pH to 7, allowing for the isolation of stable selected library members. None of these exchange processes have yet been used for the generation of DCLs of receptors.

Imines, Hydrazones, and Oximes Several groups have explored the exchange of the labile C=N bond of imines (Scheme 3.3). The exchange takes place under mild



Scheme 3.3 Imine exchange and related exchange processes explored for the preparation of DCLs.

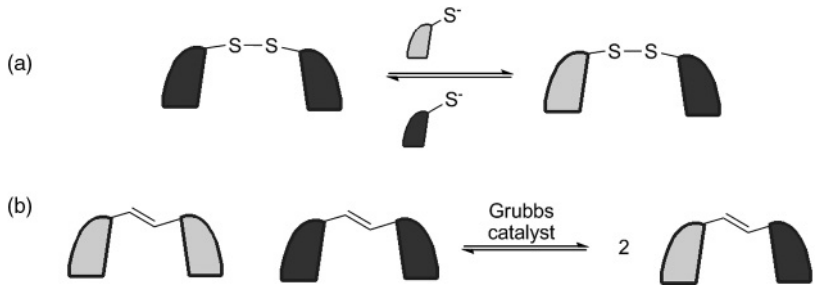
conditions: imine-based DCLs are often prepared in aqueous buffers between pH 5.0 and 8.5, or in organic media in the presence of weak acid such us NH_4PF_6 [48] or oxalic acid [49–51]. Lehn *et al*. have reported the use of lanthanide cations as efficient transimination catalysts in organic medium [52].

Several receptors have been amplified from imine-based DCLs showing that the mild conditions necessary for the exchange are compatible with the coordination of a variety of metal cations [53–56] and with the binding of a hydrophobic guests within hydrophobic cavities [57]. The main disadvantage of imine bonds is their susceptibility to hydrolysis; thus, for most purposes reduction of the products to amines is required. Provided the different library members display the same reactivity toward reduction, the amine product distribution will reflect the imine distribution [58]. However, reduction introduces changes both in geometry and electronics, compromising the binding properties of the isolated receptor.

Hydrazones are structurally related to imines but they have the advantage of being hydrolytically much more stable so no reduction step is required. Hydrazone exchange requires acidic conditions (typically trifluoroacetic acid (TFA) in $CHCl_3$), although for some substrates containing strongly electron-withdrawing groups, exchange is feasible under neutral conditions [59]. The hydrazone exchange reaction was first used for the preparation of DCLs of potential receptors by the Sanders group, who developed a family of amino acid-based bifunctional building blocks featuring one protected aldehyde unit and one acyl hydrazide unit [60]. A series of molecular amplification examples observed with these DCLs has demonstrated the compatibility of the reaction conditions with binding of metal cations and organic cations [61–67].

Eliseev *et al*. have studied the transimination of oximes. Equilibration in water is rapid at high temperatures and/or acidic conditions, whereas ambient temperature and neutral conditions switched off exchange [68]. Later the same group constructed a DCL of oximes in MeOH, but no attempts of guest-induced amplification have been made using this exchange reaction [69, 70].

Disulfides Disulfide exchange was one of the first reactions used for the preparation of receptors under reversible conditions [71] and is perhaps one of the most successful reactions for the development of receptors using DCC (Scheme 3.4a). Thiols readily oxidize to disulfides in aqueous solution upon exposure to air and thiol–disulfide exchange takes place under mild conditions [72–74]. In practice, disulfides DCLs are often prepared in water from thiol building blocks, which are allowed to oxidize to form the desired disulfides. At the same time that oxidation is taking place the reaction passes through a phase where thiol and disulfide coexist, allowing the reversible exchange process to occur through the nucleophilic attack of thiolate anion on the disulfide bond liberating another thiolate [75, 76]. Since the exchange requires thiolate anion, it can be switched off either through protonation of the thiolate nucleophile or by allowing the oxidation process to go to completion. A series of macrocyclic receptors that bind ammonium ions in water [77, 78] has been developed using DCLs inspired by the family of cyclophane



Scheme 3.4 (a) Disulfide exchange. (b) Alkene exchange.

receptors developed by Dougherty *et al*. [79–82]. Disulfide exchange can also take place in organic solvents in the presence of organic bases such as triethylamine or 1,8-diazabicyclo[5.4.0]undec-7-ene [71, 83, 84]. Such reaction conditions are compatible with molecular recognition through metal coordination [84] and through hydrogen bonds [83]. Recently, a small DCL was prepared using disulfide exchange in a two-phase system, which increases the possibilities and the scope of DCC by facilitating the combination of otherwise incompatible building blocks [85].

Alkenes With the development of efficient catalysts, olefin metathesis has become a suitable reaction for the preparation of DCLs of potential receptors (Scheme 3.4b). Grubbs' catalyst is probably the most used transition metal complex for cross-metathesis reactions due to its reasonable stability toward oxygen, water, and minor impurities that can be present in solvents, and because of its compatibility with a wide range of aliphatic and aromatic reactants, such as alkenes carrying epoxide, ester, sulfone, or aromatic aldehyde functions [86]. Template effects have been observed for the ring-closing metathesis macrocyclization in the presence of $LiClO_4$ [87]. However, the catalyst may be deactivated by functional groups that can coordinate to the catalyst (e.g., phosphines, amines, phenols, and nitriles). The functional group tolerance of the first-generation Grubbs catalyst can be enhanced by the substitution of one of the phosphine ligands by a *N*-heterocyclic carbene to produce the second-generation Grubbs catalyst [88]. This catalyst has been shown to be compatible with a pyridine-based template in a library prepared from alkene-functionalized porphyrin building blocks [89]. One practical difficulty in the use of alkene metathesis in DCC is the limited lifetime of the available catalyst, preventing the complete equilibration of complex mixtures and the separation of the catalyst to stop the exchange process [6].

Alkyne metathesis has not been explored extensively for the preparation of DCLs [90]; however, the development of new complexes that allow alkyne metathesis of highly functionalized substrates under mild conditions should stimulate a more widespread use of this reaction [6, 91].

3.2.1.2 **Building Block Design**

The exact structure of the building blocks used in the preparation of a DCL will depend on the desired characteristics for the receptor that is needed. Such characteristics will dictate potential recognition groups to be included, which will need to be compatible with the reversible chemistry and reaction conditions used. Most of the receptors discovered to date using the DCC approach have been built from building blocks inspired by previously known receptors.

One important factor to consider is building block solubility. This aspect becomes particularly important since any insoluble material present in the DCL could act as a thermodynamic trap or as a kinetic trap if the dissolution rate is slow. Although good building block solubility does not assure good library solubility, it is still worth taking solubility into account to prevent problems. A recently explored alternative is the use of two-phase systems [85]. This may enable the use of a wider range of guests and building blocks, overcoming solubility problems and thereby increasing the diversity of libraries that can be generated and explored.

Building blocks can be designed to be predisposed for macrocyclization or polymerization. Rigid building blocks wherein the number of available conformations is reduced can decrease the entropic penalty for cyclization. In contrast, steric constraints can prevent cyclization and thus drive the assembly toward linear species [49–51, 92].

An interesting relationship between building block structure and library diversity was revealed by the early work of the Sanders group on DCLs generated through transesterification. Transesterification under thermodynamic control of building blocks **3** and **4** (Figure 3.2) produced all four possible cyclic trimers in a statistical 1 : 3 : 3 : 1 ratio [33]. Hardly any dimers or tetramers were formed. In contrast, cyclizations under kinetic control produced a mixture of trimers, tetramers, and higher oligomers.

The authors suggested that the marked preference for the formation of cyclic trimer under thermodynamic conditions was associated with the building block backbone rigidity. Increasing flexibility of the possible macrocycles replacing one of the building blocks by the extended building block **5** allows the formation of cyclic dimers, trimers, and tetramers, giving a total of 11 detectable species as a

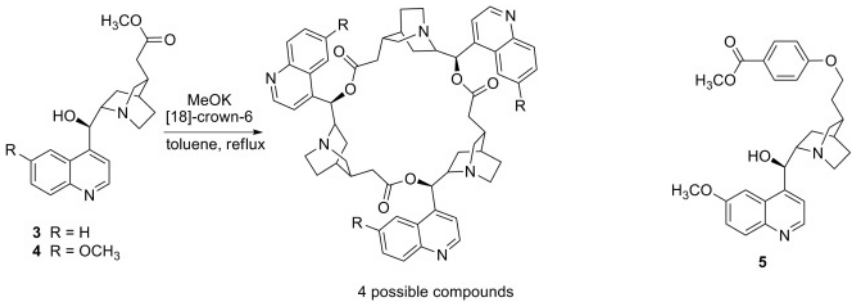


Figure 3.2 Macrocyclization under thermodynamic control.

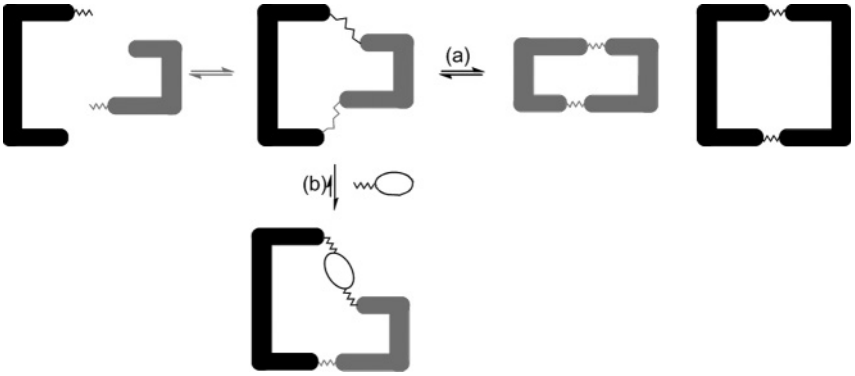


Figure 3.3 (a) Self-sorting of rigid building blocks of different bite size. (b) Addition of small and flexible building blocks gives access to mixed species.

mixture of hetero-oligomers and homo-oligomers [35, 93]. Library diversity was further increased by incorporating additional building blocks with increased flexibility, reduced dimensions, and both [92, 94]. The main conclusion from these studies was that rigid concave building blocks tend to self-sort when they have different "bite sizes" (Figure 3.3a). Ring strain destabilizes any formed mixed species, dramatically decreasing their concentration. In this situation, library diversity can still be increased by the addition of small and flexible building blocks that can act as diversifiers bridging rigid building blocks of different bite sizes (Figure 3.3b).

The number and the diversity of building blocks that have been successfully used to date to prepare DCLs of potential receptors are surprisingly low. There are almost more different receptors isolated from DCLs than building blocks used to produce those DCLs.

This may be considered as an indication of the efficiency of the approach; however, it may also be an indication of how demanding the approach is in terms of building block design, solubility, and so on.

3.2.1.3 **Building Block Concentration**

The total building block concentration in DCLs of potential receptors is a key factor because it influences solubility, rates of the reversible reactions, and, in the case of libraries of oligomers, it can also affect the relative concentration of library members.

Since chain extension is a second-order reaction, whereas cyclization is a first-order process, the dependence of each process on the concentration is different. At low concentrations cyclization is faster than chain extension. There is a critical concentration below which the equilibrium composition of the system consists entirely of small macrocycles, the distribution of which depends on the building block concentration [95]. Under such conditions the product distribution is also modulated by ring strain, which may destabilize individual macrocycles [96], and

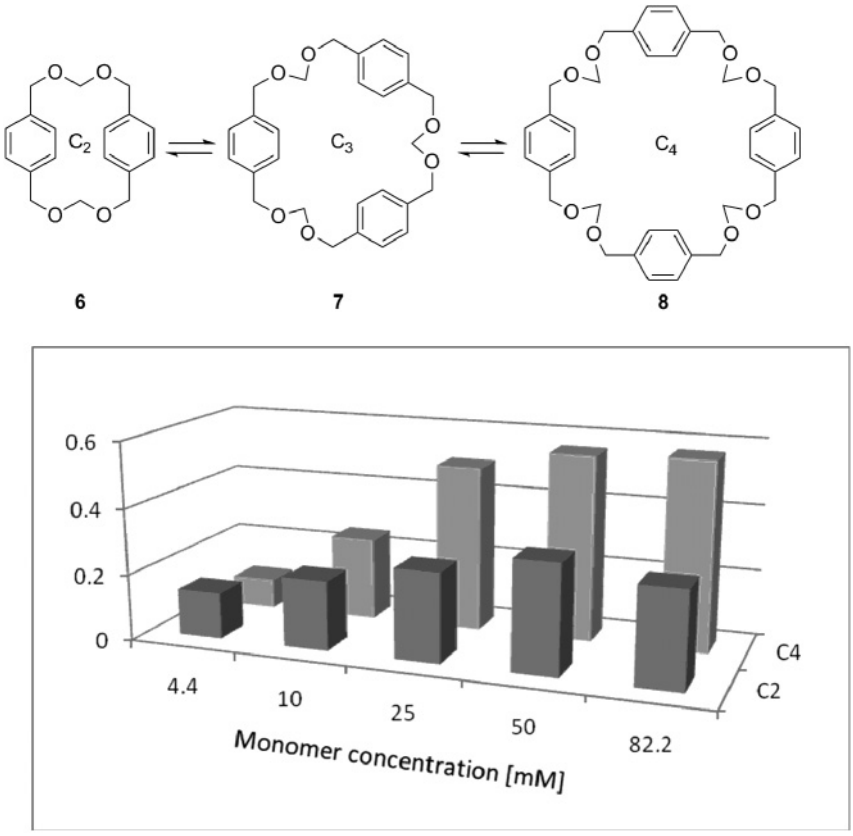


Figure 3.4 Influence of monomer concentration on the relative concentration of cyclophanes C_2 and C_4.

intramolecular interactions, which may stabilize selected macrocycles [83]. Intermolecular recognition between library members may also affect stability of the involved molecules [97, 98].

The influence of building block concentration on the equilibrium composition of a simple dynamic system of cyclophanes was illustrated by Mandolini *et al.* [45]. High dilution favors the lower cycles at the expense of high-molecular-weight materials, as shown by a comparison of the equilibrium concentrations and yields of oligomers C_2 (**6**) and C_4 (**8**) (Figure 3.4).

3.2.1.4 **Assessing Equilibrium**

Molecular amplification is the increase in the total concentration of a library member at equilibrium, produced by addition of a template molecule. As a consequence, to detect real amplifications it is important to be certain that the equilibrium has been reached.

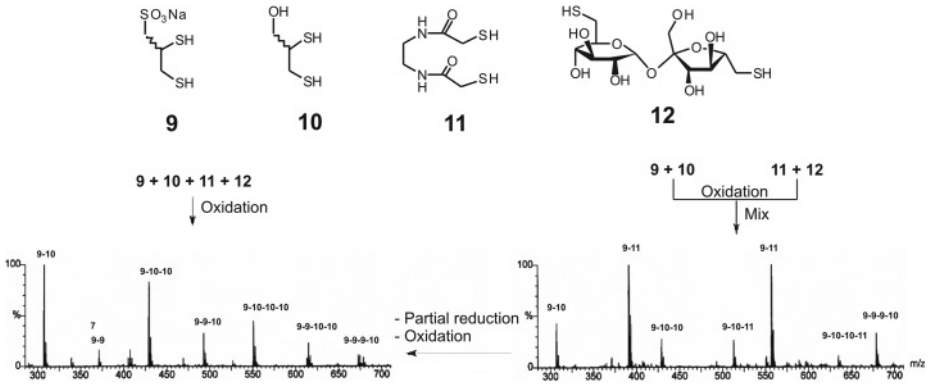


Figure 3.5 Preparation of a DCL from different starting points to demonstrate that equilibrium has been reached.

The most common way to demonstrate that an equilibrium has been established is to obtain the same product distribution preparing the library from different starting points. It is rather improbable (but not strictly impossible) that the same product distribution is obtained if the mixtures have not reached equilibrium. For example, the product distribution in a library made by mixing building blocks **9**–**12** at the start of the experiment should be the same as that obtained upon mixing two sublibraries that contain a subset of the building blocks **9** and **10**, and **11** and **12**, respectively (Figure 3.5) [76].

An equivalent experiment for DCLs prepared from a single building block is to isolate one library member and demonstrate that upon exposing it to the exchange conditions, it regenerates a library identical to that from which it was isolated [45, 62, 63].

It is also possible to move the library composition away from the equilibrium composition by altering an experimental variable such as temperature [99–101], concentration [45], pressure, or solvent and check that the library returns to the original composition after reverting to the initial experimental conditions. For example, the relative concentration at equilibrium of cyclophanes C_2–C_4 (Figure 3.4) can be shifted toward higher oligomers by increasing the concentration of building blocks from 25 to 50 mM. Further dilution of the mixture shifts the product distribution back to the initial concentrations [45].

Alternatively, a shift in the original equilibrium composition can be induced by complexation of some library members by a template and then disrupting the interaction between the template and the library members [67].

3.2.2 Templating of the Library

3.2.2.1 Ensuring Meaningful Amplifications

The power of DCC for the discovery of new synthetic receptors depends in part on the original conception that the best hosts in a DCL would be most amplified upon

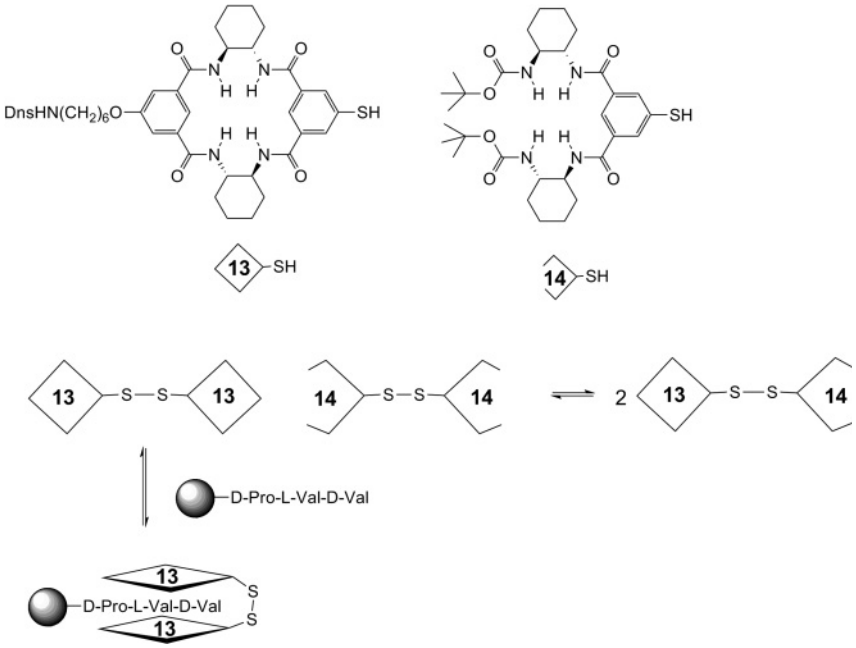
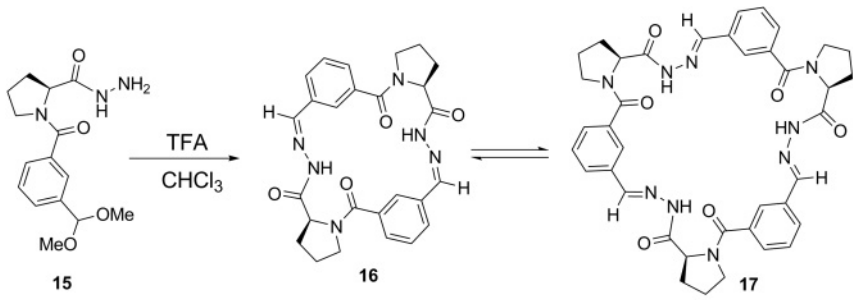


Figure 3.6 Amplification of the best receptor induced by a solid supported template.

addition of a guest. Early results on amplification of receptors from simple model systems supported that idea. In the pioneering work by Hioki and Still, amplification of the best receptor was observed in a simple dynamic system [71]. Two organic monothiols **13** and **14** were connected by means of a disulfide bridge to generate three potential receptors for small peptides (Figure 3.6). Under conditions that allow disulfide exchange approximately equal amounts of the species **13–13**, **13–14**, and **14–14** were observed. Addition of an excess of a polymer-supported tripeptide as a target resulted in a pronounced re-equilibration with amplification of the homodimers **13–13** and **14–14**. Since an immobilized target was employed, the authors were also able to separate the bound and nonbound receptors. This allowed the isolation the high-affinity receptor **13–13** in 97.5% purity. Although binding constants were not measured accurately, it was clear from qualitative data that the best receptor within the system was amplified selectively.

Based on the research by Kubik and Goddard on L-proline-derived cyclic peptides that bind ammonium salts [102], Sanders *et al.* prepared a small DCL derived from building block **15**. Exposure of the DCL to the cationic guest acetylcholine resulted in the amplification of the expected trimeric receptor **17** (Figure 3.7) [66]. Although the trimer macrocycle was expected to be the best receptor as it resembles the compound in Kubik's original report, binding affinities of trimer and dimer **16** for the template were not compared because of solubility problems.

Despite these and various other examples of molecular amplification of receptors described during the last 12 years, the "amplification of the fittest" concept



Template	Yield of dimer	Yield of trimer **17**
---------	90%	10%
Acetylcholine	12%	88%

Figure 3.7 Amplification of the expected receptor induced by acetylcholine.

was never unquestionably proved. On the other hand, investigations by Severin *et al.* [19, 103, 104] and Sanders *et al.* [105–107] have demonstrated that "amplification of the fittest" is not always the case. It is now well established that the correlation between binding and amplification can break down when two or more library members: (i) bind a template that is in excess and (ii) include in their structures different numbers of one building block that is in short supply (see Chapter 2, Section 2.3 for more details). Under such conditions those library members constituted by a smaller number of the scarce building block have a competitive advantage over those members that contain larger numbers of that particular building block. Consequently, smaller oligomers tend to have an advantage over higher oligomers and hetero-oligomers over homo-oligomers.

In the presence of a modest amount of template the number of possible binding events is limited by the number of template molecules. Thus, binding events that produce the highest energy gain are preferred and the best host will be amplified selectively. Thus, if the objective is to ensure a good correlation between binding and amplification, the concentration of the template in the DCL needs to be well below the concentration of the competing receptors. For more detailed guidance, see Chapter 2, Section 2.3.

A clear case where competing hetero-oligomers suppress amplification of homo-oligomers was demonstrated experimentally by Saur and Severin [19]. The authors studied lithium binding by ruthenium and iridium complexes **18–21**. Binding constants for the complexes **18** and **21** are 4.4×10^3 and $1.1 \times 10^{-1}\,M^{-1}$ respectively. Based on this data it was expected that the mixed aggregates **19** and **20** (containing one and two Cp*Ir fragments respectively) show a lower Li^+ affinity than the homotrimer **18**. Four member DCLs were produced by mixing **18** and **21** (Figure 3.8). Lithium sulfate was added to the DCLs, and the binding was studied by ^{7}Li-nuclear magnetic resonance (NMR) spectroscopy. When a small amount of lithium was added, it was found to bind preferentially to **18**. When successively larger

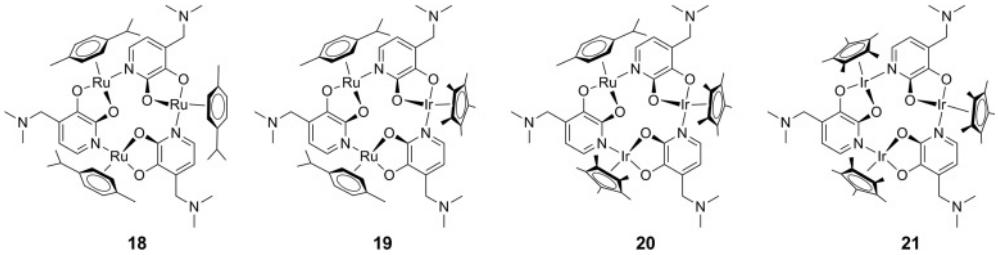


Figure 3.8 Members of the DCL used to demonstrate the effect of template concentration on amplification selectivity.

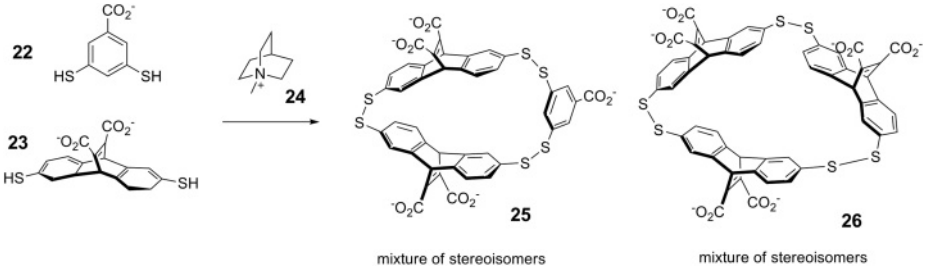


Figure 3.9 DCL used to demonstrate the effect of template concentration on equilibrium composition.

excesses of lithium were added, up to 40 equiv., the lithium was increasingly found to be bound to **19**, demonstrating that adding excessive amounts of template can indeed cause DCLs to select suboptimal library members.

Otto and Sanders reported later a quantitative experimental demonstration of the breakdown in the correlation between affinity and amplification in a DCL prepared from disulfides **22** and **23** [107]. Introducing template **24** into the library induced amplification of receptors **25** and **26** (Figure 3.9). Binding constants for the complexes of **24** with **25** and **26** are 5.0×10^4 and $7.9 \times 10^4\,M^{-1}$, respectively. At low template concentration, the amplification of **26** was more efficient than for **25**, reflecting the fact that **26** is a slightly better binder. An increase in the template concentration enhanced the amplifications of both receptors to different extents and beyond a certain concentration the weaker binding heterotrimer **25** becomes the most amplified compound. It is important to note that according to the experimental results and the simulations carried out to date, such breakdown in correlation is produced when two receptors with similar binding constant compete. When the binding constant for the best receptor is around 20-fold higher than for the others, the best receptor is amplified even in the presence of modest excess of template.

To date there is no experimental example of a system where a smaller weaker binding oligomer outcompetes a stronger binding higher oligomer. It appears that no system has yet been studied where binding constants for oligomers of different

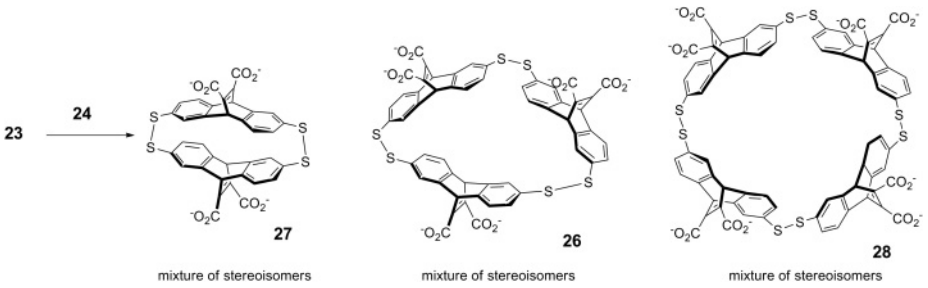


Figure 3.10 DCL used to analyze the effect of template concentration on equilibrium composition.

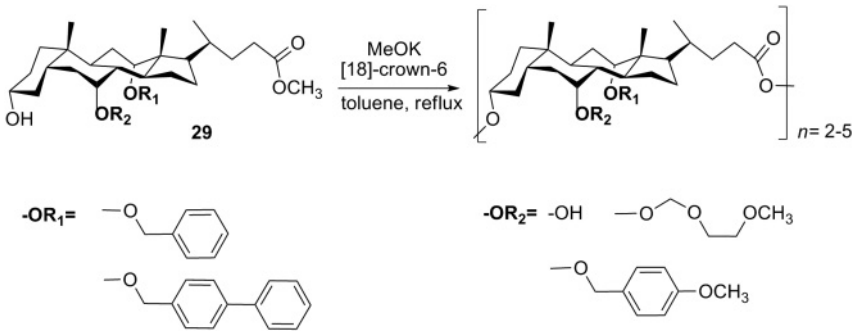
sizes are similar. The only case where macrocycles of different sizes have been analyzed in detail is the library prepared from dithiol **23** to form a mixture of dimer **27**, trimer **26**, and tetramer **28**, where the binding constant for the trimer is $7.9 \times 10^4\,M^{-1}$ and the tetramer is $1.3 \times 10^6\,M^{-1}$ (Figure 3.10). Calculations suggest that in such system a breakdown in the correlation between binding constant and amplification would occur for template concentrations higher to 500 mM [107].

A different way to guarantee that no building block is scarce, and thus maintain a good correlation between binding and amplification, is to work under experimental conditions where the free building blocks (and not the potential receptors) are the dominant species in solution (virtual libraries). Similarly, DCLs can be designed that contain a binding sublibrary coexisting with a nonbinding sublibrary that can act as a building block reservoir [104, 108]. Alternatively, it can be useful to design the DCL in a way that all the library members contain one common building block, which is used in substoichiometric amounts relative to the other building blocks [104]. The utility of each of these alternative ways of ensuring a high enough concentration of critical building blocks will depend on the library design.

In summary, to ensure selective amplification of the best receptor it is important that the concentration of template molecules does not demand more molecules of the preferred building blocks than the system can provide. This can be achieved by using low concentration of templates or by keeping an *in situ* building block feedstock. This feedstock can be free building blocks (virtual libraries) or building blocks as part of nonbinding assemblies (sublibraries).

A second original concept from the early years of DCC was that exposure of a given DCL of receptors to a variety of guests that amplify the same receptor would produce responses from the libraries that correlate with the binding of those guests to the amplified receptor. Again, some early results on amplification of receptors from relatively simple DCLs supported that idea.

The first macrocyclic receptor generated using DCC was prepared by Sanders *et al*. using transesterification. Initial studies used building blocks based on cholates functionalized with a methyl ester and a hydroxyl group (Scheme 3.5) [31]. Transesterification of compound **29** was catalyzed by potassium methoxide/dicy-



Scheme 3.5 Generation of a DCL of oligocholates by transesterification.

clohexyl [18]crown-6 in refluxing toluene to produce an equilibrium mixture of cyclocholate macrocycles including trimers, tetramers, and pentamers. Proof of the thermodynamic nature of the system was obtained by purifying the trimers and tetramers and allowing them to re-equilibrate, regenerating the original product distribution.

Adding alkali metal iodide salts as templates induced modest changes in product distribution [36]. The most significant shift in library composition was induced by sodium iodide, which doubled the concentration of tetramer and pentamer in the reaction mixture. Binding constants were not measured; however, the relative affinities of the amplified cyclophane macrocycles for the different metal alkali were studied by electrospray ionization-mass spectrometry (electrospray ionization ESI-mass spectrometry MS), finding good agreement with the main amplification observed. The yield of cyclic tetramer was increased by the addition of sodium in the cases of the "MEM" and di(*p*-methoxybenzyl) monomers, and these were found to bind sodium preferentially in the ESI-MS study [109].

In a different analysis, when the simple dynamic system depicted in Figure 3.7 was exposed to a series of ammonium cations as templates, different amplification factors (AFs) were observed for the cyclic trimer **17**. Binding constants for the complexes between the selected trimer and the different templates correlate well with the effects these templates have on the product distribution [63]. Unfortunately, the binding constants that are driving amplification cannot be measured for this system since the necessary acidity to reproduce exact conditions for binding will catalyze the degradation of receptor **17** to produce cyclic dimer in an extent that will depend on the amount of template present.

The assumption that the differences in amplification of a given receptor induced by different guests correlates with the binding preferences of this receptor does not take into account that amplification is the consequence of a binding process under competition conditions. The response from a library will result from the binding profile of a given template with every library member – two templates that bind to the same extent to one particular library member A and to a different extent to other library member(s) may induce different amplifications of A.

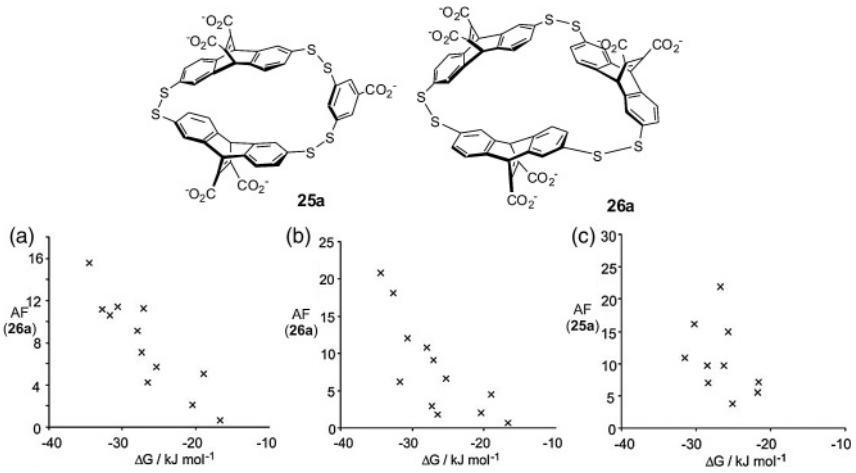


Figure 3.11 AFs as a function of the host–guest Gibbs binding energy of (a) receptor **26a** in DCLs prepared from building block **23**, and (b) receptor **26a** and (c) receptor **25a** in DCLs prepared from building blocks **22** and **23**.

A recent study clearly shows that the correlation between host–guest binding and the AF for a given host across a broad range of guests is not necessarily good [110]. The correlation between the AF for one particular host produced by a series of guests and the host–guest binding energy may be different for different DCLs (DCLs that include different compositions of competing receptors).

Figure 3.11a shows the experimental results for the amplification of receptor **26a** induced by a series of ammonium cations as a function of the host–guest binding energy in DCLs prepared from building block **23** (Figure 3.9). Figure 3.11b shows the corresponding data in the DCLs prepared from building blocks **22** and **23**. Amplification of host **26a** in the DCLs prepared from only one building block appears to correlate fairly well with the host–guest binding energy. However, amplification of the same host in the more complex DCLs prepared from two building blocks shows an increased amount of scatter.

In addition, the correlation between binding energy and the AFs for two different hosts induced by the same set of templates in the same DCL can also be significantly different. As explained previously, Figure 3.11b shows the experimental results for the amplification of receptor **26a** induced by the series of guests as a function of the host–guest binding energy in DCLs prepared from building blocks **22** and **23**. Figure 3.11c shows the corresponding data for host **25a**. Although amplification of the host **26a** shows some correlation, the amplification of host **26a** in the same libraries is hardly correlated at all to the guest binding energy.

This lack of good correlation is mainly the result of each guest having its own unique pattern of affinities for the various competing receptors in the mixture. In conclusion, assessing the absolute binding affinity of a certain guest for a certain host from the extent to which the guest induces the amplification of this host is

not necessarily reliable in a quantitative sense, but is in most cases still qualitatively useful. Most importantly, when strong amplifications or strongly selective amplifications are observed, the selected receptors are indeed strong or highly selective binders, even though the reverse is not necessarily true [110].

3.2.2.2 Identification of Amplified Compounds

The main objective of library analysis is to detect changes in concentration of individual library members. This will allow determining if the library has reached equilibrium and if molecular amplification is being produced by addition of the template.

In some libraries, changes in concentration can be detected by direct analysis of the DCLs through NMR or MS; however, more complex libraries can require the previous separation through chromatographic techniques.

^{1}H-NMR has been used mainly for the analysis of small DCLs where each library member has one distinctive signal that is used for comparing concentrations [44, 45, 54, 111, 112]. Direct NMR analysis of library changes usually has the advantage that the interaction between guest and host is not disrupted so it is possible to observe effects produced by the guest on the library (i.e., complexation-induced changes in chemical shift) in addition to the changes in concentration [19, 113].

MS, in particular with ESI, has been an important tool for the analysis of DCLs of receptors. Its utility has been shown by analyzing libraries of potential receptors based on imines [56], hydrazones [60, 114], disulfides [76], acetals [44], and esters [35]. When the libraries are more complex, high resolution may be necessary. Furthermore, MS-MS enables analysis in cases involving sequence isomers [114] and regioisomers [70].

ESI-MS has also proved to be a powerful technique for the study of intermolecular processes. This characteristic has allowed the use of ESI-MS to detect intermolecular interactions that drive molecular amplification in DCLs. This has been used mainly for DCLs templated with cation guests like acetylcholine or cinchona alkaloids [61, 63, 66].

High-performance liquid chromatography (HPLC) has been one of the main tools for library analysis of small DCLs of receptors. The combination of HPLC and MS is probably the most powerful analytical technique for the analysis of changes in composition of libraries of covalent receptors. It allows reliable comparison of composition to detect amplification and, at the same time, it potentially allows the identification of the receptor (provided all receptors have different molecular weights). Liquid chromatography can also be coupled to polarimetry to detect enantioselective amplifications from racemic libraries of receptors. Since achiral compounds and racemates are polarimetry silent; only receptors enriched in one enantiomer (i.e., from a diastereoselective host–guest complex) will give a signal in a laser polarimeter detector [115].

Independent of the analytical methodology used, detecting good binders within a DCL requires the comparison of AFs for the various library members, rather than absolute differences in concentration (these two metrics are equivalent in libraries where all library members have similar concentration in the absence of

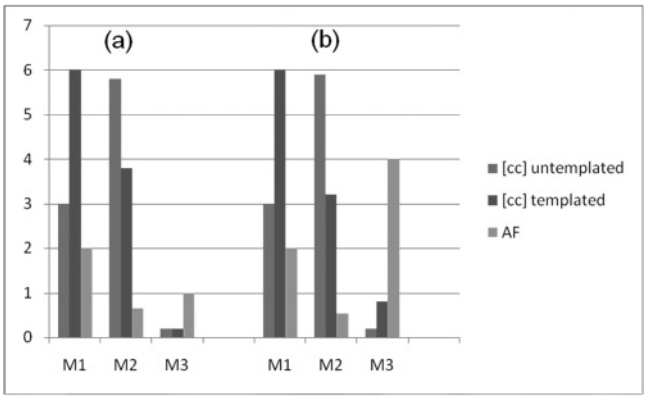


Figure 3.12 AFs and concentration for each library member in two different DCLs. (a) M1, the library member with the highest AF, is the same as the library member with highest difference in concentration. (b) M3, the library member with the highest AF, is different from the library member with highest difference in concentration (M1).

template, but this is rarely the case). Significant amplification of certain compounds that were already present in the untemplated DCL in relatively high concentration has a tendency to capture the attention, at the expense of library members with low abundances but with better AFs. Such could be the case in the hypothetical three-membered DCL depicted in Figure 3.12. In Figure 3.12a, library member M1 is the most amplified (AF = 2). In this particular case, AF is coincident with the absolute change in concentration which is also highest for M1. In Figure 3.12b it may not be immediately obvious from the raw data that although the concentration difference due to templation for M1 is higher than for M3, the AF for M3 is twice the AF for M1.

3.2.3 Isolation of Amplified Receptors

The final step in the discovery of synthetic receptors through DCC is the isolation of the receptor in order to study its binding properties and to use it for some application. This of course demands that the receptor is stable in the absence of the template so the exchange process needs to be stopped as discussed in Section 3.2.1.1. Apart from the traditional separation techniques, there are some strategies that can be applied to facilitate purification.

3.2.3.1 Biased Libraries

Having identified the selected hosts, the library composition can be tuned to favor their formation. A second generation of biased libraries can be prepared that contain only those building blocks that were selected by the guest in the appropriate ratio. In this way the guest-induced amplification is used as a preparative tool for the high yield synthesis of the selected receptors.

For example, receptors **25**, **26**, and **28** (Figures 3.9 and 3.10) were all isolated from biased DCLs that contained only those building blocks that are part of the receptors. While receptor **25** constituted only 5–10% of the library material at the screening stage, in a biased library made from building blocks **22** and **23** in a 1 : 2 ratio, the desired host is formed in 60–65% yield. This strategy worked even better for receptors **26** and **28** (Figure 3.10). In this way DCC can be used as a practical method not only for screening, but also for receptor synthesis.

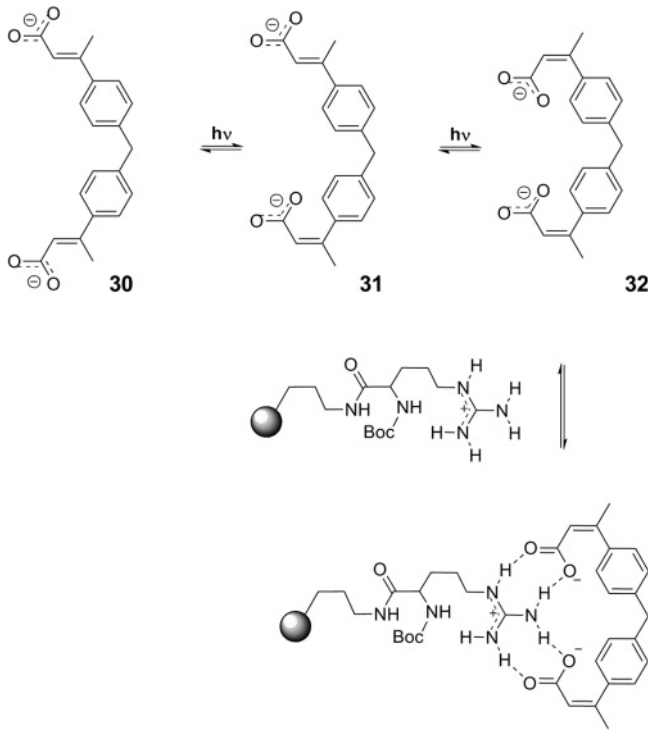
3.2.3.2 **Solid Supported Templates**

The use of templates supported on polymer beads can become a very useful tool for the development of synthetic receptors through the dynamic combinatorial approach. Since the template is attached to a solid support, the untemplated species will eventually become washed from the solid support by filtration leaving only strong binders appended to the beads. The species attached to the solid phase may be washed off under different solvent conditions that disrupt the template–receptor interaction and subsequently identified. The development of selection techniques using guests appended to polymer beads is pivotal to the management and success of highly diverse DCLs of receptors.

The first receptor for an organic cation using DCC was amplified and isolated by Eliseev and Nelen using a solid supported template [116, 117]. A simple dynamic system was designed that included three photoisomerizable dicarboxylate receptors **30–32**. These can interconvert through *cis–trans* isomerization by irradiation and have different affinities for the guanidinium guest. The experimental setup included an isomerization chamber where the library was irradiated and a selection chamber where the library was pumped through an affinity column bearing immobilized arginine. Various sorbents were explored as possible supports for the guanidinium derivatives. As expected, the nonspecific adsorption varied significantly along the series. Silica gel was identified as the preferred sorbent because of its low nonspecific adsorption, low flow resistance, and high specific surface. Multiple rounds of equilibration and selection resulted in efficient amplification and isolation of the best receptor **32** (Scheme 3.6).

The experimental method introduced by Eliseev involved physical separation of the equilibration of receptors and the binding events with the guest. Later, Hioki and Still described one of the first examples of the use of molecular recognition of an immobilized template to shift a receptor equilibrium and isolate the best receptor (Section 3.2.2.1, Figure 3.6) [71].The best-binding receptor was conveniently isolated from the substrate-carrying beads by washing them first with $CHCl_3$ to eliminate unbound material and then with dimethyl formamide to extract any peptide-binding compounds.

In a later study the group of Sanders demonstrated that amplification of the trimeric receptor **17** (Section 3.2.2.1, Figure 3.7) previously induced by benzyltrimethyl ammonium iodide, could be also achieved with an immobilized version of this template. Since the template induces an increased yield and keeps the product bound and separated from library members with lower affinity, the procedure represents a combination of template synthesis and affinity chromatography in one

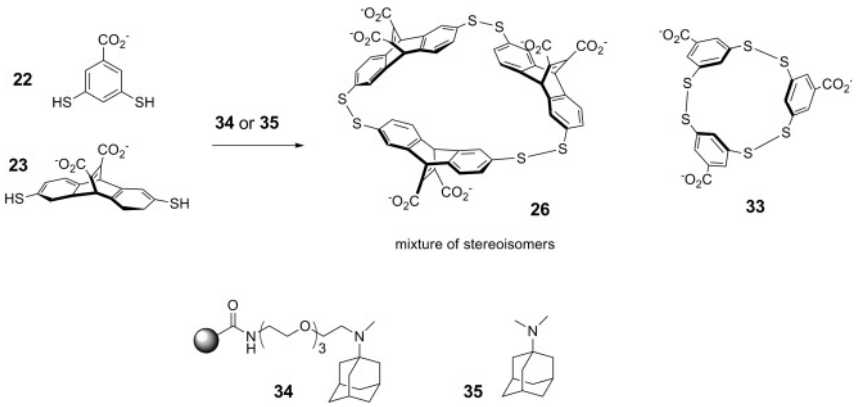


Scheme 3.6 Amplification and isolation of the best receptor from a small DCL using a solid supported template.

single step. Even if template efficiency is modest and many other library members are present the one-pot synthesis, the amplification and isolation procedure was efficient [64]. Owing to the preference of the resin-bound guest for its best receptor, this species could be separated from other library members by filtration.

More recently Besenius *et al.* described a one-pot amplification and isolation of the receptors **26** and **33** using the immobilized analog **34** of guest **35** (Scheme 3.7) without simultaneously binding of the higher, more charged oligomers that would occur if an ion-exchange-type mechanism was taking place [118].

When the template **34** was prepared through polymerization with dimethylacrylamide the immobilized template produced a similar amplification than observed in homogeneous phase with **35**; however, when acrylamide was used no amplification was observed at all with the resin. The authors speculate that extensive hydrogen bonding between amide hydrogens and carbonyl groups within the acrylamide based resin may inhibit local access to the immobilized template or that the dimethylacrylamide-based resin may provide a more favorable hydrophobic microenvironment around the template. The work clearly illustrates the importance of the polymer backbone on host–guest chemistry under heterogeneous conditions – an aspect that has not received significant consideration.



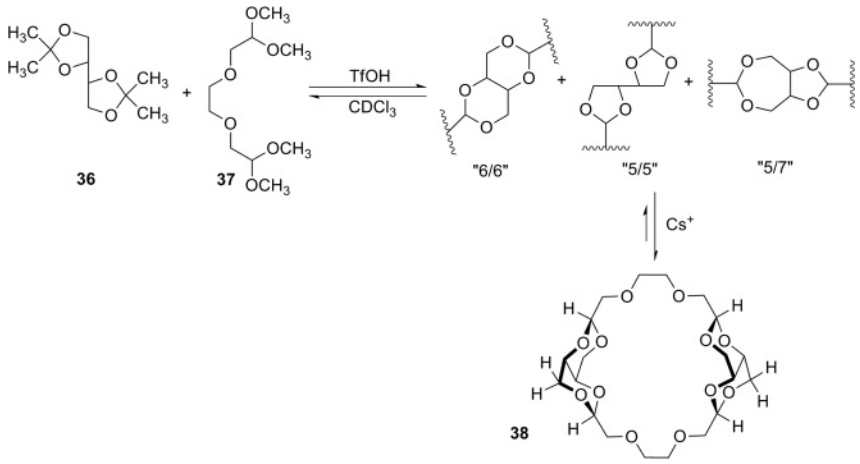
Scheme 3.7 Amplification and isolation of receptors from a DCL using a solid supported template.

3.3
Selected Examples

3.3.1
Synthetic Receptors for Metal Cations

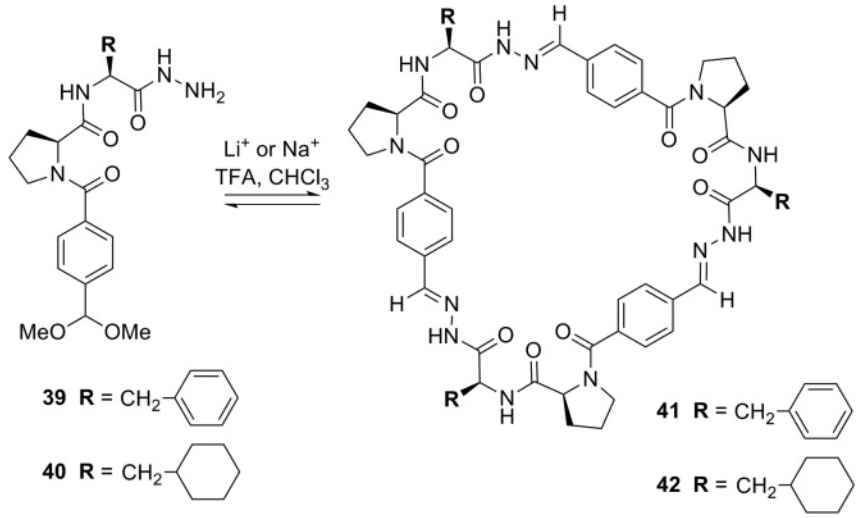
The application of metal ions in order to favor the formation of macrocycles has a long history in chemistry (see also Chapter 1). The complexation by a metal ion may be used to direct the synthesis away from insoluble polymers to soluble cyclic adducts and shift otherwise unfavorable equilibria toward bond formation [119]. It is not surprising that the initial attempts to prepare macrocyclic receptors using DCC involved metal cations as templates. The first series of macrocyclic receptors generated using DCC were prepared by Sanders *et al*. using transesterification. Initial studies used building blocks based upon cholates functionalized with a methyl ester and a hydroxyl group (Scheme 3.5) [31, 36, 109].

Fuchs, *et al*. produced a dynamic library of chiral crown ethers based on tetraoxadecalin cores using acetal exchange [44]. As shown in Scheme 3.8, the transacetalation of the diacetonide of D-threitol (**36**) with diacetal **37** in the presence of TfOH in deuterated chloroform ($CDCl_3$) leads to a complex mixture of cyclic and linear products. These are formed by linking units derived from **36** and **37** through two fused six-membered diacetal rings ("6/6"), two fused five-membered diacetal rings ("5/5"), and fused five- and seven-membered rings ("5/7"). When $CsPF_6$ was added to the reaction mixture, the formation of [2 + 2] macrocycles was strongly favored and more than 95% of these macrocycles contained two fused six-membered diacetal rings (**38**). The same experiment carried out in MeCN led to a poorer amplification. No binding constants were reported for this host–guest system. Receptor **38** was also prepared using kinetically controlled chemistry to compare the preparative efficiency of the dynamic approach with a kinetic one. The isolated overall yield using the kinetic approach was considerably lower.



Scheme 3.8 Amplification of a crown-ether-like receptor for Cs^+.

Three receptors for alkali cations have been prepared by the Sanders group using hydrazone-exchange-based DCLs. The group synthesized building blocks based on dipeptides (Scheme 3.9) that are equipped with a hydrazide group on one end and an aldehyde (protected as dimethyl acetal) on the other. Based on building blocks **39** or **40** small DCLs of macrocyclic pseudopeptides were generated [62, 65]. The DCL compositions were affected by the presence of NaI and LiI. In both libraries, the amplification of cyclic trimers **41** and **42** was efficient, leading to essentially quantitative yields of these 42-membered macrocycles. Binding constants for

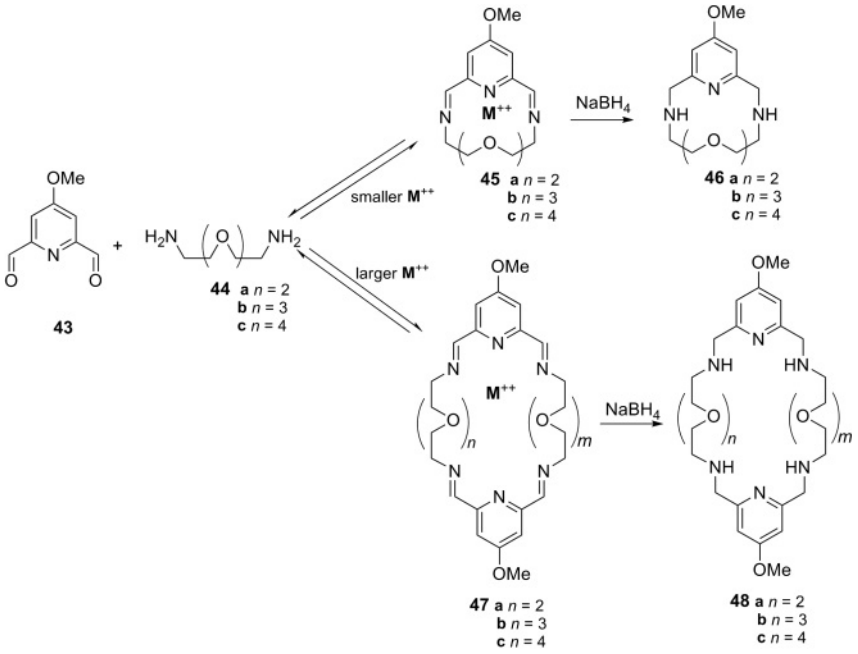


Scheme 3.9 Building blocks and macrocyclic receptors for Li^+ and Na^+.

receptors were determined using ^{1}H-NMR and isothermal titration calorimetry. Both receptors bind Na^+ stronger than Li^+ – a result that correlates well with the templating experiments. Both receptors are highly flexible molecules that undergo substantial conformational rearrangement upon binding to guests. Such induced-fit type receptors are difficult to access using a design approach.

A series of related building blocks have been prepared based on this design [114, 115, 120–122]. The diversity of the DCLs prepared from compounds analogous to **39** suggest that changes in the structure such us the order, the side chain, or the chirality of one or more of the amino acids constituents of the building block can have significant effects on the library behavior that are often difficult to predict [120]. Despite the relatively large number of dynamic libraries prepared based on this building block design, no other receptors for metal cations have been reported to date.

Storm and Lüning have used imine exchange to investigate the possibility of controlling the selection of different macrocyclic receptors from the same DCL by the use of different templates [53, 54]. When dialdehyde **43** (Scheme 3.10) was mixed with diamines of varying chain length **44a–c** in water, a mixture of linear and cyclic oligomers was produced. In the absence of template the only macrocycle that was detected after reduction of the imines to amines with excess of sodium borohydride was **46b** in 9% yield. Addition of different metal ion templates directed the mixture to the formation of different [1 + 1] or [2 + 2] macrocyclic products

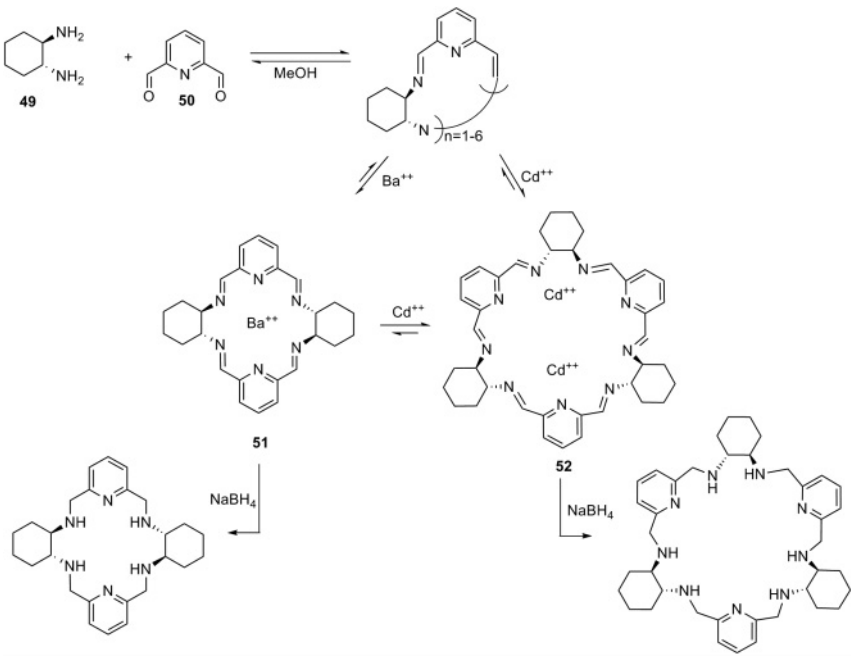


Scheme 3.10 Building blocks and receptors for various metal cations amplified from imine DCLs.

(**45** and **47**) depending on the template size. Addition of Mg^{2+} to a mixture of **43** and **44** resulted in the amplification of diimine macrocycle **45a** up to 86% (based on the yield of its reduced counterpart **46a**). Larger template ions like Ca^{2+} and Sr^{2+} reduced the yields of **45a** (60 and 45%, respectively) producing a larger [2 + 2] macrocycle **47a**, in 22 and 34% yield, respectively, which is not detected in Mg^{2+} templated libraries.

The use of the three template metal ions and the three diamines at the same time led to the formation of several target macrocycles. For instance, by simultaneously adding Mg^{2+}, Ca^{2+}, and Sr^{2+}, compounds **45a–c** were amplified (**45a** = 60–84%, **45b** = 22%, and **45c** = 27–62%).

Gotor *et al*. have used imine exchange for the metal-ion-templated synthesis of *N*-macroheterocycles from (*R*,*R*)-cyclohexane-1,2-diamine **49** and pyridine-1,2-dicarboxaldehyde **50**. In the absence of template a mixture of macrocycles ranging from the [2 + 2] to [6 + 6] products is observed in MeOH. However, the main product depended on the added cation: addition of Ba^{2+} led to the [2 + 2] cyclic imine **51**, whereas addition of Cd^{2+} resulted in almost quantitative formation of the cyclic [3 + 3] derivative **52** (Scheme 3.11).



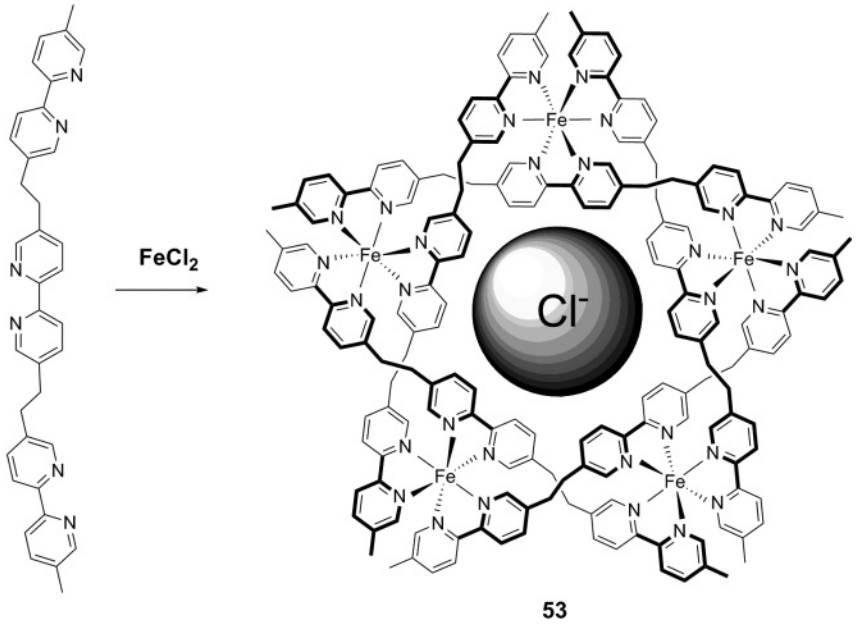
Scheme 3.11 Receptors for Cd^{2+} and Ba^{+} and building blocks used in their preparation.

The macrocycle **51** can be converted into **52** by addition of an excess of Cd^{2+} salt to the complex of **51** with Ba^{2+}. Upon reduction with $NaBH_4$, the resulting [3 + 3] amine macrocycle was the only product observed [55].

ESI-MS, ultraviolet, and NMR were used to study templating effects on the DCLs, and proved that metal ion binding is the driving force. Recently, the same group reported a highly diastereoselective amplification induced by Cd^{2+} of the heterochiral form of macrocycle **52** starting from a mixture of (*R*,*R*)- and (*S*,*S*)-*trans*-cyclohexane-1,2-diamine [56].

3.3.2
Synthetic Receptors for Anions

One of the papers that introduced the concept of DCC described the influence of anions on the composition of a multicomponent self-assembly of circular double helicates [123, 124]. Lehn *et al.* explored how the coordination of several tris-(bipyridine) strands with hexacoordinated metal ions may generate a dynamic mixture of oligomeric circular helicates of different size. The actual complex obtained depends on the conditions (Scheme 3.12). Thus, the pentameric entity **53** is expressed quantitatively in the presence of chloride anions, owing to the strong binding of a single chloride ion in the central cavity. The corresponding hexameric species is formed when the larger tetrafluoroborate or sulfate anions are used.



53

Scheme 3.12 Pentameric helicate amplified by chloride.

Kubik *et al.* have used DCC for the optimization of a synthetic receptor for anions. Kubik had previously designed the cyclic peptide **54** which binds to anions in aqueous solvents [125] (Figure 3.13). Based on the observation that **54** forms a

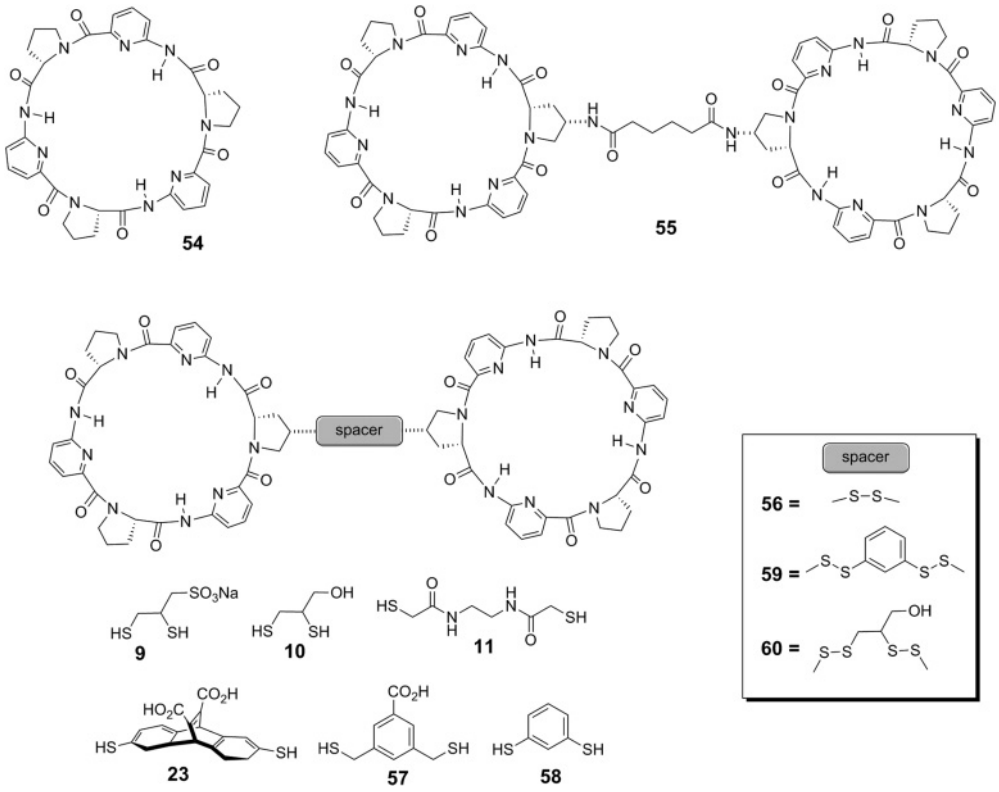


Figure 3.13 Optimization of a receptor for anions using a design approach (**55**) and using DCC (**59** and **60**).

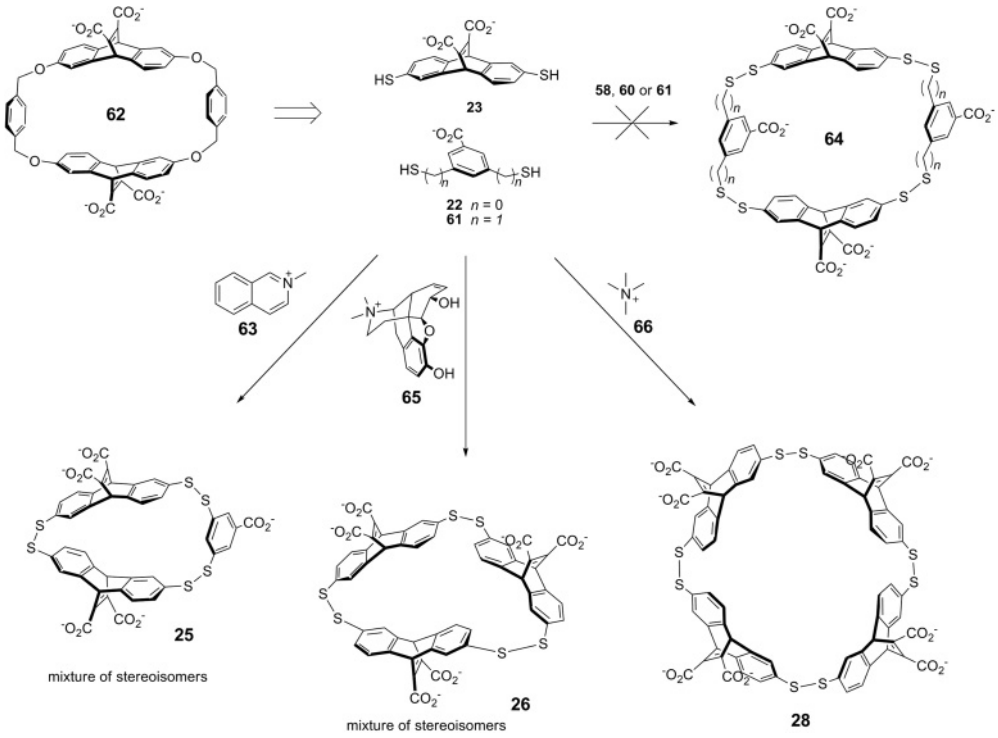
2:1 sandwich complex with the guest, a new receptor **55** was developed in which the two peptide rings are covalently connected. The linker was designed based on the crystal structure of the 2:1 iodide complex of **54** [126]. Simultaneously, a dynamic combinatorial approach was developed in which cyclopeptide disulfide **56** was allowed to choose its best spacer from the group of dithiols **9**–**11**, **23**, **57**, and **58** [8]. Addition of sulfate or iodide anions into the resulting DCL induced amplification of two new synthetic anion receptors **59** and **60**. These receptors were an order of magnitude more efficient than designed receptor **55** in binding sulfate and iodide anions.

A detailed study of these complexes showed that a significant part of their binding affinity derived from hydrophobic interactions between the two peptide rings that did not directly involve the anion [127]. The authors propose that such intrareceptor interactions may well be the key to the development of more efficient synthetic receptors [3, 7]. Since prediction of these interactions may be problematic, DCC represents an attractive method for developing these structurally complex molecules.

3.3.3
Synthetic Receptors for Organic Guests

Several organic molecules have been used to produce molecular amplification of receptors under thermodynamic conditions. A variety of receptors for organic cations have been discovered using dynamic combinatorial approaches, including receptors for alkaloids, for transition state analogs (TSAs), and for some biologically relevant guests. As discussed in Section 3.2, the research carried out using guanidinium and *n*-alkyl ammonium cations as templates has addressed several fundamental aspects related to the practical application of DCLs such as the relationship between binding constants and amplifications, and the use of immobilized templates [63, 64, 66, 77, 78, 110, 116–118].

In most cases these receptors were amplified from libraries designed to include among their members some compounds that resemble previously known receptors for the templates used. Such is the case for a series of macrocyclic receptors **25**, **26**, and **28** developed by Otto and Sanders (Scheme 3.13). These receptors were identified from amplifications observed in DCLs prepared from building blocks **22**, **23**, and **61** inspired by the successful cyclophane receptor **62**, developed by the



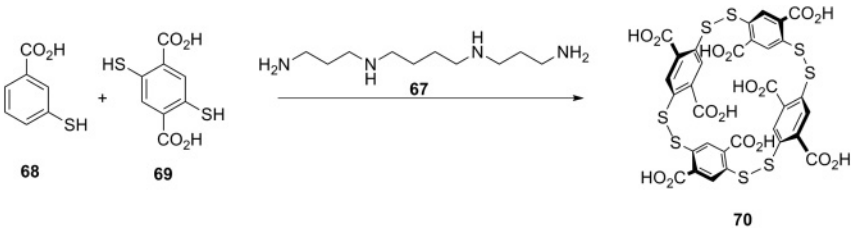
Scheme 3.13 Exposing a DCL made from three different building blocks to different templates results in the amplification of three unexpected receptors.

Dougherty laboratory [79]. Initially, guest **63**, which is one of the best binders to the original Dougherty host, was selected as a template with the expectation that it would amplify a macrocycle of the type of **64** analogous to the cyclophane **62**, but surprisingly the macrocycle **25** was amplified instead. The lack of amplification of **64** may be due to an increased rigidity of the -S–S- units in the disulfide macrocycles as compared to the -CH_2–O- units in **62**. Flexibility is important as host **62** is reported to bind guest **63** in a partially collapsed conformation [18, 79, 80, 82, 128–137] and it is likely that more rigid disulfide analogs of **62** are unable to adopt a similar collapsed conformations. In addition, the increased length and different bond angles of the disulfide linkages may cause a difference in binding.

Exposing the library to *N*-methyl morphine **65**, which is considerably larger than guest **63**, did not result in the amplification of **64** either – this time a cyclic trimer (**26**) was amplified [77]. A more intriguing result came when the relatively small tetramethylammonium iodide **66** was used as template – it amplified the cyclic tetramer **28** [78]. Since building block **23** was used as a racemic mixture, four diastereomeric disulfides can in principle be present in the DCL. Remarkably, only the diastereomer in which the four subunits have alternating chirality was significantly amplified. Apparently only this isomer can fold up in a way that effectively surrounds the small cationic guest [78]. These results confirm once again that DCC is an effective method for identifying unpredictable receptors. Such unexpected receptors were amplified even from DCLs that had been designed to contain members that resemble know receptors.

Receptors **25** and **26** were also amplified from the same DCL when using ammonium cations as templates that are TSAs of a Diels–Alder reaction and of an acetal hydrolysis respectively [138, 139]. Kinetic experiments confirmed that both macrocycles are catalytically active (see Chapter 4 for a more detailed discussion).

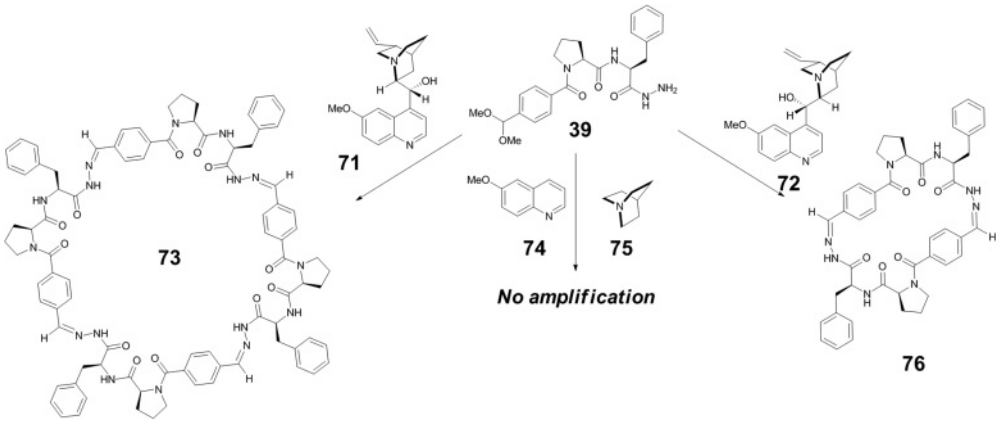
Disulfide exchange is one of the few reversible reactions currently used for DCC that operates under physiological conditions [7, 75]. This characteristic was exploited by Otto *et al.* for developing a receptor for spermine (**67**), a polyamine that plays an important role in numerous cellular processes including apoptosis and cancer. The design of the building blocks for the recognition of spermine started with identifying carboxylate–amine interactions as a potential mode of recognition. Since it was not clear whether the DCL should include linear or cyclic compounds, monothiol building block **68** and dithiol **69** were used so as to generate both types of structures (Scheme 3.14). When the DCL from **68** and **69** was



Scheme 3.14 Amplification of a receptor for spermine.

exposed to spermine, cyclic tetramer **70** was amplified. This compound turned out to have a remarkably high affinity for spermine ($K = 4.5 \times 10^7\,M^{-1}$ in 3 mM Tris buffer, pH 7.4) and binds it by forming a pseudorotaxane-type complex. Binding is strong enough to enable the receptor to sequester spermine from one of its natural hosts, DNA. Binding of spermine to DNA is known to induce a change in the helicity of some DNA sequences from the normal right-handed form to a left-handed helix. Receptor **65** was able to reverse this process, regenerating the original right-handed DNA [140].

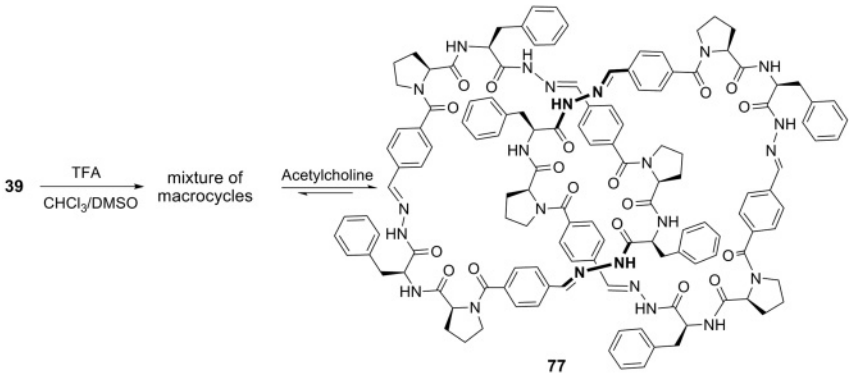
Apart from these examples, there are some other receptors that have been discovered by chance. Such is the case of the responses induced by ammonium templates on the DCL prepared from the established building block **39** based on the dipeptide L-Pro–L-Phe (Scheme 3.15). The small DCL made from **39** was affected by the presence of the cinchona alkaloids quinine (**71**) and quinidine (**72**) [61]. When the library was exposed to **71** a significant shift in the product distribution toward the cyclic tetramer **73** at the expense of the other macrocycles was observed. The template molecule quinine can be regarded as a quinoline moiety (**74**) attached to one quinuclidine (**75**) moiety by a one-carbon bridge. In an attempt to determine which of those moieties was responsible for the molecular amplification observed, **74** and **75** were tested, individually and simultaneously, as potential templates. The lack of any amplification observed for either template or their combination suggests that the binding is driven by an interaction involving both moieties in the template. In order to test the importance of the relative positions of the moieties, templating by quinidine (**72**), a diastereomer of quinine that possesses a different configuration in two of the four stereogenic centers, was tested. No amplification of the cyclic tetramer was observed; on the contrary, the concentration of this species along with other macrocycles in the mixture decreased to feed amplification of the cyclic dimer **76** that increased in abundance from 9 to 45% of the library material. Direct experimental measurement of binding



Scheme 3.15 Amplification of different receptors by diastereomeric templates.

constants for these host–guest systems is not possible since the acidic conditions required to produce the protonated guests induce library equilibration.

The same small DCL was the origin of possibly one of the most impressive and surprising template effects observed to date: the molecular amplification of a [2] catenane with the assistance of the small neurotransmitter acetylcholine [141]. Addition of acetylcholine hydrochloride to the reaction mixture slowly produces a change in the composition of the mixture by shifting the equilibrium to favor the formation of the [2]catenane **77** (Scheme 3.16). The catenane is the result of assembling six molecules of **39** into two interlocked cyclic trimers. With the assistance of the template and after 44 days of equilibration, this improbable product could be synthesized in one pot from **39** in 67% isolated yield.



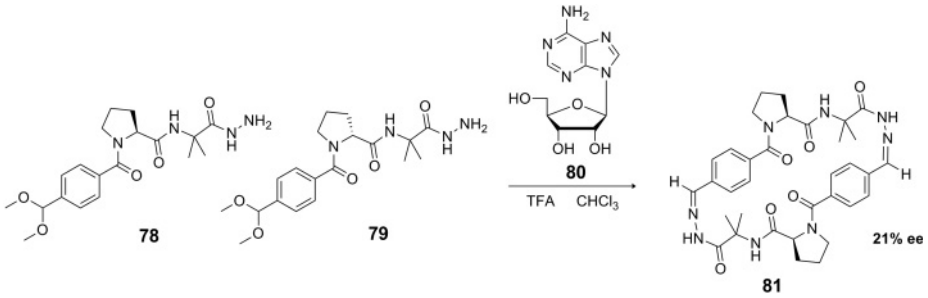
Scheme 3.16 Amplification of a [2]catenane by acetylcholine.

Although two stereoisomeric catenanes can be formed, formation of only one of them was observed. The catenane binds acetylcholine chloride with an exceptionally high affinity ($1.4 \times 10^7\ M^{-1}$ in 95 : 5 chloroform/dimethyl sulfoxide (DMSO)). Although the exact mode of binding has not been established in detail, several observations suggest that the trimethylammonium moiety of acetylcholine is the primary recognition site. The catenane binds choline, acetylcholine, and butyrylcholine in a similar way; however, replacing one methyl group in acetylcholine by ethyl reduces the binding constant by two orders of magnitude.

Discovery of this unpredictable and highly complex host through stereoselective amplification is a perhaps the most clear expression of the power of DCC as a tool for the development of new synthetic receptors.

Gagne *et al*. have recently reported the discovery of an enantioselective receptor for (–)-adenosine from a racemic DCL prepared from building blocks **78** and **79** (Scheme 3.17). When (–)-adenosine **80** was introduced into the library as a template, the amplification of the (*S*,*S*) dimer **81** with an enantiomeric excess of 21% was observed by HPLC equipped with a laser polarimeter detector.

In all the previous examples, once amplification of a receptor had been achieved, the reversible reaction was stopped and the template molecule removed to give a

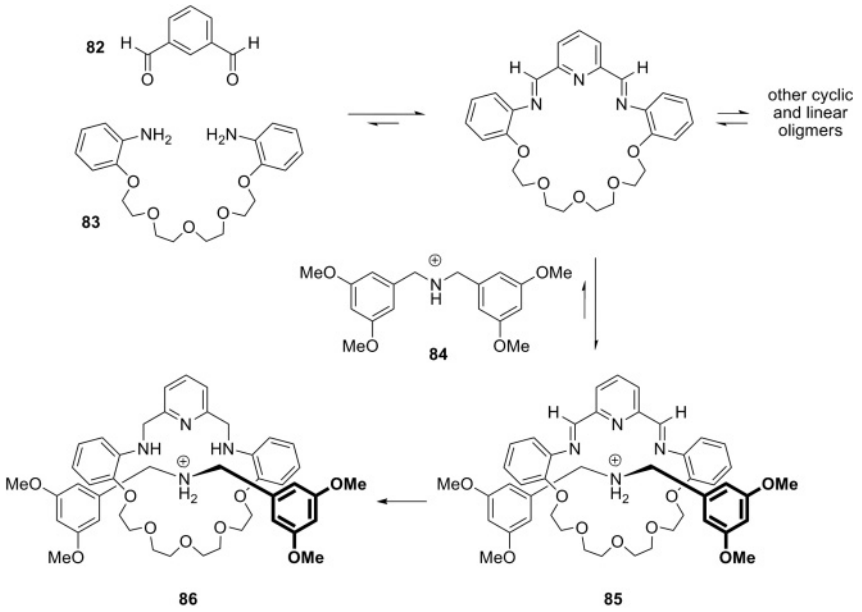


Scheme 3.17 Enantioselective amplification (ee = enantiomeric excess) of a receptor for (–)-adenosine.

kinetically stable biased mixture of products from which the receptor of interest can be isolated. There are cases, however, where the template molecule stays mechanically locked to the amplified library member. Stoddart *et al.* [142] reported the template-directed preparation of a [2]rotaxane under thermodynamic conditions by forming a macroheterocycle around a dumbbell-shaped organic template. In this case, the known interaction between a protonated dialkylammonium center and a crown ether [143] was used as stabilizing recognition motif for the rotaxane, whereas imine exchange was used as connecting reversible reaction. On mixing of dialdehyde **82** and diamine **83** in acetonitrile, a mixture of interconverting cyclic and linear oligomers was obtained. Addition of the thread **84** shifts the equilibrium of the interconverting imines toward the formation of the [2]rotaxane **85**. Subsequent reduction of the imino groups affords the kinetically stable [2]rotaxane **86** in practically quantitative yields (Scheme 3.18).

The generality of the dynamic clipping of imine-containing macrocycles around dialkylammonium ions was studied through the generation of a DCL from diamine **83** and dialdehydes **50**, **82**, and **87** (Schemes 3.11, 3.18, and Figure 3.14). The effect of three dumbbell-shaped templates (**84**, **88** and **89**, Figure 3.14) with different donor acceptor nature of the terminal benzyl groups on the formation of [2]rotaxanes was investigated [144]. This study showed that the approach is extremely sensitive to small changes in the structure of the macrocycle and dumbbell-shaped components, which, in turn, have dramatic effects on the kinetics and thermodynamics of the assemblies. Furan-containing macrocycles derived from dialdehyde **87** and π-electron-deficient dumbbell-shaped ions (like **89**) are the components of choice to form the thermodynamically most-stable [2]rotaxanes, although pyridine containing macrocycles form [2]rotaxanes that are more kinetically stable [145].

The effectiveness of these stabilizing recognition motifs and connecting reversible chemistry for the construction of interlocked molecules was explored further by the same group by preparing more complex systems. Varying the geometry of the templates and dialdehydes used it was possible a 87% yield synthesis of the dendritic [4]rotaxane **90** and a 75% yield synthesis of the cyclic [4]rotaxane **91** (Figure 3.15) [146, 147].



Scheme 3.18 Amplification of a [2]rotaxane.

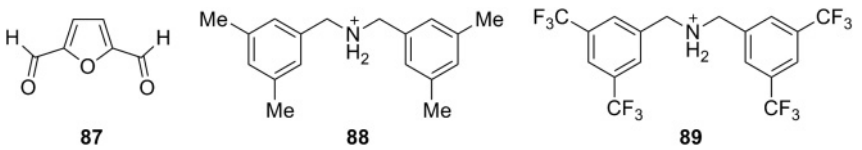


Figure 3.14 Different building blocks used to generate a DCL of rotaxanes.

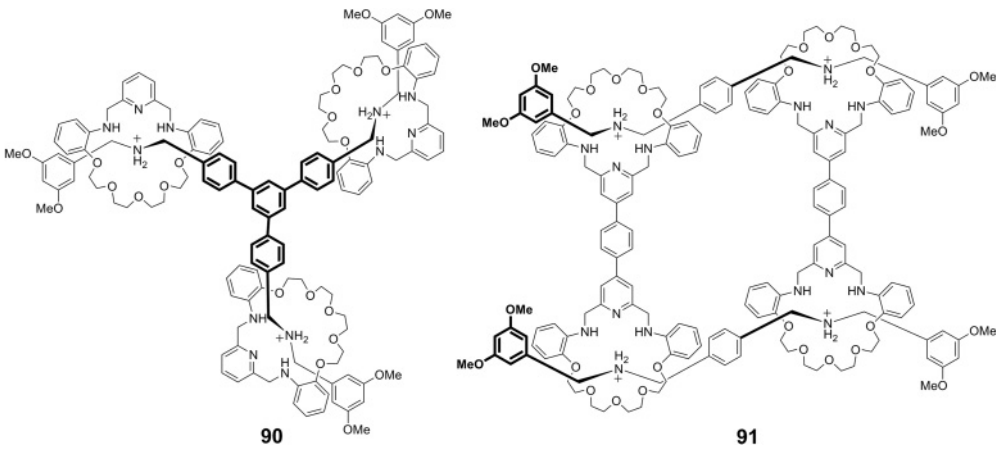


Figure 3.15 Dendritic rotaxanes prepared through DCC.

Related approaches have been used for the preparation of other mechanically interlocked molecules using different reversible processes, such as olefin metathesis [98, 148, 149] and disulfide exchange [150, 151].

Folded, linear receptors that bind to rod-like guests have been developed by the Moore lab using DCLs based on imine exchange [57]. Mixing building blocks **92** and **93**, a DCL of oligomers of different lengths was generated (Figure 3.16). When using acetonitrile as a solvent the library members can fold to form a tubular cavity

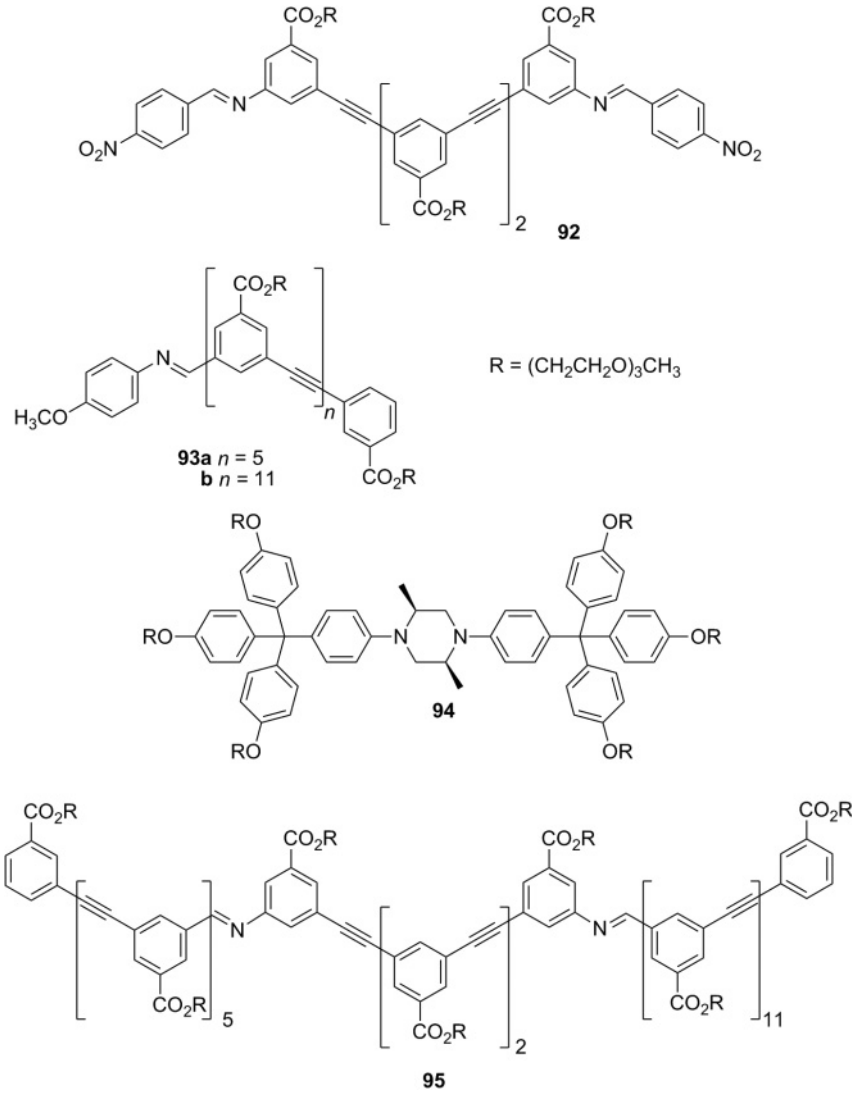


Figure 3.16 Imine oligomers used as building blocks and rod-shaped guest used to amplify a linear receptor.

in their interior that has the potential to act as a binding site for a guest. Exposure of the DCL to rod-shaped guest **94** resulted in the amplification of receptor **95**, which has a cavity length that matches the length of the guest. These results suggest that tailored size-selective recognition of guests using such systems should be possible.

3.4 Conclusions

To date, most of the receptors developed using DCC have been discovered by chance or have been inspired by receptors previously developed by design. Templation of DCLs inspired by known designed receptors has led to the discovery of several new receptors with improved binding affinity. In these cases DCC could be considered more as an optimizing tool than as source of novel host structures.

In order to develop truly novel receptors it is necessary to improve library diversity and to identify rules to guide the design of libraries to target specific guests.

The structural diversity of most DCLs described is modest. The majority of the libraries have been prepared from one to three building blocks which contain very similar recognition groups. Although a significant number of potential receptors with different shapes may be obtained with such a simple setup, it is clear that the properties of complex dynamic mixtures of receptors with significant differences in recognition groups have not been explored yet. The use of diverse recognition groups within one DCL will probably be hampered by appearance of intermolecular interactions between library members. The effect that such intermolecular interactions between library members can have on the molecular amplification of selected receptors by guests has not yet been analyzed in any detail.

In addition, the type and diversity of recognition groups that can be incorporated into a given library is restricted by the reaction conditions required for the exchange. In order to produce reusable receptors, the reaction conditions should not interfere with the binding events involved in molecular amplification. Such interference includes any effect of the reaction conditions on the binding event: negative effects or positive effects.

In the best scenario the driving force for amplification of a good receptor from a DCL produced by a given guest is the noncovalent interaction between the two species (in the case of bimolecular complexes). The strength of such interaction will be influenced by experimental conditions such as surrounding molecules (solvent, reagents, and catalysts), temperature, and so on. Those experimental conditions will define the exact molecules that are involved in the molecular recognition process (e.g., the acidity or basicity of the medium will determine whether the species are neutral or charged, etc.). The only way to be certain that the binding

observed in the DCL can be used for further applications is to reproduce the whole set of conditions. However, by definition, the receptor will not be a stable under such conditions. As a consequence, two main strategies have been used to stabilize the amplified receptor in a way that it can be isolated and actually applied: (i) to alter the receptor or (ii) to alter the environment.

A typical change to freeze interconverting receptors is the covalent transformation of the receptor (reduction of the C=N bond of imines or covalent trapping of receptors assembled by noncovalent interactions). Such transformation may affect binding strength in a significant manner compromising the usefulness of the receptor.

Changes in the environment usually include removal of the catalyst of the exchange reaction. In principle, if the catalyst does not participate in binding, such changes should not represent a problem. However, when the catalyst is a proton or a strong base, then it is quite likely that, apart from activating the reacting groups involved in the reversible covalent chemistry, it changes the protonation state of the guest and/or receptor. Typical stabilizing changes in environment include adjusting the acidity of the medium (hydrazones, acetals, etc.) or removal of a metal catalyst (alkene metathesis, etc). Even if the connecting bonds are not covalently transformed by those changes, the isolated receptors will not necessary be the same as those that were originally amplified. Thus, although it can be advantageous that hydrazones and acetals are stable in neutral media, we should ask ourselves: are the receptor and the guest that originally interacted in the DCL stable in neutral media? Or in more general terms: is the binding event stable? In this respect, disulfide exchange is one of the most promising exchange processes since it can be active under neutral conditions and removal of thiolate is enough to stabilize the receptor. Such removal can be achieved simply by complete oxidation, which normally happens anyway during the DCL experiment or during receptor isolation (unless there is a free thiol moiety in the receptor).

Despite these limitations, in the last decade DCC has produced several unpredictable and structurally complex receptors with exceptionally high affinities.

One advantage of the approach is that resynthesis and isolation of the receptors is required only after the amplification results suggest that the selected molecule binds the guest.

In addition, the strategy also provides a high-yield synthetic method for the selected receptor. The use of reversible chemistry under thermodynamic control in combination with solid supported templates may speed up this time-consuming step in the discovery process of new receptors.

The development of such tools increases the potential of DCC as a source of synthetic receptors for specific applications in the near future. However, the practical utility of receptors developed using DCC has not been fully demonstrated to date.

The next challenge for the chemists working in the field is to establish DCC as an alternative strategy to solve specific problems that are difficult to solve by other traditional strategies.

References

1 Sanders, J.K.M. (1998) *Chem. Eur. J.*, **4**, 1378–1383.
2 Houk, K.N., Leach, A.G., Kim, S.P., and Zhang, X. (2003) *Angew. Chem. Int. Ed.*, **42**, 4872–4897.
3 Otto, S. (2006) *Dalton Trans.*, 2861–2864.
4 Srinivasan, N. and Kilburn, J.D. (2004) *Curr. Opin. Chem. Biol.*, **8**, 305–310.
5 Still, W.C. (1996) *Acc. Chem. Res.*, **29**, 155–163.
6 Corbett, P.T., Leclaire, J., Vial, L., West, K.R., Wietor, J.-L., Sanders, J.K.M., and Otto, S. (2006) *Chem. Rev.*, **106**, 3652–3711.
7 Otto, S. and Severin, K. (2007) *Top. Curr. Chem.*, **277**, 267–288.
8 Otto, S. and Kubik, S. (2003) *J. Am. Chem. Soc.*, **125**, 7804–7805.
9 de Bruin, B., Hauwert, P., and Reek, J.N.H. (2006) *Angew. Chem. Int. Ed.*, **45**, 2660–2663.
10 Allen, F.H., Motherwell, W.D.S., Raithby, P.R., Shields, G.P., and Taylor, R. (1999) *New J. Chem.*, **23**, 25–34.
11 Crego Calama, M., Hulst, R., Fokkens, R., Nibbering, N.M.M., Timmerman, P., and Reinhoudt, D.N. (1998) *Chem. Commun.*, 1021–1022.
12 Timmerman, P., Vreekamp, R.H., Hulst, R., Verboom, W., Reinhoudt, D.N., Rissanen, K., Udachin, K.A., and Ripmeester, J. (1997) *Chem. Eur. J.*, **3**, 1823–1832.
13 Wu, A.X., Mukhopadhyay, P., Chakraborty, A., Fettinger, J.C., and Isaacs, L. (2004) *J. Am. Chem. Soc.*, **126**, 10035–10043.
14 Crego Calama, M., Timmerman, P., and Reinhoudt, D.N. (2000) *Angew. Chem. Int. Ed.*, **39**, 755–758.
15 Goral, V., Nelen, M.I., Eliseev, A.V., and Lehn, J.-M. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 1347–1352.
16 Albrecht, M., Janser, I., Runsink, J., Raabe, G., Weis, P., and Frohlich, R. (2004) *Angew. Chem. Int. Ed.*, **43**, 6662–6666.
17 Telfer, S.G., Yang, X.J., and Williams, A.F. (2004) *Dalton Trans.*, 699–705.
18 Pang, K.L., Guo, D., Duan, C.Y., Mo, H., and Meng, Q.J. (2003) *Inorg. Chem.*, **42**, 5453–5455.
19 Saur, I. and Severin, K. (2005) *Chem. Commun.*, 1471–1473.
20 Stulz, E., Scott, S.M., Bond, A.D., Teat, S.J., and Sanders, J.K.M. (2003) *Chem. Eur. J.*, **9**, 6039–6048.
21 Shoji, O., Okada, S., Satake, A., and Kobuke, Y. (2005) *J. Am. Chem. Soc.*, **127**, 2201–2210.
22 Huc, I., Krische, M.J., Funeriu, D.P., and Lehn, J.-M. (1999) *Eur. J. Inorg. Chem.*, 1415–1420.
23 Cai, M.M., Shi, X.D., Sidorov, V., Fabris, D., Lam, Y.F., and Davis, J.T. (2002) *Tetrahedron*, **58**, 661–671.
24 (a) Pantos, G.D., Wietor, J.L., and Sanders, J.K.M. (2007) *Angew. Chem.*, **119**, 2288–2290.
(b) Pantos, G.D., Wietor, J.L., and Sanders, J.K.M. (2007) *Angew. Chem. Int. Ed.*, **46**, 2238–2240.
(c) Pantos, G.D., Wietor, J.L., and Sanders, J.K.M. (2008) *Angew. Chem. Int. Ed.*, **47**, 2689–2692.
25 Campos-Fernandez, C.S., Schottel, B.L., Chifotides, H.T., Bera, J.K., Bacsa, J., Koomen, J.M., Russell, D.H., and Dunbar, K.R. (2005) *J. Am. Chem. Soc.*, **127**, 12909–12923.
26 Stulz, E., Ng, Y.-F., Scott, S.M., and Sanders, J.K.M. (2002) *Chem. Commun.*, 524–525.
27 Try, A.C., Harding, M.M., Hamilton, D.G., and Sanders, J.K.M. (1998) *Chem. Commun.*, 723–724.
28 Vilar, R. (2003) *Angew. Chem. Int. Ed.*, **42**, 1460–1477.
29 Gimeno, N. and Vilar, R. (2006) *Coord. Chem. Rev.*, **250**, 3161–3189.
30 Cardullo, F., Crego Calama, M., Snellink-Ruel, B.H.M., Weidmann, J.L., Bielejewska, A., Fokkens, R., Nibbering, N.M.M., Timmerman, P., and Reinhoudt, D.N. (2000) *Chem. Commun.*, 367–368.
31 Brady, P.A., Bonar-Law, R.P., Rowan, S.J., Suckling, C.J., and Sanders, J.K.M. (1996) *Chem. Commun.*, 319–320.

32 Rowan, S.J., Hamilton, D.G., Brady, P.A., and Sanders, J.K.M. (1997) *J. Am. Chem. Soc.*, **119**, 2578–2579.

33 Rowan, S.J., Brady, P.A., and Sanders, J.K.M. (1996) *Angew. Chem. Int. Ed. Engl.*, **35**, 2143–2145.

34 Rowan, S.J., Brady, P.A., and Sanders, J.K.M. (1996) *Tetrahedron Lett.*, **37**, 6013–6016.

35 Rowan, S.J. and Sanders, J.K.M. (1998) *Org. Chem.*, **63**, 1536–1546.

36 Brady, P.A. and Sanders, J.K.M. (1997) *J. Chem. Soc., Perkin Trans.* **1**, 3237–3253; structural corrections: *J. Chem. Soc., Perkin Trans.* **1**, 1998, 2119.

37 Amatore, C., Jutand, A., Meyer, G., and Mottier, L. (1999) *Chem. Eur. J.*, **5**, 466–473.

38 Kaiser, G. and Sanders, J.K.M. (2000) *Chem. Commun.*, 1763–1764.

39 Larsson, R., Pei, Z.C., and Ramström, O. (2004) *Angew. Chem. Int. Ed.*, **43**, 3716–3718.

40 Woll, M.G. and Gellman, S.H. (2004) *J. Am. Chem. Soc.*, **126**, 11172–11174.

41 Leclaire, J., Vial, L., Otto, S., and Sanders, J.K.M. (2005) *Chem. Commun.*, 1959–1961.

42 Eldred, S.E., Stone, D.A., Gellman, S.H., and Stahl, S.S. (2003) *J. Am. Chem. Soc.*, **125**, 3422–3423.

43 Hoerter, J.M., Otte, K.M., Gellman, S.H., Cui, Q., and Stahl, S.S. (2008) *J. Am. Chem. Soc.*, **130**, 647–654.

44 Fuchs, B., Nelson, A., Star, A., Stoddart, J.F., and Vidal, S.B. (2003) *Angew. Chem. Int. Ed.*, **42**, 4220–4224.

45 Cacciapaglia, R., Di Stefano, S., and Mandolini, L. (2005) *J. Am. Chem. Soc.*, **127**, 13666–13671.

46 Sutton, L.R., Donaubauer, W.A., Hampel, F., and Hirsch, A. (2004) *Chem. Commun.*, 1758–1759.

47 Wipf, P., Mahler, S.G., and Okumura, K. (2005) *Org. Lett.*, **7**, 4483–4486.

48 Rowan, S.J. and Stoddart, J.F. (1999) *Org. Lett.*, **1**, 1913–1916.

49 Oh, K., Jeong, K.S., and Moore, J.S. (2001) *Nature*, **414**, 889–893.

50 Zhao, D.H. and Moore, J.S. (2002) *J. Am. Chem. Soc.*, **124**, 9996–9997.

51 Zhao, D.H. and Moore, J.S. (2003) *Macromolecules*, **36**, 2712–2720.

52 Giuseppone, N., Schmitt, J.L., Schwartz, E., and Lehn, J.M. (2005) *J. Am. Chem. Soc.*, **127**, 5528–5539.

53 Lüning, U. (2004) *J. Inclusion Phenom. Macrocycl. Chem.*, **49**, 81–84.

54 Storm, O. and Lüning, U. (2002) *Chem. Eur. J.*, **8**, 793–798.

55 Gonzalez-Alvarez, A., Alfonso, I., Lopez-Ortiz, F., Aguirre, A., Garcia-Granda, S., and Gotor, V. (2004) *Eur. J. Org. Chem.*, 1117–1127.

56 Gonzalez-Alvarez, A., Alfonso, I., and Gotor, V. (2006) *Chem. Commun.*, 2224–2226.

57 Nishinaga, T., Tanatani, A., Oh, K.C., and Moore, J.S. (2002) *J. Am. Chem. Soc.*, **124**, 5934–5935.

58 Huc, I. and Lehn, J.-M. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 2106–2110.

59 Nguyen, R. and Huc, I. (2003) *Chem. Commun.*, 942–943.

60 Cousins, G.R.L., Poulsen, S.A., and Sanders, J.K.M. (1999) *Chem. Commun.*, 1575–1576.

61 Bulos, F., Roberts, S.L., Furlán, R.L.E., and Sanders, J.K.M. (2007) *Chem. Commun.*, 3092–3093.

62 Roberts, S.L., Furlán, R.L.E., Otto, S., and Sanders, J.K.M. (2003) *Org. Biomol. Chem.*, **1**, 1625–1633.

63 Furlán, R.L.E., Ng, Y.-F., Cousins, G.R.L., Redman, J.E., and Sanders, J.K.M. (2002) *Tetrahedron*, **58**, 771–778.

64 Roberts, S.L., Furlán, R.L.E., Cousins, G.R.L., and Sanders, J.K.M. (2002) *Chem. Commun.*, 938–939.

65 Furlán, R.L.E., Ng, Y.-F., Otto, S., and Sanders, J.K.M. (2001) *J. Am. Chem. Soc.*, **123**, 8876–8877.

66 Cousins, G.R.L., Furlán, R.L.E., Ng, Y.-F., Redman, J.E., and Sanders, J.K.M. (2001) *Angew. Chem. Int. Ed.*, **40**, 423–427.

67 Furlán, R.L.E., Cousins, G.R.L., and Sanders, J.K.M. (2000) *Chem. Commun.*, 1761–1762.

68 Polyakov, V.A., Nelen, M.I., Nazarpack-Kandlousy, N., Ryabov, A.D., and Eliseev, A.V.J. (1999) *Phys. Org. Chem.*, **12**, 357–363.

69 Nazarpack-Kandlousy, N., Zweigenbaum, J., Henion, J., and Eliseev, A.V.J. (1999) *Comb. Chem.*, **1**, 199–206.

70 Nazarpack-Kandlousy, N., Nelen, M.I., Goral, V., and Eliseev, A.V.J. (2002) *J. Org. Chem.*, **67**, 59–65.

71 Hioki, H. and Still, W.C. (1998) *J. Org. Chem.*, **63**, 904–905.

72 Wallace, T.J., Schriesheim, A., and Bartok, W. (1963) *J. Org. Chem.*, **28**, 1311–1314.

73 Lees, W.J. and Whitesides, G.M. (1993) *J. Org. Chem.*, **58**, 642–647.

74 Gilbert, H.F. (1995) *Methods Enzymol.*, **251**, 8–28.

75 Ramström, O. and Lehn, J.-M. (2000) *ChemBioChem*, **1**, 41–47.

76 Otto, S., Furlán, R.L.E., and Sanders, J.K.M. (2000) *J. Am. Chem. Soc.*, **122**, 12063–12064.

77 Otto, S., Furlán, R.L.E., and Sanders, J.K.M. (2002) *Science*, **297**, 590–593.

78 Corbett, P.T., Tong, L.H., Sanders, J.K.M., and Otto, S. (2005) *J. Am. Chem. Soc.*, **127**, 8902–8903.

79 Ngola, S.M., Kearney, P.C., Mecozzi, S., Russell, K., and Dougherty, D.A. (1999) *J. Am. Chem. Soc.*, **121**, 1192–1201.

80 Kearney, P.C., Mizoue, L.S., Kumpf, R.A., Forman, J.E., McCurdy, A., and Dougherty, D.A. (1993) *J. Am. Chem. Soc.*, **115**, 9907–9919.

81 Stauffer, D.A., Barrans, R.E., and Dougherty, D.A. (1990) *J. Org. Chem.*, **55**, 2762–2767.

82 Petti, M.A., Sheppod, T.J., Barrans, R.E., and Dougherty, D.A. (1988) *J. Am. Chem. Soc.*, **110**, 6825–6840.

83 ten Cate, A.T., Dankers, P.W.D., Sijbesma, R.P., and Meijer, E.W. (2005) *J. Org. Chem.*, **70**, 5799–5803.

84 Kieran, A.L., Bond, A.D., Belenguer, A.M., and Sanders, J.K.M. (2003) *Chem. Commun.*, 2674–2675.

85 Pérez-Fernández, R., Pittelkow, M., Belenguer, A.M., and Sanders, J.K.M. (2008) *Chem. Commun.*, 1738–1740.

86 Toste, F.D., Chatterjee, A.K., and Grubbs, R.H. (2002) *Pure Appl. Chem.*, **74**, 7–10.

87 (a) Marsella, M.J., Maynard, H.D., and Grubbs, R.H. (1997) *Angew. Chem.*, **109**, 1147–1150.
(b) Marsella, M.J., Maynard, H.D., and Grubbs, R.H. (1997) *Angew. Chem. Int. Ed. Engl.*, **36**, 1101–1103.

88 Giger, T., Wigger, M., Audetat, S., and Benner, S.A. (1998) *Synlett*, 688–691.

89 van Gerven, P.C.M., Elemans, J.A.A.W., Gerritsen, J.W., Speller, S., Nolte, R.J.M., and Rowan, A.E. (2005) *Chem. Commun.*, 3535–3537.

90 Zhang, W. and Moore, J.S. (2005) *J. Am. Chem. Soc.*, **127**, 11863–11870.

91 Furstner, A. and Davies, P.W. (2005) *Chem. Commun.*, 2307–2320.

92 Oh, K., Jeong, K.S., and Moore, J.S. (2003) *J. Org. Chem.*, **68**, 8397–8403.

93 Rowan, S.J. and Sanders, J.K.M. (1997) *Chem. Commun.*, 1407–1408.

94 Rowan, S.J., Lukeman, P.S., Reynolds, D.J., and Sanders, J.K.M. (1998) *New J. Chem.*, **22**, 1015–1018.

95 Jacobsen, H. and Stockmayer, W.H. (1950) *J. Chem. Phys.*, **18**, 1600–1606.

96 Hubbard, P.A., Brittain, W.J., Mattice, W.L., and Brunelle, D.J. (1998) *Macromolecules*, **31**, 1518–1522.

97 Xu, S. and Giuseppone, N. (2008) *J. Am. Chem. Soc.*, **130**, 1826–1827.

98 Kidd, T.J., Leigh, D.A., and Wilson, A.J. (1999) *J. Am. Chem. Soc.*, **121**, 1599–1600.

99 Giuseppone, N. and Lehn, J.-M. (2006) *Chem. Eur. J.*, **12**, 1715–1722.

100 Giuseppone, N. and Lehn, J.-M. (2004) *J. Am. Chem. Soc.*, **126**, 11448–11449.

101 Giuseppone, N., Kuks, G., and Lehn, J.-M. (2006) *Chem. Eur. J.*, **12**, 1723–1735.

102 Kubik, S. and Goddard, R. (1999) *J. Org. Chem.*, **64**, 9475–9486.

103 (a) Grote, Z., Scopelliti, R., and Severin, K. (2003) *Angew. Chem.*, **115**, 3951–3955. (b) Grote, Z., Scopelliti, R., and Severin, K. (2003) *Angew. Chem. Int. Ed.*, **42**, 3821–3825.

104 Severin, K. (2004) *Chem. Eur. J.*, **10**, 2565–2580.

105 Corbett, P.T., Otto, S., and Sanders, J.K.M. (2004) *Chem. Eur. J.*, **10**, 3139–3143.

106 Corbett, P.T., Otto, S., and Sanders, J.K.M. (2004) *Org. Lett.*, **6**, 1825–1827.

107 Corbett, P.T., Sanders, J.K.M., and Otto, S. (2005) *J. Am. Chem. Soc.*, **127**, 9390–9392.

108 Saur, B., Scopelliti, R., and Severin, K. (2006) *Chem. Eur. J.*, **12**, 1058–1066.

109 Brady, P.A. and Sanders, J.K.M. (1998) *New J. Chem.*, 411–417.

110 Corbett, P.T., Sanders, J.K.M., and Otto, S. (2008) *Chem. Eur. J.*, **14**, 2153–2166.

111 Cacciapaglia, R., Di Stefano, S., Mandolini, L., Mencarelli, P., and Ugozzoli, F. (2008) *Eur. J. Org. Chem.*, 186–195.

112 Cacciapaglia, R., Di Stefano, S., and Mandolini, L. (2006) *Chem. Eur. J.*, **12**, 8566–8570.

113 Kubota, Y., Sakamoto, S., Yamaguchi, K., and Fujita, M. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 4854–4856.

114 Poulsen, S.A., Gates, P.J., Cousins, G.R.L., and Sanders, J.K.M. (2000) *Rapid Commun. Mass Spectrom.*, **14**, 44–48.

115 Voshell, S.M., Lee, S.J., and Gagne, M.R. (2006) *J. Am. Chem. Soc.*, **128**, 12422–12423.

116 Eliseev, A.V. and Nelen, M.I. (1997) *J. Am. Chem. Soc.*, **119**, 1147–1148.

117 Eliseev, A.V. and Nelen, M.I. (1998) *Chem. Eur. J.*, **4**, 825–834.

118 Besenius, P., Cormack, P.A.G., Ludlow, R.F., Otto, S., and Sherrington, D.C. (2008) *Chem. Commun.*, 2809–2811.

119 Laughrey, Z.R. and Gibb, B.C. (2005) *Top. Curr. Chem.*, **249**, 67–125.

120 Liu, J., West, K.R., Bondy, C.R., and Sanders, J.K.M. (2007) *Org. Biomol. Chem.*, **5**, 778–786.

121 Matsumoto, M. and Nicholas, K.M. (2007) *J. Org. Chem.*, **72**, 9308–9313.

122 Bornaghi, L.F., Wilkinson, B.L., Kiefel, M.J., and Poulsen, S.A. (2004) *Tetrahedron Lett.*, **45**, 9281–9284.

123 Hasenknopf, B., Lehn, J.-M., Boumediene, N., Dupont-Gervais, A., Van Dorsselaer, A., Kneisel, B., and Fenske, D. (1997) *J. Am. Chem. Soc.*, **119**, 10956–10962.

124 Hasenknopf, B., Lehn, J.-M., Kneisel, B.O., Baum, G., and Fenske, D. (1996) *Angew. Chem. Int. Ed. Engl.*, **35**, 1838–1840.

125 Kubik, S., Goddard, R., Kirchner, R., Nolting, D., and Seidel, J. (2001) *Angew. Chem. Int. Ed.*, **40**, 2648–2651.

126 Kubik, S., Kirchner, R., Nolting, D., and Seidel, J. (2002) *J. Am. Chem. Soc.*, **124**, 12752–12760.

127 Rodriguez-Docampo, Z., Pascu, S.I., Kubik, S., and Otto, S. (2006) *J. Am. Chem. Soc.*, **128**, 11206–11210.

128 Petti, M.A., Shepodd, T.J., and Dougherty, D.A. (1986) *Tetrahedron Lett.*, **27**, 807– 810.

129 Shepodd, T.J., Petti, M.A., and Dougherty, D.A. (1986) *J. Am. Chem. Soc.*, **108**, 6085–6087.

130 Shepodd, T.J., Petti, M.A., and Dougherty, D.A. (1988) *J. Am. Chem. Soc.*, **110**, 1983–1985.

131 Stauffer, D.A. and Dougherty, D.A. (1988) *Tetrahedron Lett.*, **29**, 6039–6042.

132 Dougherty, D.A. and Stauffer, D.A. (1990) *Science*, **250**, 1558–1560.

133 (a) Stauffer, D.A., Barrans, R.E., and Dougherty, D.A. (1990) *Angew. Chem.*, **102**, 953–956. (b) Stauffer, D.A., Barrans, R.E., and Dougherty, D.A. (1990) *Angew. Chem. Int. Ed. Engl.*, **29**, 915–918.

134 McCurdy, A., Jimenez, L., Stauffer, D.A., and Dougherty, D.A. (1992) *J. Am. Chem. Soc.*, **114**, 10314–10321.

135 Forman, J.E., Marsh, R.E., Schaefer, W.P., and Dougherty, D.A. (1993) *Acta Crystallogr. A*, **49**, 892–896.

136 Forman, J.E., Barrans, R.E., and Dougherty, D.A. (1995) *J. Am. Chem. Soc.*, **117**, 9213–9228.

137 Ngola, S.M. and Dougherty, D.A. (1996) *J. Org. Chem.*, **61**, 4355–4360.

138 Brisig, B., Sanders, J.K.M., and Otto, S. (2003) *Angew. Chem. Int. Ed.*, **42**, 1270–1273.

139 Vial, L., Sanders, J.K.M., and Otto, S. (2005) *New J. Chem.*, **29**, 1001–1003.

140 Vial, L., Ludlow, R.F., Leclaire, J., Perez-Fernandez, R., and Otto, S. (2006) *J. Am. Chem. Soc.*, **128**, 10253–10257.

141 Lam, R.T.S., Belenguer, A., Roberts, S.L., Naumann, C., Jarrosson, T., Otto, S., and Sanders, J.K.M. (2005) *Science*, **308**, 667–669.

142 Glink, P.T., Oliva, A.I., Stoddart, J.F., White, A.J.P., and Williams, D.J. (2001) *Angew. Chem. Int. Ed.*, **40**, 1870–1875.

143 Ashton, P.R., Campbell, P.J., Chrystal, E.J.T., Glink, P.T., Menzer, S., Philp, D., Spencer, N., Stoddart, J.F., Tasker, P.A., and Williams, D.J. (1995) *Angew. Chem. Int. Ed. Engl.*, **34**, 1865–1869.

144 Horn, M., Ihringer, J., Glink, P.T., and Stoddart, J.F. (2003) *Chem. Eur. J.*, **9**, 4046–4054.

145 Aricó, F., Badjic, J.D., Cantrill, S.J., Flood, A.H., Leung, K.C.-F., Liu, Y., and Stoddart, J.F. (2005) *Top. Curr. Chem.*, **249**, 203–259.

146 Aricó, F., Chang, T., Cantrill, S.J., Khan, S.I., and Stoddart, J.F. (2005) *Chem. Eur. J.*, **11**, 4655–4666.

147 Leung, K.C.-F., Aricó, F., Cantrill, S.J., and Stoddart, J.F. (2005) *J. Am. Chem. Soc.*, **127**, 5808–5810.

148 Guidry, E.N., Cantrill, S.J., Stoddart, J.F., and Grubbs, R.H. (2005) *Org. Lett.*, **7**, 2129–2132.

149 Hamilton, D.G., Feeder, N., Teat, S.J., and Sanders, J.K.M. (1998) *New J. Chem.*, 1019–1021.

150 Wang, W., Wang, L.Q., Palmer, B.J., Exarhos, G.J., and Li, A.D.Q. (2006) *J. Am. Chem. Soc.*, **128**, 11150–11159.

151 Furusho, Y., Oku, T., Hasegawa, T., Tsuboi, A., Kihara, N., and Takata, T. (2003) *Chem. Eur. J.*, **9**, 2895–2903.

4
Dynamic Combinatorial Chemistry for Catalytic Applications

Pierre-Alain R. Breuil and Joost N.H. Reek

4.1
Introduction

The use of dynamic combinatorial chemistry (DCC) can be considered as a paradigm shift in chemistry as it involves dynamic mixtures of compounds rather than pure entities that were traditionally aimed for in the area of chemistry [1]. In analogy to natural systems, new properties can emerge from these mixtures, validating the recent interest in this new area of research. Dynamic mixtures of compounds become particularly valuable if they can be combined with a proper selection process with which a specific compound with desired properties can be identified. In a traditional combinatorial approach, a large library of analog compounds is synthesized and subsequently evaluated in a parallel fashion. In a dynamic combinatorial approach, only a small number of building blocks are generally synthesized, from which a large virtual (dynamic) library can be constructed. A proper selection pressure is required to shift the equilibrium of the dynamic library, and to select and amplify a member with the desired properties. Especially in research areas where rational design has met limited success and progress is based on trial-and-error and combinatorial approaches, DCC can provide new tools to find solutions for standing problems. Initially, the approach was developed for the identification of novel receptor molecules. The selector in these examples is obviously the target for which a receptor is desired. Along with this development, novel analytical tools have been developed as well as theoretical understanding and the DCC field has shifted to other applications that are described in the various chapters of this book. One of the underdeveloped fields of applications of DCC so far is catalysis. In this chapter, we discuss the potential of DCC in the field of catalysis. Catalysts accelerate a reaction without being consumed and there are many different types of catalysts that have different modes of action. In the area of supramolecular chemistry, many "supramolecular catalysts" have been developed [2]. Generally, their working principles are strongly related to receptor molecules; they both can host a substrate or several substrates, but in a supramolecular catalyst they are converted within the cavity of the receptor. Transition metal catalysts, on the other hand, generally create a new reaction

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

pathway that is not available in the absence of the metal catalyst, via elementary steps that occur at the metal site (oxidative addition, transmetallation, migration, elimination, reductive elimination, to mention a few) [3]. Although the application of DCC to supramolecular catalysts seems easier, the implications of DCC with transition metal catalysis are much larger as it may directly result in practical applications. Theoretically, there is no difference in finding the best catalysts from a dynamic library of transition metal catalysts and supramolecular catalysts, as we will see. For an efficient selection of catalyst via DCC one needs:

- A dynamic library of catalytic systems from a limited number of building blocks.
- The library should be adaptive or responsive to additives.
- A proper selection procedure. This in principle implies that an additive should be found to which the dynamic library responds and from the response the best catalyst(s) (in terms of activity or selectivity) should present itself.

Finding the proper selection procedure is clearly a challenge. If one considers an energetic pathway for a chemical transformation as depicted in Figure 4.1, the objective is to lower the energy barrier of this reaction. This can be achieved by stabilizing the transition state and/or elevating the energy level of the intermediate (or substrates) prior to this step. The overall effect in both cases would be an overall lower energy barrier and thus a higher reaction rate. Ideally, the selection procedure should be done with the transition state, but the inherent instability makes this impossible. Therefore, a transition state analog (TSA) should be designed that closely resembles the transition state structure. Selection procedures can also be envisioned using the intermediate of a reaction, in which case an inverse correlation between the stability and the reactivity is expected. The more stable the intermediate complex, the larger the energy barrier of the reaction and the slower the reaction.

Marcus theory describes the reaction pathway as the result of two intersecting parabola [4], the transition state being on the intersection of these. If the reaction

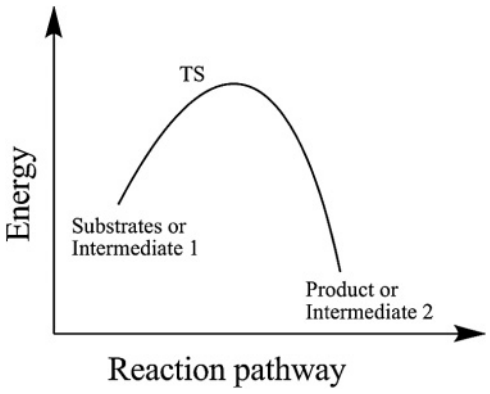


Figure 4.1 General energy profile of a reaction path.

goes via an early transition state, the transition state is substrate-like (or similar to intermediate 1), whereas a late transition state is more product-like (or intermediate 2). This difference has consequences for the design of the selection procedure, as will be explained below. In this explanation we focus on activity, but the extension to selectivity is a matter of comparison of two competing pathways. If the reaction proceeds via an early transition state, its stabilization implies also significant stabilization of the intermediate (Figure 4.2a). The decrease of the energy barrier depends on the difference of energy between the two stabilized species, which is expected to be small. In contrast, if the reaction proceeds via a late transition state, the transition state stabilization will hardly affect the intermediate 1 (Figure 4.2b) and is anticipated to be very effective. The stabilization of a late transition state (Figure 4.2b) also leads to the stabilization of the product (or intermediate 2), which may lead to selection of catalysts that show product inhibition, detrimental for the reaction rate. A selection procedure which is based on stabilization of a TSA is therefore most effective for reactions that proceed via a late transition state.

A selection procedure based on the reaction intermediate (or an analog) is based on the destabilization of the intermediate with respect to the transition state. In an early transition state reaction path, the destabilization of the transition state will be close to that of the intermediate and is therefore less effective as the

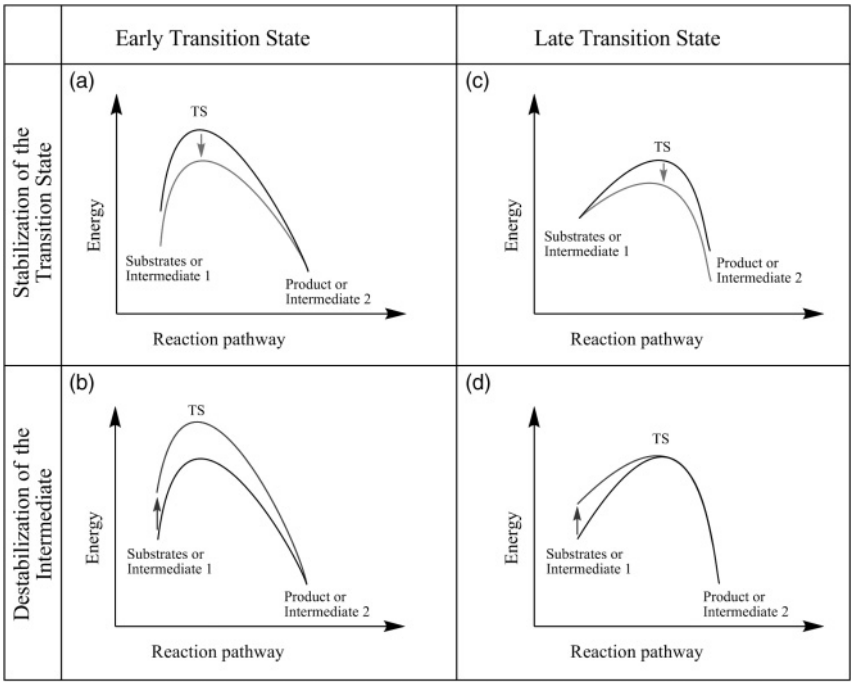


Figure 4.2 Selection procedures based on TSAs (a and c) and intermediates (b and d) in reactions that proceed via early (a and b) or late (c and d) transition states.

difference in overall energy barrier will be small (Figure 4.2c). The same strategy applied to a reaction that proceeds via a late transition state (Figure 4.2d) is expected to be more effective as this destabilization has a reduced effect on the transition state energy.

4.2
Dynamic Combinatorial Approaches to Cage Catalysts

Enzymes, Nature's catalysts [5], encapsulate multiple functionalities within their cavity where the catalytic conversion takes place, and can be extremely active and selective for a range of chemical reactions. Therefore, enzymes have served as the major source of inspiration for supramolecular catalysis, but at the same time the working principles of enzymes are still subject to debate. In 1948, Pauling [6] proposed that enzymes stabilize transition states to a larger extent than reagents in their (vibrational) ground state by means of noncovalent interactions between the functional groups in the enzyme cavity and the compounds inside the cavity. These initial ideas have inspired many scientists from various fields to explore similar approaches for synthetic systems. The main focus in the area of supramolecular catalysis has been on host–guest catalysis [2] in which a substrate is bound in a cavity possibly next to the catalytically active center. In addition, cage compounds have been used as hosts for substrate molecules. This approach has resulted in numerous interesting examples of supramolecular catalysts. Most of these cage catalysts work by (i) positioning of substrates (or bring substrates together for a bimolecular reaction) and (ii) transition state stabilization (by preorganization of substrates or interaction) [7]. Since there were precedents for selecting the best cage receptor for certain guest molecules and similar cages had been applied as catalysts, it is no surprise that the first examples in the area of dynamic combinatorial catalyst selection come from this approach.

4.2.1
Libraries of Cage Molecules and Dynamic Selection of Hosts/Guests

Molecular cages and molecular capsules are a special class of host molecules with a very defined three-dimensional structure, including a hollow interior which under judicious conditions could engage in "binding" of guest molecules by encapsulation within the enclosed internal space (i.e., the cavity). Besides mere physical entrapment, it can be easily envisioned that additional attractive forces (e.g., hydrophobic interactions, π–π interactions, or weak coordinative interactions) can aid in the selective encapsulation of particular guest molecules. The difference between capsule and cage compounds is somewhat arbitrary, but in general guest molecules can exchange from cage compounds without changing the structure of the cage, whereas for capsules a deformation or partial disassembly is required for exchange. Two types of nanometer-sized modular capsules can be distinguished: the covalent-based capsules and the noncovalent-based capsules

(i.e., self-assembled or supramolecular capsules) [8]. The latter consists of (not necessarily identical) components that, upon self-assembly, lead to the formation of the capsule. Initial research in this area was on covalent capsules such as hemicarcerands, calixarenes, and cyclotriveratrylene-based capsules. For details on molecular capsules and catalysis, we refer to [9]; here, we only discuss a few examples relevant for the current topic.

Otto *et al.* [10] developed macrocycles formed by reversible disulfide covalent bonds, leading to a dynamic mixture of cage receptors. The formation of the constituents of the library is under thermodynamic control; its composition is governed by their relative free energies, molecular recognition of added templates through host–guest interactions induces a shift in the composition of a dynamic library of macrocycles. Indeed, the exposure of two different guests as templates to the library resulted in the amplification of two different hosts that were verified to be good receptor molecules for these specific guests (Figure 4.3) (see also Chapter 3).

Rebek has constructed a variety of molecular capsules that self-assemble on the basis of hydrogen bonds. In a typical approach, concave building blocks are utilized with self-complementary binding motifs. For example, the resorcinarene displayed in Figure 4.4 has been functionalized with imide functional groups and upon dimerization via hydrogen bonds it forms a cylinder-shaped cavity that can encapsulate guests [11]. In a similar manner, glycoluril-based building blocks have been prepared that form capsules by assembly of four of these units. In order to create a dynamic library of capsules seven different building blocks were prepared, with various substituents on the aromatic ring [12]. A mixture of two different building blocks gives rise to the formation of six capsules and seven different monomers can potentially produce 613 different capsules, of which 70 can be distinguished by mass spectrometry (MS). In the presence of a guest only 11 species are observed by MS, indicating the selection of hosts for a specific guest that was used as a template.

The selection of guest from a library is also demonstrated by Rebek *et al.* [13], for which the cylinder-shaped capsule was used (Figure 4.4). The capsule can simultaneously encapsulate two different guests (i.e., selective pairwise recognition). From a mixture of benzene, toluene and *p*-xylene, the capsule almost selectively bound a pair of benzene and *p*-xylene, demonstrating an interesting selection strategy based on occupation of space and interaction between guest molecules. This type of binding event could be relevant for the stabilization of transition states, and analogs thereof, for coupling reactions.

Fujita *et al.* [14] developed an octahedral M_6L_4 capsule (Figure 4.5a) formed by self-assembly of six metal fragments and four tridentate ligands. By using slightly different building blocks, cavities with different shapes and dimensions are formed. Many different guests are bound in these types of cage compounds. As the cage is highly charged it dissolved in aqueous solution and the organic guest molecules are generally forced in by hydrophobic interactions. The cage compound has also been used as reaction vessel for various reactions [14a]. Metal–ligand interactions were also applied by Raymond *et al.* [15] to form tetrahedral

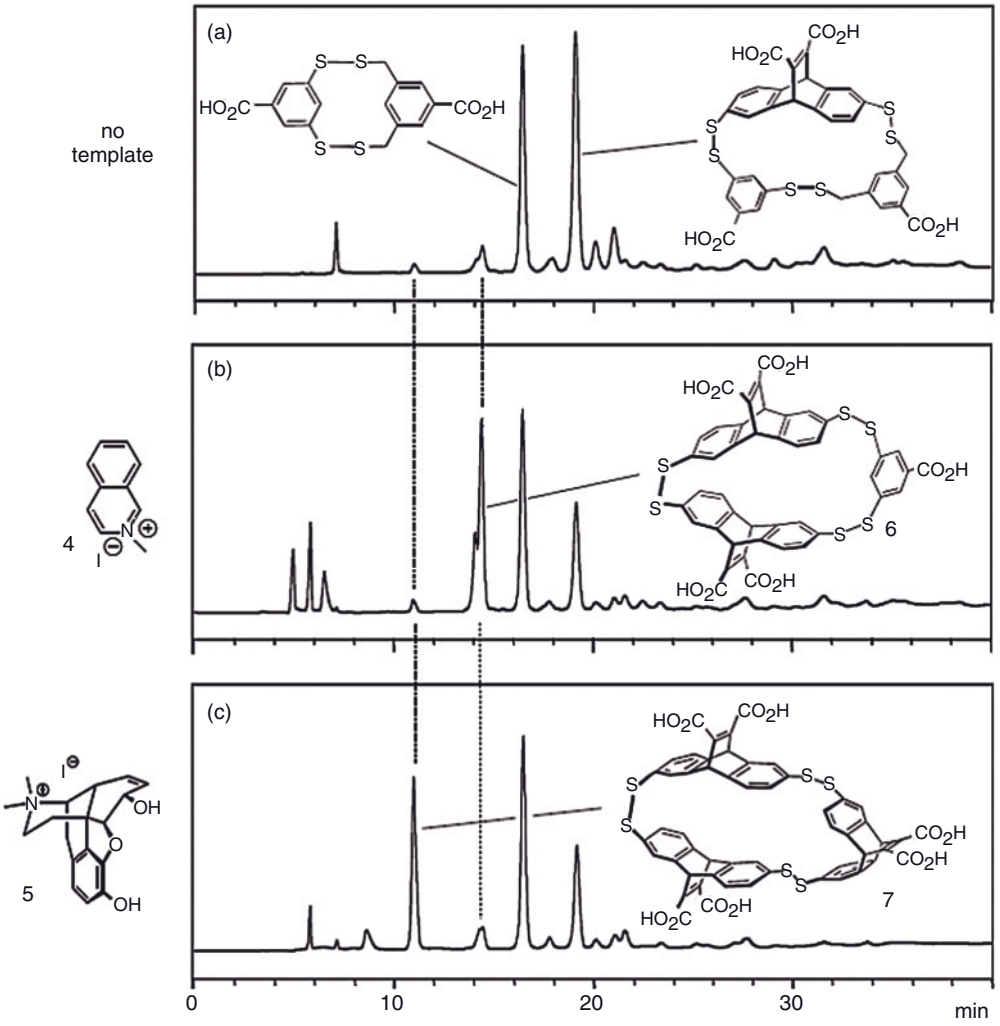


Figure 4.3 High-performance liquid chromatography analyses of the DCL made from dithiols (a) in the absence of any template, (b) in the presence of **4** inducing the amplification of host **6**, and (c) in the presence of morphine derivative **5** leading to the amplification of host **7**.

capsules (Figure 4.5b) using ditopic ligands in combination with several metals that require octahedral coordination. This assembly was successfully used for the dynamic resolution of a pair of enantiomeric guests through encapsulation. For both cages of Fujita and Raymond, many different building blocks have been prepared and in most cases used in pure form. However, they both have demonstrated that the use of a mixture of building blocks leads to a mixture of cage compounds that respond to some extent to the addition of guests, indicating that these systems are suitable for DCC.

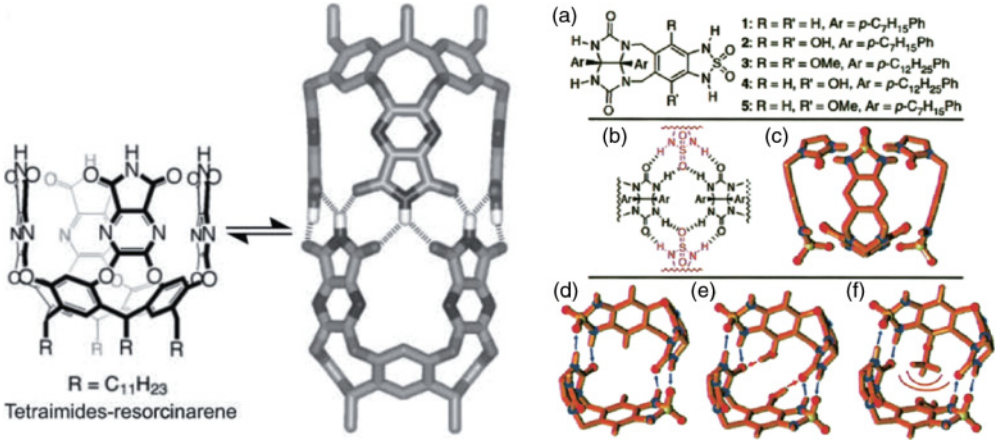


Figure 4.4 Supramolecular hydrogen bonded capsule developed by Rebek *et al.*: (left) a dimeric structure and (right) various tetrameric structures.

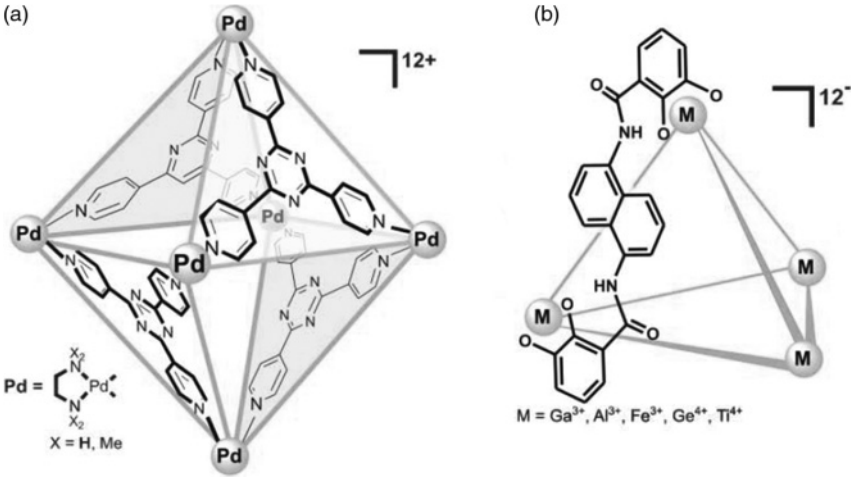
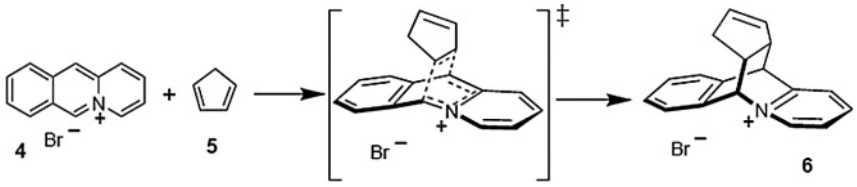


Figure 4.5 Supramolecular capsule formed by metal–ligand interactions developed by (a) Fujita *et al.* [14] and (b) by Raymond *et al.* [15].

4.2.2
Catalysis with Cage Compounds and Possible Selection Procedures

A wide array of self-assembled molecular capsules based on various building blocks and noncovalent interactions has been developed in the last decade. The nanospace within these supramolecular capsules is generally in the range of 300–500 Å^3, which is sufficient for the selective encapsulation of one large or a number of smaller molecules. The structure of the different capsules varies

significantly, and as a result guest shielding and guest exchange rates strongly depend on the capsule applied. In addition, a number of open cage compounds have been prepared and used as catalyst. A diversity of chemical processes has been carried out within molecular capsules and the effects observed so far are, although academic, very interesting. Reactions can be accelerated and the selectivity of a chemical process can be changed completely. These observations can be explained by stabilization of the reaction transition state by the capsule (based on enthalpic and entropic contributions) or by concentration effects in the case of bimolecular reactions, such as Diels–Alder reactions. More important are the unique reaction selectivities induced by the novel finite microenvironment within the capsule. The size and shape of the nanoreactor's cavity and that of the nanoreactor's gates can control the substrate selectivity by controlling the access to the cavity. In the same manner it can protect an active site located in the cavity that otherwise would be poisoned by chemicals present in solution. The regio- and chemoselectivities can also be changed by the capsule by changing the ratio of reaction rates of competing pathways. Product inhibition is a frequently encountered problem in bimolecular coupling reactions carried out within enclosed cavities. The coupling product might have a higher affinity for the capsule than the substrates, and consequently product release from the nanoreactor becomes the slowest step in the reaction. Product inhibition can prohibit the utility of nanoreactors as true catalysts. The capsules and cages displayed in this chapter have been successfully used as cage catalyst for Diels–Alder reactions [14], dipolar cycloaddition reactions, orthoformate hydrolysis [16], and 3-aza Cope rearrangement [17]. In chemical transformations such as Diels–Alder reactions and dipolar cycloaddition reactions, the transition state is similar to the final product, so is considered to be a late transition state (Scheme 4.1).



Scheme 4.1 Diels–Alder reaction demonstrating the similarity between transition state and product.

The Diels–Alder reaction between acridizinium bromide and cyclopentadiene is typically catalyzed by cage compounds. In a seminal paper by Otto *et al.* [18], selection of catalysts was performed using the reaction product as a suitable TSA to select macrocycles from a dynamic library. Exposure of the dynamic combinatorial library (DCLs) based on dithiol building blocks to the product (as TSA) leads to the selection and the amplification of two hosts among all the constituents of the dynamic library (Figure 4.6). The selected cage compounds were applied as catalysts in separate experiments and, indeed, compound **7** was demonstrated to catalyze the Diels–Alder reaction between the two substrates. The reaction rate was

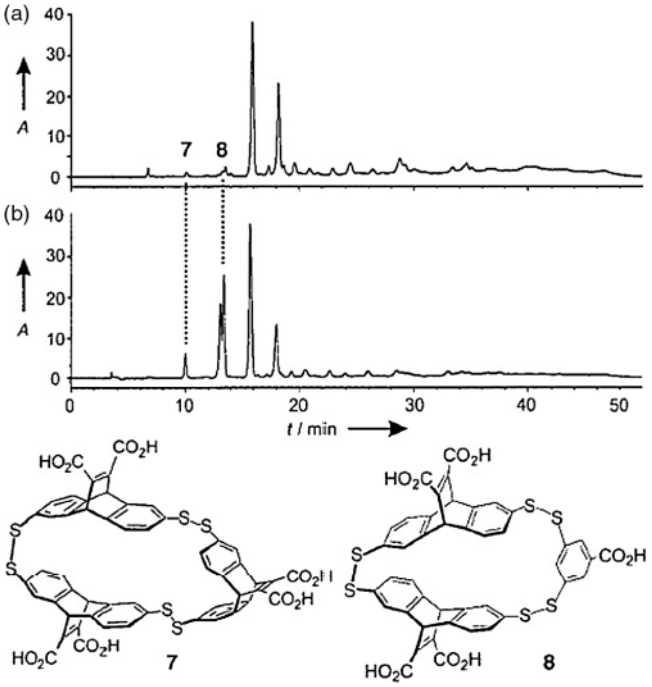
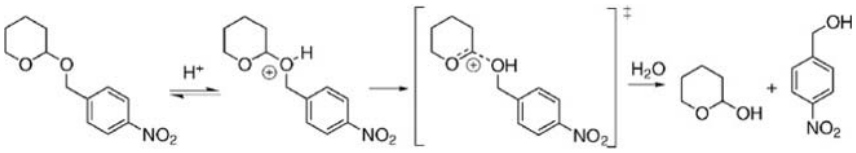


Figure 4.6 High-performance liquid chromatography analyses of the DCL made from dithiols: (a) in the absence of any template and (b) in the presence of TSA **6**, which induced the amplification of macrocycles **7** and **8** (as mixtures of stereoisomers).

increased by a factor of 10. Since the selection procedure was performed with the product, the cage compounds also have affinity for the product that is formed, potentially leading to catalyst inhibition. Indeed, if the reaction is carried out in the presence of the product, it turns out to be slower. The cage compound, however, still gave turnover, indicating that the product did not block the cavity for subsequent reactions as it was displaced by the substrates.

In another example developed by Otto *et al.* [19], macrocycles were selected from a DCL to be used as catalysts for the acetal hydrolysis reaction (Scheme 4.2). The acceleration of the hydrolysis reaction was observed in the presence of the selected macrocycle. However, it remains uncertain whether the reaction acceleration is due to the stabilization of the transition state, to the shift of the pre-equilibrium towards the protonated acetal, or a combination of both effects.



Scheme 4.2 Pathway for the hydrolysis reaction.

4.3
Dynamic Combinatorial Approaches to Transition Metal Catalysts

Transition metal catalysts generally operate via elementary steps that can occur at the metal center, typically oxidative addition, substrate coordination, migration (insertion), reductive elimination, and so on. This creates a reaction pathway by breaking down the reaction in several different steps. As a typical example, the reaction pathway of rhodium-catalyzed hydrogenation is depicted. The reaction profile of such a metal-catalyzed reaction therefore consists of various transition states and intermediates (Figure 4.7). Often one of the transition states has the highest energy barrier and represents the most difficult step, and it is this step of

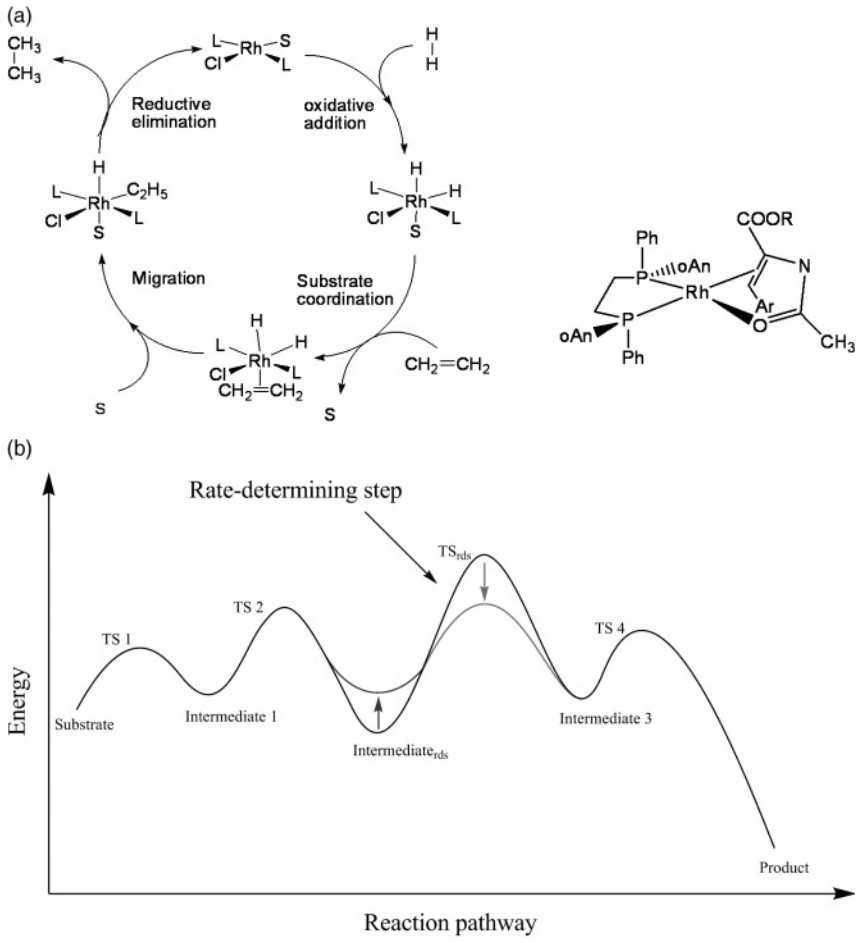


Figure 4.7 (a) General mechanism of the rhodium-catalyzed hydrogenation and typical metal complex that represents an intermediate (resting state) of the reaction and (b) general reaction pathway in transition metal catalysis.

a catalytic cycle that should be accelerated to get a faster catalyst. Typically, ligands are designed such that this rate-limiting step is accelerated, generating a catalyst that gives a faster reaction. In the event that a catalyst should be selected from a DCL of catalysts, efforts should be focused on this rate-limiting step.

Activity is only a part of the challenge, as in most reactions the selectivity of a catalytic process is also an important issue. The creation of a chiral center in a selective manner during the catalytic reaction, often referred to as asymmetric catalysis, is generally the most difficult selectivity to achieve. During one of the elementary steps of the catalytic cycle, the chiral information is transferred from the catalyst to the substrate. In the decisive step there are two competing pathways, with associated transition states that afford the (*R*)-product and the (*S*)-product, respectively. The difference in transition state energy is related to the selectivity afforded to the final product; because only relative small energy differences are required to obtain high selectivity (3 kcal/mol gives 99% enantiomeric excess), finding selective catalysts by rational design is generally impossible. A general energy profile related to a selectivity issue in a reaction pathway with a common intermediate to form both enantiomers of the product is pictured in Figure 4.8, where the (*S*)-enantiomer will be predominantly formed as the lowest energy barrier is associated with its formation. The aim of finding more selective catalysts is related to finding catalysts that have selectively lower energy barriers for this pathway. Currently, most effective approaches consist of combinatorial screening of chiral catalysts and ligand design, guided by knowledge-based intuition. A selection-based process in which a TSA selects the most selective catalyst would provide an interesting complementary approach. It requires the design of TSAs of the enantiodiscriminating step that have different interactions with the (*R*) and (*S*) chiral ligands.

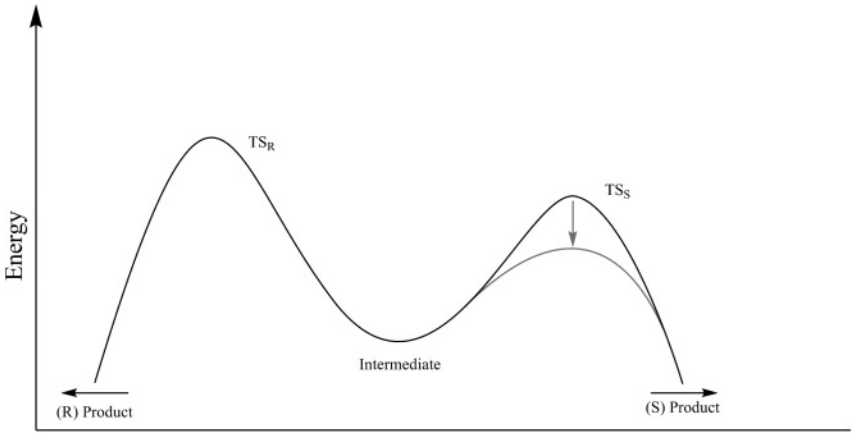


Figure 4.8 Competing reaction pathway that leads to the formation of the two enantiomeric products.

4.3.1 Dynamic Libraries of Transition Metal Catalysts

Over the years many different ligands have been explored in (asymmetric) transition metal catalysis. A logical strategy to select the most active or the most selective catalyst from a dynamic library of catalysts would feature the metal fragment as a part of the TSA, with which a DCL of ligands can be screened, as it is the ligand that creates the selectivity. The use of various ligands will be discussed briefly below. Depending on the kind of possible interactions involved in the dynamic process, the libraries present different advantages and drawbacks (see Table 4.1).

4.3.1.1 Library of Monodentate and Covalent Bidentate Ligands

Many different ligands are now commercially available or easily accessible in few synthetic steps. The achiral phosphine and diphosphine ligands, for example, have been widely and successfully used in homogeneous catalysis [20] such as hydrogenation, allylic substitution, and cross-coupling reactions. Libraries of these ligands can be used to select more active catalysts. No chemical transformations are involved in the selection process, but only exchange of ligands at the metal center that represents the TSA. The major drawback of selection processes in these libraries is the possible multiple coordination as the phosphorous coordination sites are in large excess compared to the metal center, limiting the size of the dynamic library that can be used. The same holds for chiral versions of those libraries, like BINAP, DIPAMP, or DIOP derivatives, which can be used for selection of the selective catalysts. Other classes of bidentate ligands such as P–N, P–O or P–S ligands or (bi)pyridine might also be screened using this approach.

4.3.1.2 Library of Supramolecular Bidentate Ligands

The class of supramolecular bidentate ligands, which was recently introduced (for reviews, see [21a–c]; for recent articles, see [21d–h]), comprises the use of functionalized ligand building blocks that form bidentate ligands by supramolecular interactions between the building blocks upon coordination to the metal center. This class of ligands is considered as an important breakthrough in homogeneous catalysis as the building blocks are generally easily accessible and the ligand library grows exponentially with the number of building blocks, which is ideal for combinatorial approaches. So far mainly phosphorous-based supramolecular bidentate ligands have been developed, using various supramolecular interactions: metal–ligand coordination, electrostatic interactions, and hydrogen bond interactions. If these libraries are to be applied in selection procedures, the exchange is through ligand exchange at the metal center and no chemical transformation is involved. Again, multiple coordination can be expected when using a large excess of ligands (with the advantage that a smaller number of building blocks already result in many combinations) and also the functional groups may interfere with the metal center.

4.3.1.3 Library of Dynamic Supramolecular Templates

Supramolecular bidentate ligands can also be constructed by using templates onto which the functionalized ligand building blocks are associated. These ligands seem ideal for selection procedures as a large excess of ligands is not required as the variation can be sought in the template. As such, the problem of over-coordination can be prevented. So far, only a few examples of such template-based ligands have been reported.

4.3.1.4 Library of Dynamic Covalent Linkers

The best ligands for selection procedures are those that have separated the donor atoms for coordination and functional groups for modification. The dynamic exchange does not involve the metal center, avoiding possible decomposition during the exchange and also over-coordination is prevented. The selection procedure is based on steric modifications of either (i) the (a)chiral linker and/or its chain length or (ii) the structure of the ligand. Numerous building blocks are in principle suited, such as (a)chiral diamine in combination with aldehyde-functionalized ligands. The formation of imine is known to be reversible under specific conditions. In principle, all the dynamic covalent chemistry discussed in chapter 1.4 and 3.2.1.1 should be applicable, as long as the exchange chemistry is compatible with the coordination chemistry. Although some scattered examples of these types of ligands can be found in literature, they have not been developed yet with the current application as a goal.

4.3.2 Selection Procedures via Intermediates and TSAs

The ideal template for the selection of the best catalyst is the transition state, which is by definition unstable and therefore not useful. Therefore, a TSA should be designed that, on the one hand, is sufficiently precise to mimic the transition state but, on the other hand, is sufficiently stable to be used for the selection procedure. Recent progress in molecular modeling and the ever-increasing computer power facilitates the identification of the transition state and for the design of analogs thereof.

An alternative strategy is to employ selection procedures for identifying the most unstable intermediate ($Intermediate_{rds}$, Figure 4.7b). By increasing the energy of this complex, the overall energy barrier will decrease and the reaction rate will increase. This strategy will only work for selection based on reaction rate and not on selectivity (unless the situation is more complicated with more intermediates that form different products). Importantly, new experimental tools (based on nuclear magnetic resonance (NMR), infrared, MS, etc.) and analytical techniques (kinetics, etc.) are continuously being developed, which will facilitate the identification of different intermediates. The intermediate targeted (prior to the rate-determining step) is generally accumulated in the reaction, its chemical transformation being the slowest of the reaction pathway.

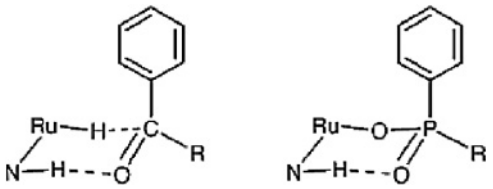


Figure 4.9 Organometallic TSAs for transfer hydrogenation reactions as developed by Polborn and Severin [23] for molecular imprinting.

4.3.2.1 **TSA Approach**

No successful example has been reported so far using a TSA in a dynamic combinatorial approach to transition metal catalyst selection. However, inspired by enzymes and molecular cages, molecularly imprinted polymers were successfully developed by Wulff *et al.* and in a small number of cases directed towards transition metal catalysis [22]. Cavities as biomimetic catalysts are created by generation of polymeric materials in the presence of a TSA as a template, which is removed after polymerization. In the presence of the substrate, the incorporation of the catalyst precursor leads to high activities, the transition state being stabilized by the polymeric cavities.

Polborn and Severin [23] recently reported ruthenium- and rhodium-based TSAs for the transfer hydrogenation reaction. These complexes were used as catalyst precursors in combination with molecular imprinting techniques. Phosphinato complexes were prepared as analogs for the ketone-associated complex. They demonstrated that the results obtained in catalysis were better in terms of selectivity and activity when these TSAs were imprinted in the polymer. This shows that organometallic complexes can indeed serve as stable TSAs (Figure 4.9).

4.3.2.2 **Selection of Catalyst Based on Intermediate Stability**

We recently studied if it is possible to device a selection strategy based on the relative stability of the intermediate of a reaction [24]. It is known that in the palladium-catalyzed allylic substitution, the rate-determining step is the attack of the nucleophile on the π-allyl-palladium species. The transition state of this step is believed to be late when carbon nucleophiles are used. In this scenario, an inverse correlation of the energy of the intermediate and the reaction rate is expected, as the transition state is more product-like (see Figure 4.10). Based on this hypothesis, the selection of catalyst among a dynamic mixture of palladium complexes was studied.

For the selection experiments, a homolog library of diphosphine ligands, which form the Pd-allyl intermediate complexes, was used. We specifically wanted to investigate the effect of the bite angle and steric properties of the aryl groups, to which end we decided to use dppe, dppp, dppb, and their corresponding *o*-tolyl-phosphino analogs. In a typical selection experiment, 1 equiv. of [Pd(crotyl)(dppe)] OTf was mixed in dichloromethane with equimolar amounts of the other free ligands (i.e., dppp and dppb) to generate a dynamic library of intermediates.

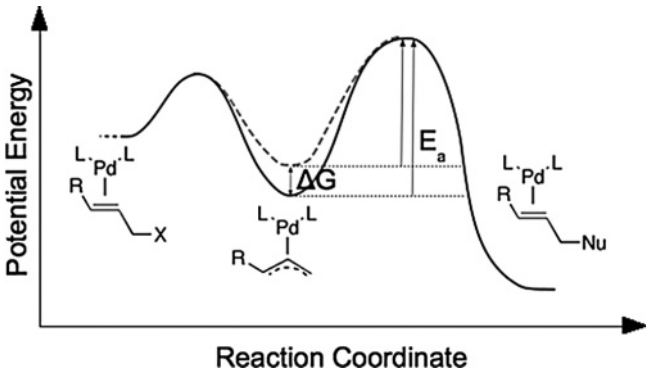
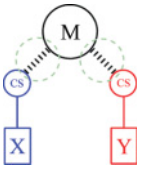
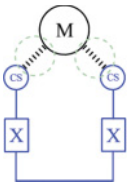
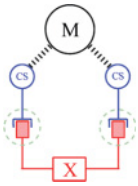
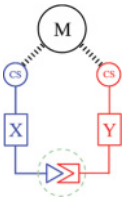
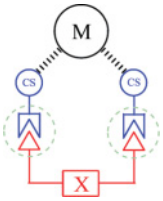


Figure 4.10 Selection procedure by intermediate destabilization in the palladium-catalyzed allylic substitution reaction.

Ligand exchange was shown to be fast by ^{31}P-NMR and after 1 h the mixture was analyzed by electrospray ionization (ESI)-MS. As there is only one equivalent of metal ion present, the ligand that forms the most stable Pd-allyl complex will compete most effectively for coordination and the corresponding complex will be most abundant in the mixture. Therefore, the abundance of the complexes in the ESI-MS spectrum is a direct reflection of the stability of the corresponding species. In order to confirm that a thermodynamic controlled product mixture was obtained, we performed the experiments also starting from the dppp and dppb Pd-allyl complexes, which resulted in identical spectra. Furthermore, to be sure the intensities can be correlated to concentrations of the observed species, we made calibration curves by injecting mixtures of preformed diphosphine Pd-allyl complexes at different ratios. The ESI-MS spectra display signals corresponding to all three Pd-allyl complexes in case of both phenyl and *o*-tolyl substituents on phosphorous. The complexes with the ethane backbone are most abundant followed by the propane and butane backbone. This trend suggests that dppb should give the fastest allylic alkylation catalyst and that the rate of the reaction increases with increasing bite angle in the investigated series of ligands. This was indeed what was observed experimentally, and the nice inverse correlation between the abundance in the MS spectrum and the activity demonstrates the principle of this selection procedure.

Table 4.1 Schematic representation of varions possible catalyst libraries that could be applied for ligand selection, including advantages and disadvantages.

Library	Advantages	Drawbacks
Monodentate ligands 	library commercially available; easy synthesis; no chemical transformations involved in the dynamic process	over-coordination with large excess of ligand
Bidentate covalent ligands 	library commercially available; no chemical transformations involved in the dynamic process	over-coordination with large excess of ligand; tedious synthesis (unless C_2-symmetric)
Covalent linkers 	library commercially available (chiral or achiral spacer, etc.)	chemical transformations involved in the dynamic process, possible interference with the metal center or the ligands
Supramolecular bidentate ligands 	no chemical transformations involved in the dynamic process; supramolecular interaction under thermodynamic control; convergent synthesis of building blocks	possible interference of functional groups with the metal center; over-coordination with large excess of ligand
Supramolecular templates 	no chemical transformations involved in the dynamic process; supramolecular interaction under thermodynamic control; convergent synthesis of building blocks	possible interference of functional groups with the metal center; limited examples reported

M, Metal center; , bond involved in the dynamic exchange; CS, coordination site (P, N, O, S, etc.); X Y, diversity of building blocks in the constituents of the dynamic library; , dynamic covalent interaction (disulfide, imine, hemi-ketal, amide, ester, etc.); , dynamic supramolecular interaction (hydrogen bond, metal–ligand interaction, ionic bond, etc.).

4.4 Conclusions

The field of dynamic combinatorial catalysis is virtually open. The first few examples, however, provided first proof-of-principle that also in this area DCC has added value. For both cage-type catalysts as well as transition metal complexes the DCC approach could have significant advantages over rational design or traditional combinatorial strategies, but at this stage it is far too early to compare these approaches. In this chapter, we mainly focus on the principles, supported by some initial results, in the hope that we will stimulate scientists to continue research in this exciting new area.

References

1 Corbett, P.T., Leclaire, J., Vial, L., West, K.R., Wietor, J.L., Sanders, J.K.M. and Otto, S. (2006) *Chem. Rev.*, **106**, 3652–3711.

2 (a) Sanders, J.K.M. (1998) *Chem. Eur. J.*, **4**, 1378–1383.
(b) Kovbasyuk, L. and Krämer, R. (2004) *Chem. Rev.*, **104**, 3161–3188.
(c) Motherwell, W.B., Bingham, M.J., and Six, Y. (2001) *Tetrahedron*, **57**, 4663–4686.

3 Crabtree, R.H. (2005) *The Organometallic Chemistry of the Transition Metals*, Wiley-VCH Verlag GmbH, Weinheim.

4 (a) Marcus, R.A. (1993) *Angew. Chem. Int. Ed. Engl.*, **32**, 1111–1121.
(b) Marcus, R.A. (1960) *Disc. Faraday Soc.*, **29**, 21.

5 Ringe, D. and Petsko, G.A. (2008) *Science*, **320**, 1428–1429.

6 Pauling, L. (1948) *Am. Sci.*, **36**, 51–58.

7 Koblenz, T.S., Wassenaar, J., and Reek, J.N.H. (2008) *Chem. Soc. Rev.*, **37**, 247–262.

8 (a) Conn, M.M. and Rebek, J., Jr. (1997) *Chem. Rev.*, **97**, 1647–1668.
(b) Rebek, J., Jr. (1996) *Chem. Soc. Rev.*, **25**, 255–264.
(c) Hof, F., Craig, S.L., Nuckolls, C., and Rebek, J., Jr. (2002) *Angew. Chem. Int. Ed.*, **41**, 1488–1508.

9 (a) Vriezema, D.M., Aragonès, M.Comellas, Elemans, J.A.A.W., Cornelissen, J.J.L.M., Rowan, A.E., and Nolte, R.J.M. (2005) *Chem. Rev.*, **105**, 1445–1489.
(b) Lützen, A. (2005) *Angew. Chem. Int. Ed.*, **44**, 1000–1002.
(c) Yoshizawa, M., Klosterman, J.K., and Fujita, M. (2009) *Angew. Chem. Int. Ed.*, **48**, 3418–3438.

10 Otto, S., Furlan, R.L.E., and Sanders, J.K.M. (2002) *Science*, **297**, 590–593.

11 Rebek, J., Jr. (2005) *Angew. Chem. Int. Ed.*, **44**, 2068–2078.

12 Hof, F., Nuckolls, C., and Rebek, J., Jr. (2000) *J. Am. Chem. Soc.*, **122**, 4251–4252.

13 Heinz, T., Rudkevich, D.M., and Rebek, J., Jr. (1998) *Nature*, **394**, 764–766.

14 (a) Yoshizawa, M., Tamura, M., and Fujita, M. (2006) *Science*, **312**, 251–254.
(b) Yu, S.Y., Kusukawa, T., Biradha, K., and Fujita, M. (2000) *J. Am. Chem. Soc.*, **122**, 2665–2666.
(c) Fujita, M., Oguro, D., Miyazawa, M., Oka, H., Yamaguchi, K., and Ogura, K. (1995) *Nature*, **378**, 469–471.

15 Fiedler, D., Leung, D.H., Bergman, R.G., and Raymond, K.N. (2004) *J. Am. Chem. Soc.*, **126 (12)**, 3674–3675.

16 Pluth, M.D., Bergman, R.G., and Raymond, K.N. (2007) *Science*, **316**, 85–88.

17 Fiedler, D., van Halbeek, H., Bergman, R.G., and Raymond, K.N. (2006) *J. Am. Chem. Soc.*, **128**, 10240–10252.

18 Brisig, B., Sanders, J.K.M., and Otto, S. (2003) *Angew. Chem. Int. Ed.*, **42**, 1270–1273.

19 Vial, L., Sanders, J.K.M., and Otto, S. (2005) *New J. Chem.*, **29**, 1001–1003.

20 van Leeuwen, P.W.N.M. (2004) *Homogeneous Catalysis: Understanding the Art*, Kluwer, Dordrecht.

21 (a) Wilkinson, M.J., van Leeuwen, P.W.N.M., and Reek, J.N.H. (2005) *Org. Biomol. Chem.*, **3**, 2371–2383.
(b) Breit, B. (2005) *Angew. Chem. Int. Ed.*, **44**, 6816–6825.
(c) Sandee, A.J. and Reek, J.N.H. (2006) *Dalton Trans.*, 3385–3391.
(d) Breuil, P.A.R., Patureau, F.V., and Reek, J.N.H. (2009) *Angew. Chem. Int. Ed.*, **48**, 2162–2165.
(e) Gulyás, H., Benet-Buchholz, J., Escudero-Adan, E.C., Freixa, Z., and van Leeuwen, P.W.N.M. (2007) *Chem. Eur. J.*, **13**, 3424–3430.
(f) Jiang, X.B., Lefort, L., Goudriaan, P.E., de Vries, A.H.M., van Leeuwen, P.W.N.M., de Vries, J.G., and Reek, J.N.H. (2006) *Angew. Chem. Int. Ed.*, **45**, 1223–1227.
(g) Birkholz, M.N., Dubrovina, N.V., Jiao, H., Michalik, D., Holz, J., Paciello, R., Breit, B., and Börner, A. (2007) *Chem. Eur. J.*, **13**, 5896–5907.
(h) Sandee, A.J., van der Burg, A.M., and Reek, J.N.H. (2007) *Chem. Commun.*, 864–866.

22 Liu, J.Q. and Wulff, G. (2004) *J. Am. Chem. Soc.*, **126**, 7452–7453.

23 (a) Polborn, K. and Severin, K. (1999) *Chem. Commun.*, 2481–2482.
(b) Polborn, K. and Severin, K. (2000) *Eur. J. Inorg. Chem.*, 1687–1692.
(c) Polborn, K. and Severin, K. (2000) *Chem. Eur. J.*, **6**, 4604–4611.

24 Wassenaar, J., Jansen, E., van Zeist, W.-J., Bickelhaupt, F.M., Siegler, M.A., Spek, A.L., and Reek, J.N.H., *Nature Chemistry* (2010), in press.

5
Dynamic Combinatorial Chemistry: Ligands for Biomolecules

Olof Ramström, Luis Amorim, Rémi Caraballo, and Oscar Norberg

5.1
Ligand Discovery

5.1.1
High-Throughput Chemistry

Already proposed in the early 1960s [1], although more pronouncedly formulated in the 1980s [2], combinatorial chemistry and related strategies have matured to become important instruments in chemistry and biology, in general, and in ligand discovery processes, in particular [3–6]. Combinatorial approaches have thus become a potent technology for creating collections of structurally related compounds that may subsequently be screened for desirable effects and properties. Initially developed to produce peptide libraries for screening against antibodies or receptors as candidates of optimal binding properties, numerous methodologies have been designed to produce chemical combinatorial libraries. Following an almost explosive development in the 1990s, especially in the pharmaceutical domain, the field evolved substantially from the development of techniques for the generation and analysis of very large compound collections to more selective approaches for the identification of lead structures. This evolution was in part driven by the lack of immediate successes in the discovery campaigns [7], where the combinatorial libraries employed often created immense datasets of relatively low quality with respect to diversity and fidelity [8, 9]. However, despite all criticism raised, combinatorial chemistry still occupies a central role in drug discovery, where speed remains of paramount importance in lead generation [10]. Furthermore, in parallel with chemical technology development, automation techniques and rapid analysis instrumentation have at the same time experienced major advancements that are of high importance in the present-day discovery processes. These developments have led to more advanced high-throughput chemistry (HTC) strategies, and such approaches are also gaining importance in the development of synthetic processes, in catalyst discovery, and in materials science.

Over time, chemical combinatorial libraries have been developed primarily using different formats of parallel syntheses, where the stepwise formation of

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

individual compounds is achieved in parallel. This can be done in discrete compartments, essentially resulting in small-scale traditional synthetic schemes run with a multitude of selected starting materials, or in the same compartment as long as the individual compounds can be easily distinguished and handled. Many protocols have for this reason implemented resin-based chemistry, where the compounds are immobilized on solid supports, each addressable using a variety of tagging methods. The so called split-mix methodologies, originally proposed by Furka [2, 11], can be used to create very large libraries progressively over several synthetic steps. Solution-phase, nonimmobilized libraries were originally explored (e.g., using multicomponent reactions), but the use of pools of discrete soluble substances in the same compartment has not experienced a rapid progress, due to difficulties in hit identification with such systems. The compounds are instead usually prepared as discrete, stable entities, virtually insensitive to changes in the environment and all compounds are subsequently screened one at a time or in small sets using automated systems towards a chosen biological target.

5.1.2
Dynamic Target Assistance

The rapid development of different combinatorial and high-throughput techniques has allowed the efficient synthesis of large arrays of compounds in a short time, and large and diverse compound collections have been prepared using a range of different structures. Nevertheless, for new collections each individual compound needs to be prepared over several synthetic steps and to be sufficiently characterized in order to avoid ambiguities in the subsequent activity screening. Optimization of the hit-rate is of high importance, and reduction of the number of false-positives and false-negatives is of significant concern. Modern automated high-throughput techniques applied to synthesis and analysis have enabled the development of such processes, where compound collections can be synthesized and screened routinely.

Identification of new ligands by high-throughput methods is in a way a trial-and-error type of strategy, albeit supported by computer modeling and *in silico* screening approaches. A more straightforward way would be to allow the target entity *per se* to select an active ligand directly from a library pool. With this scenario, the screening process would be greatly simplified and more efficient, and in principle only result in more or less potent ligands. In addition, with a self-regulating library pool, able to undergo constitutional changes during the recognition process and adapting itself to the system constraints, then the screening signal could be dramatically enhanced. This effect would lead to facilitated detection and characterization and minimize the background effects.

This more direct strategy of target-assisted library screening forms the basis of dynamic combinatorial chemistry (DCC), constituting a new paradigm in drug discovery. In contrast to the regular static approaches to combinatorial chemistry, the library can instead be produced from a set of reversibly interchanging components that maintain dynamics of the entire system. Dynamic combinatorial libraries (DCLs) are thus generated, where each library member affects, and is affected

by, all other surrounding constituents and components. The whole approach is in principle of supramolecular nature, being driven by the interactions of the library constituents with the target site, and relies on reversible reactions or interactions between sets of basic components to generate continually interchanging adducts. This gives access to "virtual" compound collections whose potentially accessible constituents are all possible, latent combinations of the components available.

5.1.3
Dynamic Fragment-Based Ligand Discovery

The DCC principle has been reviewed in several other chapters of this book, notably Chapters 1 and 2. In the following, only a very brief description will be given for clarity, with emphasis on biological systems. DCC is in principle based on fragment assembly in generating (supra)molecular assemblies of different size and nature. In general fragment-based discovery [12–14], these fragments are assembled one-by-one and tested individually, and sophisticated analytical and preparative methods have been developed for this purpose. The use of complex molecular cocktails, containing a multitude of fragments, generally results in analytical problems and these are for this reason less used. The dynamic characteristics of DCC systems, however, reduce the complexity through adaptive reconstitution, thus resulting in simplified processing. The adaptability of the systems, enabled by the reversibility of the processes, also confers its potentiality to be amplified according to Le Châtelier's principle. If one constituent of the DCL shows stronger interaction with a certain target entity than all others, this constituent will be withdrawn from the equilibrating pool and all components making up this constituent will also be masked by the interaction. In some instances the library constituents are only detectable in the presence of the selector, resulting in what may be termed "virtual combinatorial libraries."

In principle, DCC-based discovery relies on the following general steps:

1) Selection of dynamic chemistry and library fragments.
2) Establishment of dynamic library generation conditions where the fragments are allowed to form interchanging, individual molecular entities.
3) Exposing the dynamic library to target selection at the given conditions.

Generation and control of the dynamic system is essential for the entire process to work, requiring optimization and judicious selection of involved structures and conditions in response to the target species in question. Obviously, these basic steps have been adapted and extended in a variety of different related approaches.

5.2
DCC Strategies in Targeting Biological Systems

Originating from the original concept of thermodynamically controlled DCC systems, several strategies of DCL generation and screening have been developed in response to different challenges. The DCC concept has thus become extended

to a variety of useful approaches for applications in biological systems. All of these strategies are based on reversible fragment combination as a common feature, but differ in the screening/selection phase. Furthermore, a range of related methods have been demonstrated, not always relying on dynamic features. In the following section, most of these methods are schematically described. We propose the use of the double-arrow symbol to describe a dynamic system in general and a dynamic library in particular (Figure 5.1) [15].

5.2.1 Stoichiometric DCLs

The original DCC concept addressed the potential of generating the DCLs in the presence of a receptor (or a ligand) in the same compartment (Figure 5.2). The overall system is then under thermodynamic control, and library adaptation and potential amplification can be obtained. Using the equilibrium distribution of the DCL in absence of any target as a reference, all changes upon its addition are then monitored. Following Le Châtelier's principle, the entire system has to rearrange so as to produce more of the best fragment combinations at the expense of all other assemblies in the library. Upon re-equilibration, the constituent with the highest affinity will thus experience a certain degree of amplification in comparison to the situation where no target molecule was added. In order to achieve

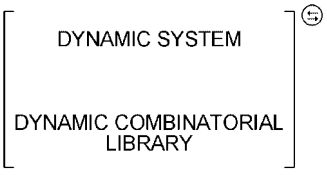


Figure 5.1 Description of dynamic systems/DCLs. All members of the system/libraries are in constant reversible exchange with each other.

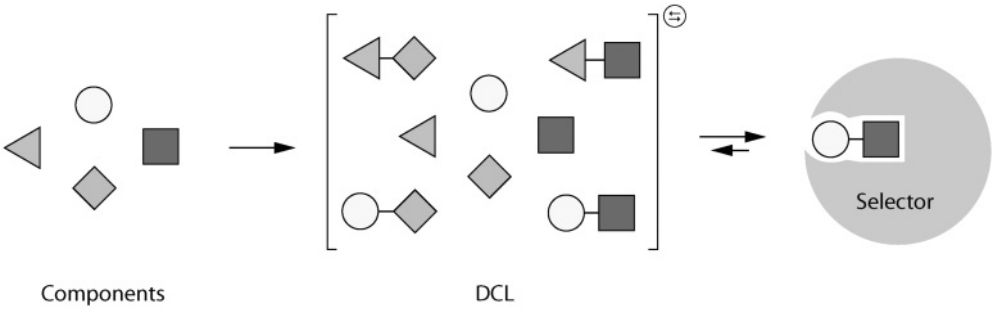


Figure 5.2 Stoichiometric DCL format.

high amplification effects, stoichiometric amounts of targets are needed in the systems.

5.2.2
Postmodified DCLs – Dynamic Reductive Amination

As will be discussed further below, imines are often employed in creating DCLs. In aqueous systems, the exchange between the involved building blocks is generally fast and to some extent controlled by pH. However, the formed imines are normally kinetically unstable entities, and cannot easily be isolated and studied. For this reason, a strategy often used in HTC strategies – reductive amination – has been employed also in DCC systems. Thus, the formed imines are subjected to reduction using cyanoborohydride species to generate stable amines – a process that can be done *in situ* during DCL generation and screening (Figure 5.3). This is a type of postmodification approach, where the active species is further processed for ease of handling. Advantages with this method are the improved stability of the final hits and the straightforward mode of operation. Disadvantages are the generation of species that are not initially selected by the selector and may not have the same effect. This may, however, also lead to a lower degree of product inhibition and the potential use of substoichiometric amounts of target.

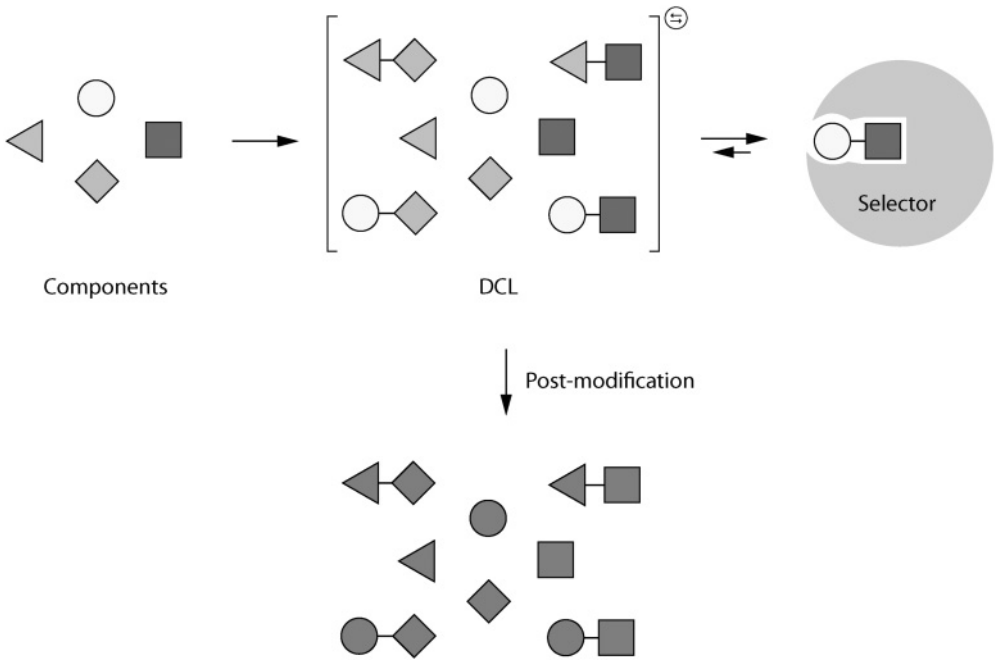


Figure 5.3 Postmodified DCLs.

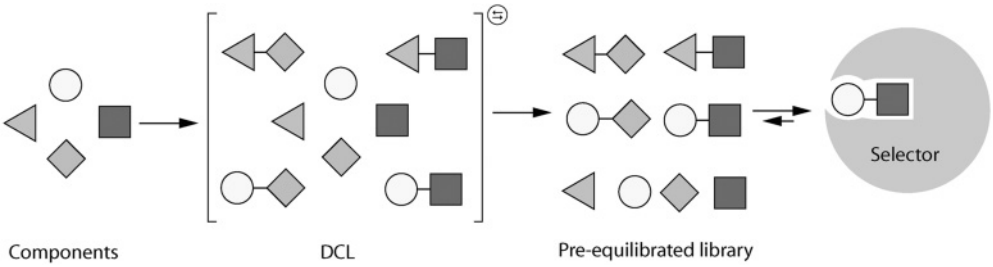


Figure 5.4 Pre-equilibrated DCL generation and screening.

5.2.3
Pre-Equilibrated and Iterated DCLs – Dynamic Deconvolution and Panning

An alternative approach, that in part addresses the issue of avoiding stoichiometric amounts of target, is denoted as pre-equilibrated DCLs. In this case, the generation and screening steps are separated from each other, and the DCLs are formed under dynamic conditions, whereas the identification/screening is performed under static conditions. Stoichiometric amounts of target entity are in this case not required, thus allowing the use of sensitive biological target species that are available in low quantities. Screening may be accomplished using common assay methods and identification of high-affinity fragment combinations can be made using dynamic deconvolution protocols. Owing to the dynamic nature of the library, such deconvolutions may be significantly simplified, enabling rapid identification of the best fragments (Figure 5.4). The pre-equilibration procedure is also highly useful when the dynamic reaction requires equilibration conditions that are incompatible with the biological target.

The pre-equilibration strategy can furthermore be used to achieve ligand amplification when run repeatedly (Figure 5.5). The DCLs are thus generated in one compartment under appropriate dynamic conditions and subsequently allowed to interact with the target species. The binding event can take place either in the same reaction chamber or separately, depending on the conditions. Separation of the unbound fragments from the high-affinity fragment combinations is needed, often using immobilized or entrapped target entities. The unbound species are then retransferred to the reaction chamber, rescrambled, and again allowed to interact with binding site. After several rounds of such a "dynamic panning protocol," the accumulated active species may be easily identified.

5.2.4
Dynamic Combinatorial X-Ray Crystallography

Rapid and reliable identification of the best fragment combination for a given target species is certainly a challenge and methods for instant identity matching are desired. This challenge was addressed using a direct X-ray screen of the DCLs in the presence of the target (Figure 5.6). Protein crystals were exposed to

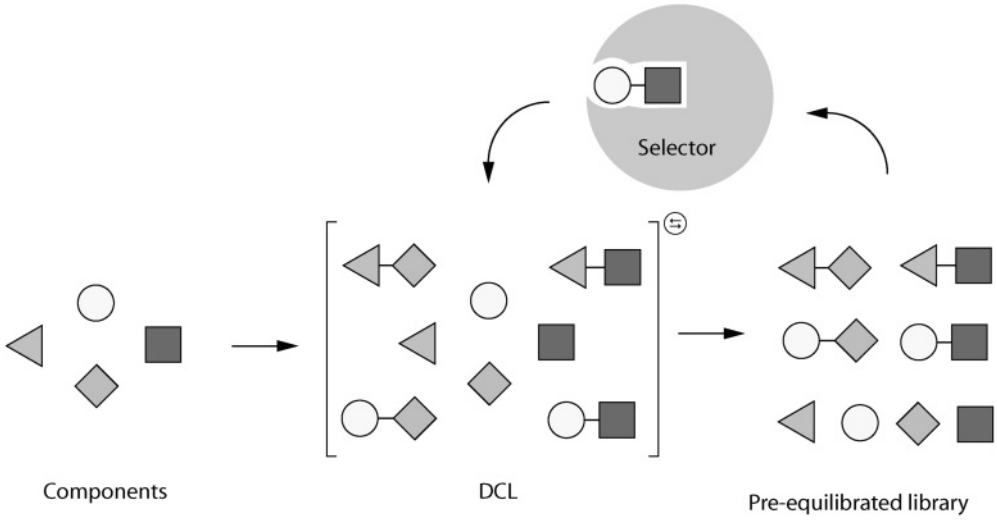


Figure 5.5 Pre-equilibrated DCL generation and iterative screening.

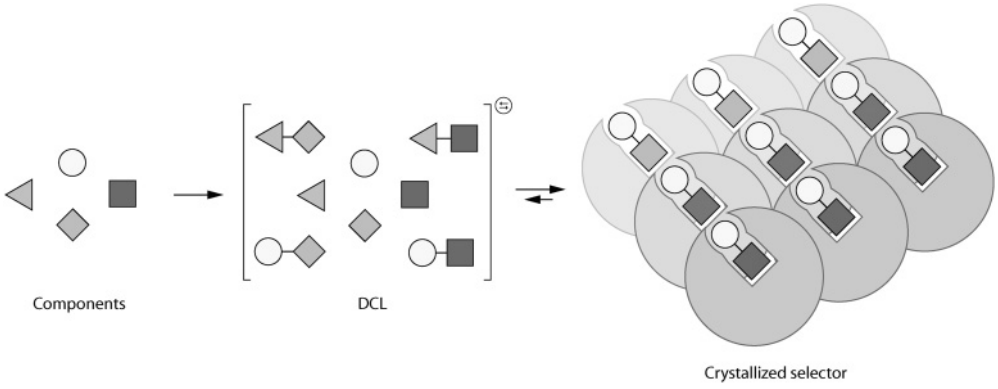


Figure 5.6 DCX.

pre-equilibrated combinatorial libraries in a nondisrupting solution and the best combinations were bound to the protein active sites *in situ* in the crystals. The DCL fragment combinations could be observed directly by X-ray crystallography and interpretation of the resulting electron-density maps yielded information of the most potent ligand. Based on the efficient site-mapping properties of X-ray crystallography, this dynamic combinatorial X-ray crystallography (DCX) method has the advantage of simultaneous ligand identification and ligand quality estimation.

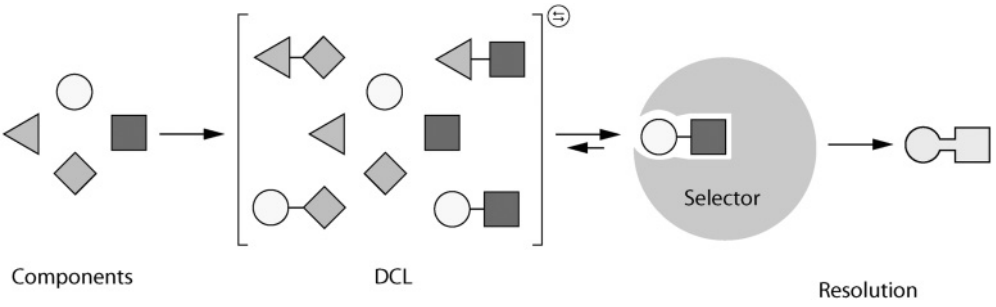


Figure 5.7 DCR.

5.2.5
Dynamic Combinatorial Resolution – Kinetically Resolved DCLs

A highly attractive method for rapid ligand identification is the direct coupling of consecutive irreversible reactions to the thermodynamically controlled DCL system. Generation and binding of fragment combinations are in this case enabled following the original concept, but an irreversible process following binding then results in kinetic resolution of the entire system (Figure 5.7). This dynamic combinatorial resolution (DCR) requires only catalytic amounts of target species, and generally the selection and kinetic resolution is carried out by the same target species. This method results in complete amplification of the best fragment combinations, considerably simplifying the identification process, and is particularly useful for enzyme substrate screening.

5.2.6
Dynamers

Not only ligands can be assembled and identified by the DCC method, but also mimics of biomacromolecules themselves. This development has only more recently been initiated, where reversible linkages are used to form dynamic polymeric structures – so-called "dynamers" – in some cases based on analogs of natural building blocks (Figure 5.8) [16–18]. To date, these dynamers have not been applied to any biological interaction studies other than self-organization, in itself highly interesting, but these materials have the potential to be used in a variety of applications, ranging from slow release to biosensing.

5.2.7
Related Approaches

5.2.7.1 Deletion Approach

Rather than selecting and amplifying the best binder as proposed by the dynamic approaches, an interesting, pseudodynamic deletion alternative has been introduced. In this case, the unbound constituents are deleted from the libraries (Figure

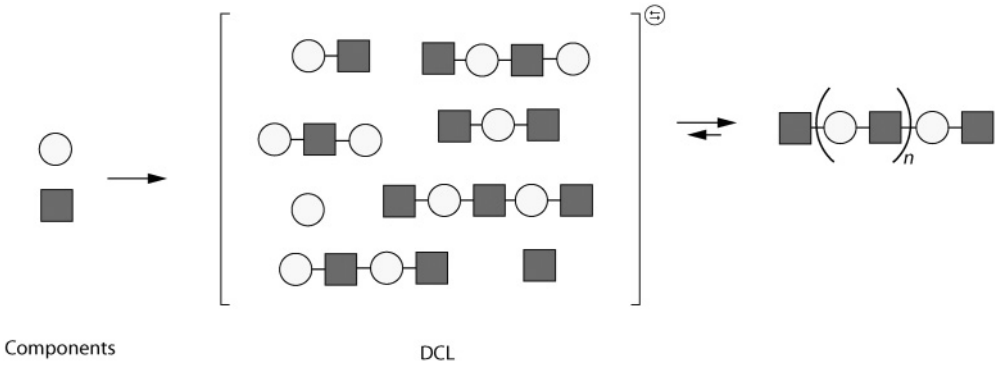


Figure 5.8 Dynamer formation.

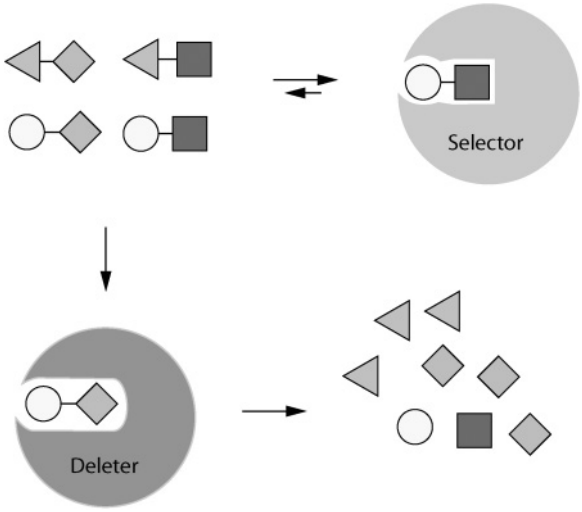


Figure 5.9 Deletion approach.

5.9). The libraries are in this case not dynamic in nature, and with this concept the formation and destruction of the libraries are separate irreversible reactions. Ligand identification is enabled by the kinetic deletion of poorer binders, where binding to the target entity shields the best binders from the deletion process. This results in an increase in the relative ratio of good and poor ligands, and identification is based on the decrease of certain fragment products compared to those building up poorer constituents in a reverse identification methodology. As the name implies, the libraries are in this case not truly dynamic in nature, inasmuch as the constituents are not spontaneously reformed in the process. Nevertheless, an improvement of the concept where the constituents were continuously resynthesized during the deletion process has also been demonstrated.

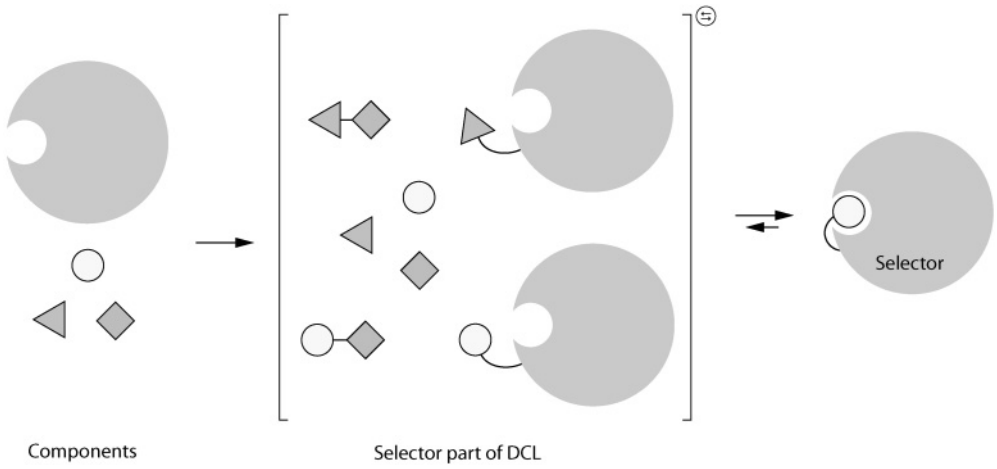


Figure 5.10 Tethering approach.

5.2.7.2 **Tethering Approach**

A highly interesting related technique that is related to DCC is the so-called "tethering" approach (Figure 5.10). In order to screen for fragments of ligands to a specific binding site, the fragments are tethered covalently to the vicinity of the binding site through disulfide bond chemistry and the inhibitory activity of the linked fragment measured. The technique in this case requires a free thiol moiety in close proximity to the binding site and, in cases where a suitable thiol functionality is absent from protein surface, the target generally needs to be engineered to possess a cysteine residue in that position. Obviously this needs to be arranged such as not to interfere with the structure of the protein or the binding. Due to the reversible nature of the thiol–disulfide interchange, the process can also be performed under dynamic conditions and in this case the biological target *per se* is made part of the library. The technique is furthermore especially useful for identifying weakly binding fragments of a potential ligand.

5.2.7.3 **In Situ "Click" Approach**

Another fragment-based approach that makes use of the target species *per se* in the selection process is the synthesis of ligands directly in the target binding site (Figure 5.11), also conceptualized as *in situ* "click" chemistry in using the Huisgen 1,3-dipolar cycloaddition reaction between azides and alkynes. This strategy requires activated fragments that are essentially inert at dilute conditions, but react with each other when adequately positioned together in the binding site. The proximity effect created by hosting the fragments in the binding site thus increases the reaction rate compared to the background reaction in solution. Although being a highly attractive technique, it depends on highly selective reactions that mainly occur with certain reactant pairs and is furthermore sensitive to product inhibition.

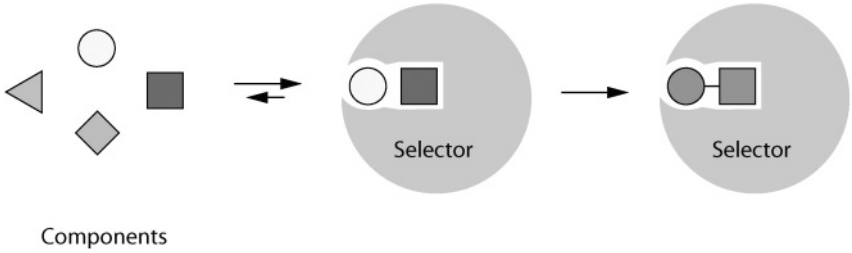


Figure 5.11 *In situ* "click" approach.

5.3
Dynamic Diversity Generation for Biological Systems

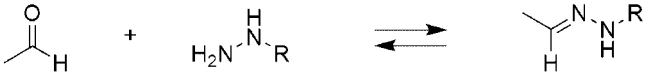
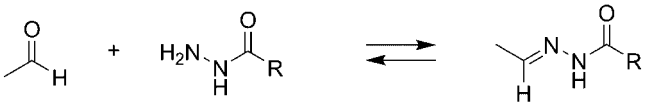
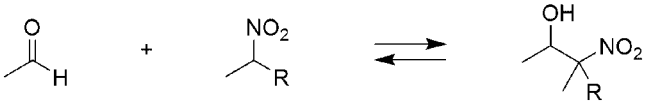
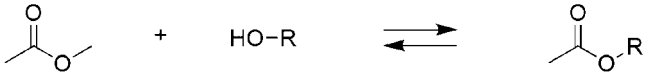
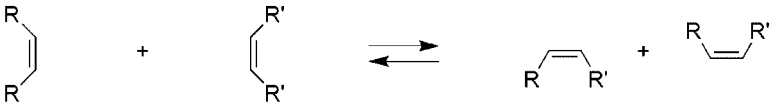
5.3.1
Dynamic Chemistry

The chemistry of DCL formation has been covered in previous chapters in this volume, notably Chapters 1 and 3. Thus, DCLs can be generated using essentially any type of reversible chemical mechanism, provided the interconverting states can be properly controlled and the final products identified. The most important processes involve molecular/supramolecular interchanges, of either reversible covalent or noncovalent character. Functional groups enabling reversible covalent bonds are of special importance in this sense and a number of them are presented in Table 5.1. The list contains a relatively wide range of chemistries that have all been used in DCL generation. For biological applications in aqueous systems, however, a few reaction types are most preferred: transimination, acyl hydrazone exchange, and thiol–disulfide exchange. In principle, the reversible chemistry needs to be compatible with the overall process, not interfering or damaging the target species. Although the preferred chemistries are potentially interfering (e.g., through amino and thiol functionalities of the target species), conditions can generally be chosen to overcome this problem. This can be achieved by control of pH, temperature, solvent composition, auxiliary reagents, and so on. The fragments must, however, be capable of connecting reversibly with one another under the chosen conditions and for the major part potentially able to interact with the target species.

5.3.2
DCL Design

For a DCL to be efficiently produced, the fragments need to fulfill several important characteristics. First and foremost, each component in the library must possess functional groups capable of undergoing reversible exchange to other components. This function can be either symmetric, so that each interacting

Table 5.1 Reversible reaction/interaction types used in DCC system applied to biomacromolecules.

Reaction type	
Transimination	
Hydrazone exchange	
Acyl hydrazone exchange	
Oxime exchange	
Aldol exchange	
Nitroaldol exchange	
Transesterification	
Transthiolesterification	
Transamidation	
Conjugate exchange	
Thiol–disulfide exchange	
Alkene metathesis	
Boronate exchange	
Metal coordination	

partner carries the same functionality (A–A), or orthogonal, where pools of elements carry two different, interacting functionalities (A–B). Both of these strategies possess their advantages and shortcomings, but in general the orthogonal approach maintains better control of the process.

The DCL fragments may furthermore be of interactional and/or organizational character. The former type secures optimal recognition to the target entity, whereas the latter forms the molecular scaffold. The library constituents should be selected to cover the geometrical and functional space of the target binding site and, similar to traditional library design, the choice of fragments may be assisted by careful study of the crystal structure of the target entity. The orientation and assembly of the interactional groups are of fundamental importance for the DCL process to function appropriately. The recognition groups need to be optimally arranged in the binding site resulting in high-affinity constituents. Organizational fragments establishing the core geometry and the topicity of the DCL constituents are therefore of high importance. Of perhaps equal importance is the need for compact dynamic chemistries, where the resulting dynamic functionality occupies minimum space, allowing more freedom in choosing fragments that more closely fill up the binding site space.

5.4 Applications of DCC in Biological Systems

As mentioned, biological systems are very challenging targets for dynamic combinatorial systems. Not only are biological entities often insufficiently stable at non-optimal conditions, sensitive to harsh treatment for longer periods of time, but they are furthermore less available in large, and pure, amounts. Most studies have thus to be made in well-defined buffer systems, greatly limiting the choice of dynamic chemistry used.

For this reason, several reversible chemistries that are efficient in organic phase (e.g., catalyzed by acids or bases) are less convenient in aqueous media. Instead, reactions occurring under sufficiently mild conditions need to be employed. These restrictions have led to the preference of but a few different reaction types, and, to date, mainly imines (including hydrazones, oximes, etc.), disulfides, and metal coordination are employed in these systems. However, other reaction types, such as transthiolesterification, conjugate addition, alkene metathesis, and aldol addition/condensation, have been demonstrated, but are so far less preferred in these systems.

The interest in identifying new and efficient interaction partners to biological entities is obviously very high and the use of an *in situ* molecular evolution method in order to allow the target to select its own ligands is highly attractive. Thus, despite – or perhaps as a result of – these challenges, a multitude of systems have been addressed. Consequently, although DCC has only been developed during a relatively short period of time, its potential in biological systems has been demonstrated in a number of examples (Table 5.2).

Table 5.2 Applications of DCC with biological target species or self-recognition of biological constituents.

Target	Reversible chemistry	Library size	Hit(s)	Reference
Proteins				
enzymes				
trypsin	alcohol–boronate exchange	ND	tripeptidyl-boronate	[19]
carbonic anhydrase	transimination	12	sulfamoylbenzaldimine	[20]
carbonic anhydrase	metathesis	ND	bis-sulfonamide derivatives	[21]
carbonic anhydrase	transamidation	4	sulfonamide-containing dipeptide	[22]
carbonic anhydrase	transamidation	8	dipeptide	[23]
AChE	acyl hydrazone exchange	66	bis-pyridinium structure	[24]
AChE	transthiolesterification	10	substrates	[25]
AChE	transthiolesterification	25	substrates	[26]
HPr kinase/phosphatase	acyl hydrazone exchange	440	2-aminobenzimidazole	[27]
cyclin-dependent kinase 2	hydrazone exchange	ND	oxindole structures	[28]
Aurora kinase	thiol–disulfide exchange	ND	3-amidobenzamide structure	[29]
glutathione *S*-transferase	conjugate addition	5	glutathione derivatives	[30]
glutathione *S*-transferase	conjugate addition	ND	inhibitors	[31]
neuraminidase	transimination	>40000	Tamiflu analogs	[32]
neuraminidase	transimination	ND	Tamiflu analogs	[33]
hen egg white lysozyme	transimination	6	*N*-acetylglucosamine derivative	[34]
hen egg white lysozyme	transimination	12	*N*-acetylglucosamine derivative	[35]
P. cepacia lipase	nitroaldol reaction	16	β-nitroacetates	[36]
galactosyl transferase	transimination	ND	UDP-galactose mimics	[37]
galactosyl transferase	transimination	ND	UDP-galactose mimics	[38]
subtilisin	thiol–disulfide exchange	3	thiocolchicine/ podophyllotoxin conjugates	[39]
β-lactamase	thiol–disulfide exchange	ND	*N*-benzoyl-D-cysteine	[40]
other proteins				
anti-β-endorphin	transamidation	ND	peptides	[41]

Table 5.2 *Continued.*

Target	Reversible chemistry	Library size	Hit(s)	Reference
fibrinogen	transamidation	ND	peptides	[41]
GalNAc-specific lectins	metal coordination	4	tris-GalNAc	[42]
GalNAc-specific lectins	metal coordination	4	tris-GalNAc	[43]
Con A	thiol–disulfide exchange	21	bis-mannoside	[44]
Con A	acyl hydrazone exchange	>474	tris-mannoside	[45]
peanut lectin	thiol–disulfide exchange	>10	divalent galactoside species	[46]
wheat germ agglutinin	thiol–disulfide exchange	13	*N*-acetylglucosamine derivative	[47]
Con A	thiol–disulfide exchange	105	bis-mannosides	[48]
human galectins	thiol–disulfide exchange	21	glycosyldisulfides	[49]
gal-selective plant lectins	thiol–disulfide exchange	21	glycosyldisulfides	[49]
wheat germ agglutinin	aldol reaction	4	sialic acid	[50]
wheat germ agglutinin	aldol reaction	3	sialic acid	[51]
calmodulin	thiol–disulfide exchange	15	inhibitor	[52]
gramidicin A	thiol–disulfide exchange	3	disulfide phospholipids	[53]
Nucleotides				
DNA	transimination/metal coordination	ND	salicylaldimine-Zn(II) complex	[54]
DNA	transimination	36	salicylaldimine-Zn(II) complex	[55]
RNA	metal coordination	>27	salicylamide-Cu(II) complex	[56]
DNA/RNA	transimination	ND	oligonucleotide derivatives	[57]
DNA/RNA	transimination	15	oligonucleotide derivatives	[58]
DNA	thiol–disulfide exchange	6	peptide derivatives	[59]
DNA	thiol–disulfide exchange	54	peptide derivative	[60]

Table 5.2 *Continued.*

Target	Reversible chemistry	Library size	Hit(s)	Reference
RNA	thiol–disulfide exchange	11325	peptide derivatives	[61]
RNA	transimination	>1015	oligonucleotide derivatives	[62]
DNA	thiol–disulfide exchange	7	polyamides	[63]
DNA	thiol–disulfide exchange	5	oxazole-peptide macrocycles	[64]
Cells				
MRSA	thiol–disulfide exchange	3828	psammaplin A analogs	[65]
Ac_2-L-Lys-D-Ala-D-Ala	metathesis/thiol–disulfide exchange	36	vancomycin dimers	[66]
Ac_2-L-Lys-D-Ala-D-Ala	metathesis/thiol–disulfide exchange	36	vancomycin dimers	[67]

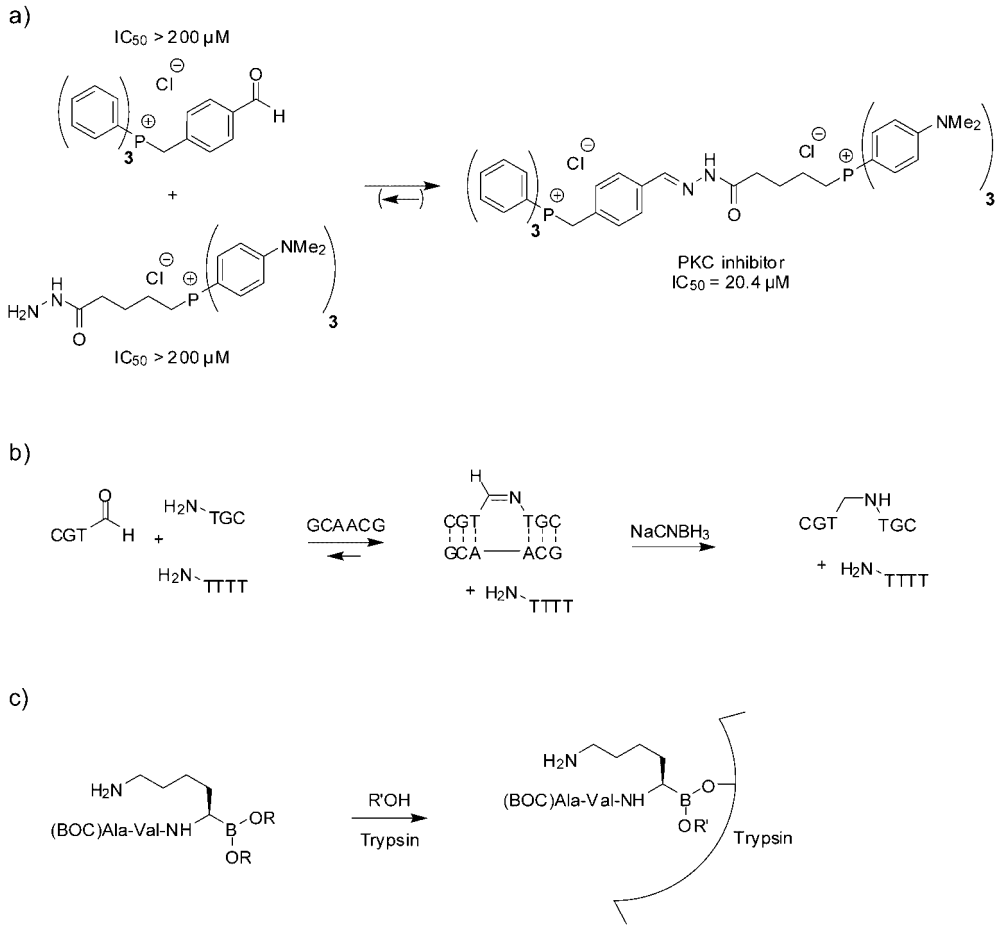
ND, not determined.

As can be seen from Table 5.2, nearly all different classes of biomolecules have been targeted, including enzymes and receptor proteins, oligonucleotides, and whole cells, and libraries have also been constructed using a range of different building blocks. Most of these have been based on non-natural construction elements, but attempts have also been made with amino acids, nucleotides, and carbohydrates.

In addition to biological macromolecules and cells, smaller entities such as cell surface ligands or other discrete substances may be targeted with receptor DCLs. Several of these reports have been discussed in Chapter 3 and will not be covered here.

5.4.1 Early Examples

Predecessors to the DCC concept and other fragment-based ligand discovery methods can be found in a variety of biological systems. Acyl hydrazone chemistry was for example used in a series of highly interesting studies by Rideout *et al.* [68–71], proposing the concept of "self-assembling drugs," where fragments of inhibitors for different enzymes were assembled *in situ* through acyl hydrazone formation (Scheme 5.1a). Reversibility of the acyl hydrazone bond was in this case not an issue, but the reaction type was rather used as a means to chemoselectively



Scheme 5.1 Early examples of potentially dynamic systems targeting biological entities. See text for details PKC, protein kinase C.

generate the active drug from nonactive fragments at the site of action. This strategy bears some resemblance to what has been proposed to be *in situ* synthesis of activated ligand fragments in enzyme-binding sites [72, 73]–the *in situ* "click" approach [74–78], mentioned above. Another early example was presented by Goodwyn and Lynn [79], employing reversible imine formation, and subsequent reduction, in DNA-templated oligonucleotide synthesis (Scheme 5.1b). In these cases, however, no libraries were actually generated and a system that bears more resemblance to the DCC concept was presented by Stroud *et al.* [19] in a strategy called episelection. Tripeptides carrying a terminal boronate group were crystallized in the presence of trypsin and a few simple alcohols–methanol, ethanol, or isopropanol–and the enzyme could preferentially select one of the boronic esters formed (Scheme 5.1c).

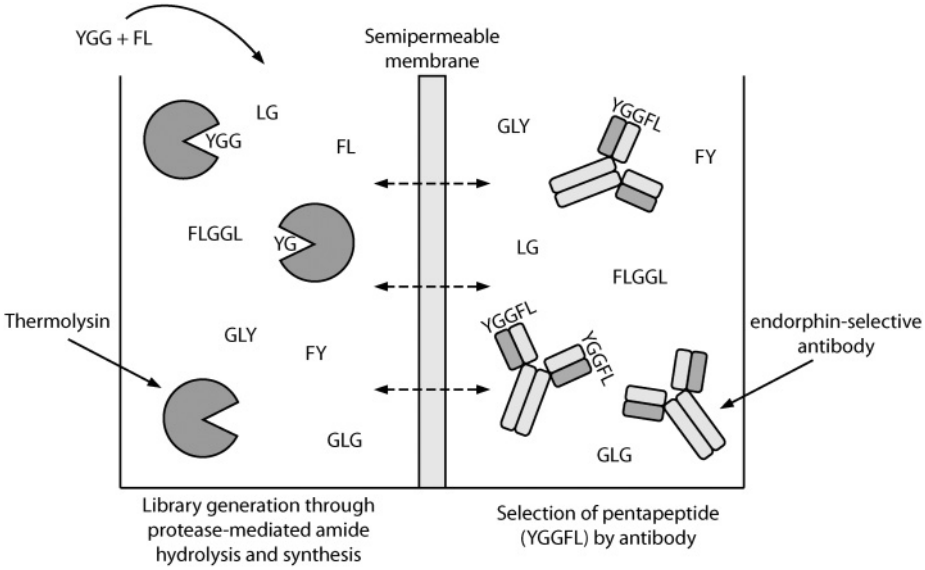


Figure 5.12 Generation and screening of β-endorphin peptides generated by protease-catalyzed transamidation.

A factual demonstration of the DCC concept in a biological system, although without a comprehensive conceptual discussion, employed protease-catalyzed transamidation to generate ligands to β-endorphin (Figure 5.12) [41]. A broadly specific protease (thermolysin) was used together with short peptides (YGG + FL) in a water-rich organic media to establish simultaneous synthesis and hydrolysis, and a peptide library was in this way generated. The process was subsequently performed in presence of a monoclonal antibody specific for the N-terminus of β-endorphin (YGGFL), and binding peptides could be selected and amplification detected. To protect the antibody from digestion, a semipermeable membrane was used to compartmentalize the scrambling process from the recognition process. Another biological entity targeted using the same principle was fibrinogen, also able to select discrete peptides from the transamidation pool [41].

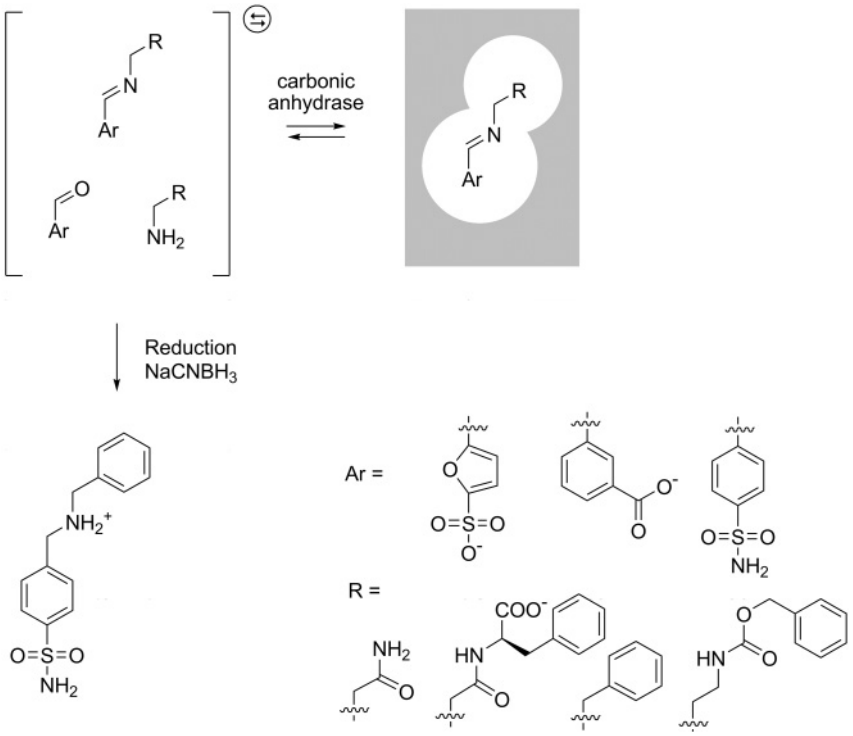
5.4.2 Proteins

5.4.2.1 Enzymes

In comparison to other biomacromolecules, enzymes have been the most attractive targets for DCC protocols. Biocatalysts are preferred for several reasons. (i) They are obviously integral parts of important biochemical pathways that may be of pharmacological relevance. (ii) They are often soluble entities that can be expressed in larger amounts, some of which commercially available. (iii) They offer very attractive features from an analytical point of view and inhibition can be easily

monitored even with small amounts of protein. These are some reasons why several examples have addressed enzymes as targets.

In a sense, the seminal account on the use of enzymes in DCC employed carbonic anhydrase, a Zn(II)-containing enzyme involved in the interconversion between CO_2 and bicarbonate (Scheme 5.2) [20]. As mentioned above, imine formation and exchange (transimination) are especially attractive reactions that are fairly compatible with water, characterized by rapid formation and fast exchange rates. This chemistry was thus adopted, and an imine system based on three different aldehydes and four different primary amines, resulting in a library of 12 different imine constituents, was generated. One of the possible combinations in the DCL was chosen to resemble known inhibitors of this enzyme and the library contained a sulfonamide group, potent in interacting with the Zn(II)-binding site, as well as lipophilic moieties for potential interactions with the neighboring hydrophobic site. An excess of amine was applied in order to overcome the effects from amino groups at the surface of the enzyme, otherwise potentially participating in the imine generation. In addition, cyanoborohydride was added to freeze out the formed imines by reduction to the corresponding secondary amines, thus demonstrating the postmodification strategy mentioned above. The designed imine combination was found to bind preferentially to the enzyme and its formation was



Scheme 5.2 Identification of carbonic anhydrase inhibitor by dynamic reductive amination.

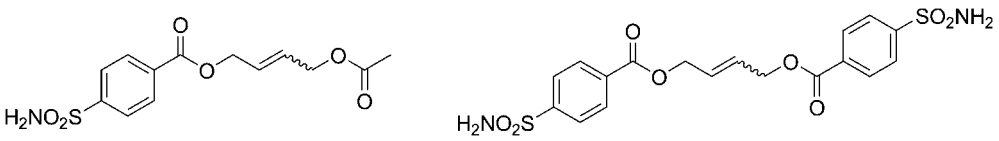


Figure 5.13 Examples of inhibitors identified from a DCL targeting carbonic anhydrase.

furthermore markedly amplified with respect to the concentration in the absence of any protein.

Carbonic anhydrase is a reasonably well-behaved, low-cost protein that is especially suited for model studies and this enzyme has been further targeted with other protocols. For example, bovine carbonic anhydrase II was more recently screened with a DCL prepared using alkene cross-metathesis [21]. In this case, the sulfonamide recognition motif, functionalized as an allyl ester, was combined with a collection of vinyl- or allyl-derivatized components. Cross-metathesis using first- or second-generation Grubbs catalysts then produced a DCL of different alkene structures. The enzyme was in this case allowed to interact with the library after the metathesis, in a pre-equilibrated strategy, and several combinations with promising binding affinities could be identified (Figure 5.13).

This enzyme has also been subjected to the deletion strategy mentioned above [22, 23]. Series of sulfonamide-containing dipeptides were allowed to interact with carbonic anhydrase and the binding monitored. When a protease was added in a second compartment, separated from the recognition by a semipermeable membrane, the less active dipeptides were hydrolyzed faster than the more active dipeptides. In this way, the identification of the best binder could be more easily discerned. In the first example, the total amount of dipeptides was gradually consumed in the process, thus reducing the signal. This was addressed in the second study [23], when a third compartment was introduced in which the dipeptides were continuously synthesized (Figure 5.14).

Acetylcholinesterase (AChE), a serine hydrolase that catalyzes the hydrolysis of the neurotransmitter acetylcholine to acetate and choline at neuromuscular synapses, is another well-studied enzyme. This enzyme has two binding sites, both of which are selective for positively charged functionalities, such as quaternary ammonium groups. One active site is located at the bottom of a deep gorge and a so-called peripheral site is situated near the rim of this gorge. By bridging the two sites, very efficient inhibitors can be found and the AChE from the electric ray *Torpedo marmorata* was first targeted using an acyl hydrazone system in a pre-equilibrated system (Figure 5.15) [24]. A DCL composed of interconverting acyl hydrazones was generated starting from 13 different hydrazide and aldehyde building blocks, some of which contained quaternary ammonium groups, forming a library of 66 different species. Of all possible acyl hydrazones formed, active compounds containing two terminal cationic recognition groups separated by a spacer of appropriate length could be rapidly identified using a dynamic deconvolution procedure based on the sequential removal of starting building blocks. A

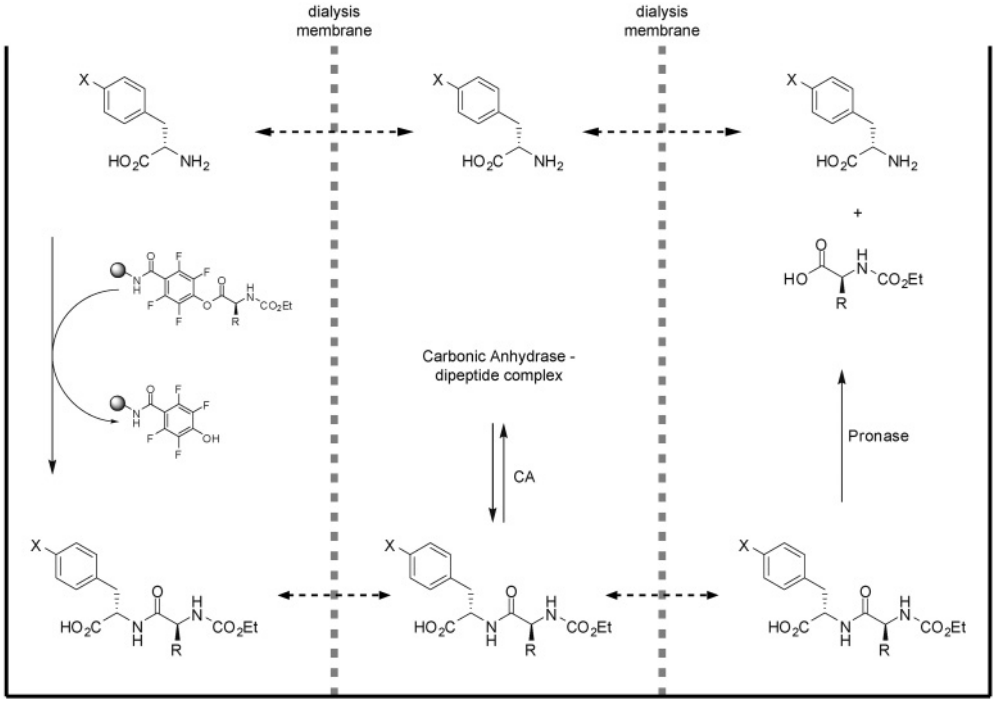
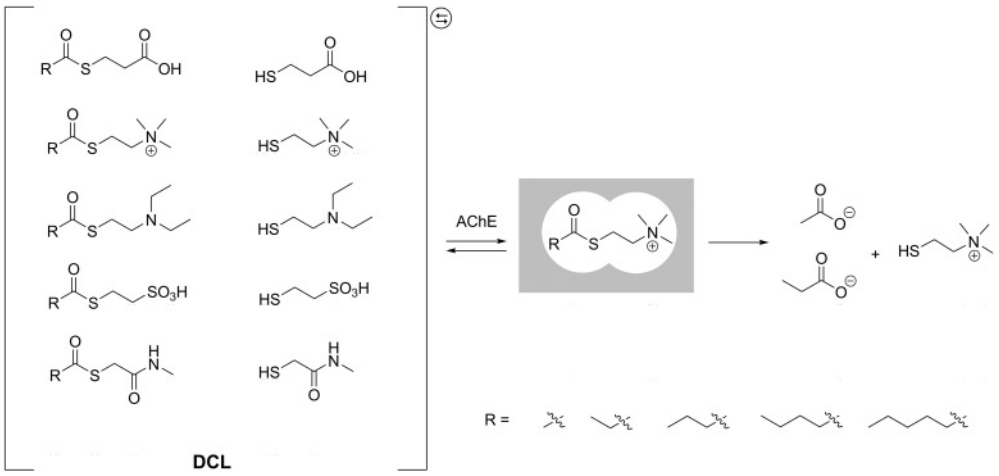


Figure 5.14 Deletion method to carbonic anhydrase inhibitors.

very potent bis-pyridinium inhibitor with a K_i value in the nanomolar range was selected from the process (K_i = 1.09 nM, αK_i = 2.80 nM) and the contribution of various structural features could be evaluated.

The DCR concept of applying kinetic resolution to DCLs was also applied to AChE [25, 26]. In this case, transthiolesterification was first used in DCL generation and libraries could be efficiently formed at neutral pH in aqueous buffer systems. Treatment of the thiolester DCL with the enzyme AChE resulted in the best substrates being first recognized by the enzyme and subsequently hydrolyzed (Scheme 5.3). This resulted in loss of the acyl component from the DCL and forced the DCL to reconstitute to generate more of the hydrolyzed species. Over time, two of the thiolesters, the acetyl and propionyl thiocholine, respectively, were selected by the enzyme, with the acetyl species being more rapidly hydrolyzed than the propionyl counterpart. After prolonged reaction times, less adapted esters were also affected, whereas all other thiolesters remained untouched in accordance with the known specificity of the enzyme. The screening of the DCL proved in this case straightforward, enabling both the identity of the enzyme substrates and the relative order of efficiency to be established. In addition, 100% selection/amplification was achieved for the selected thiolesters. The DCR process was also tested for a range of other hydrolases under the same set of conditions, including butyrylcho-



Scheme 5.3 AChE targeted with a dynamic thiolester DCR protocol.

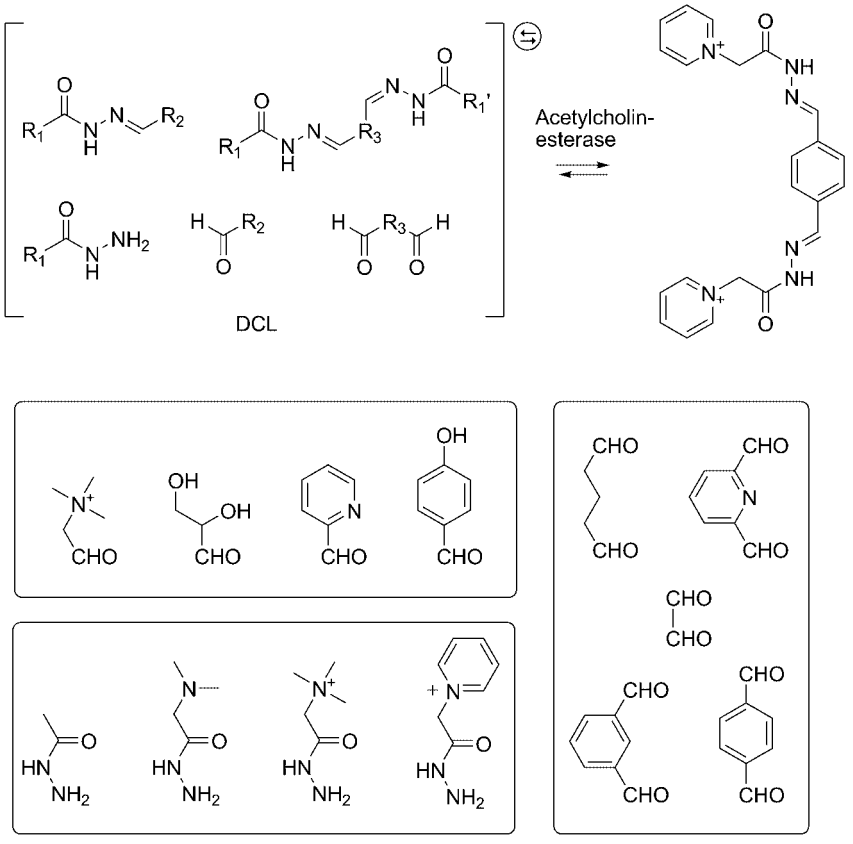
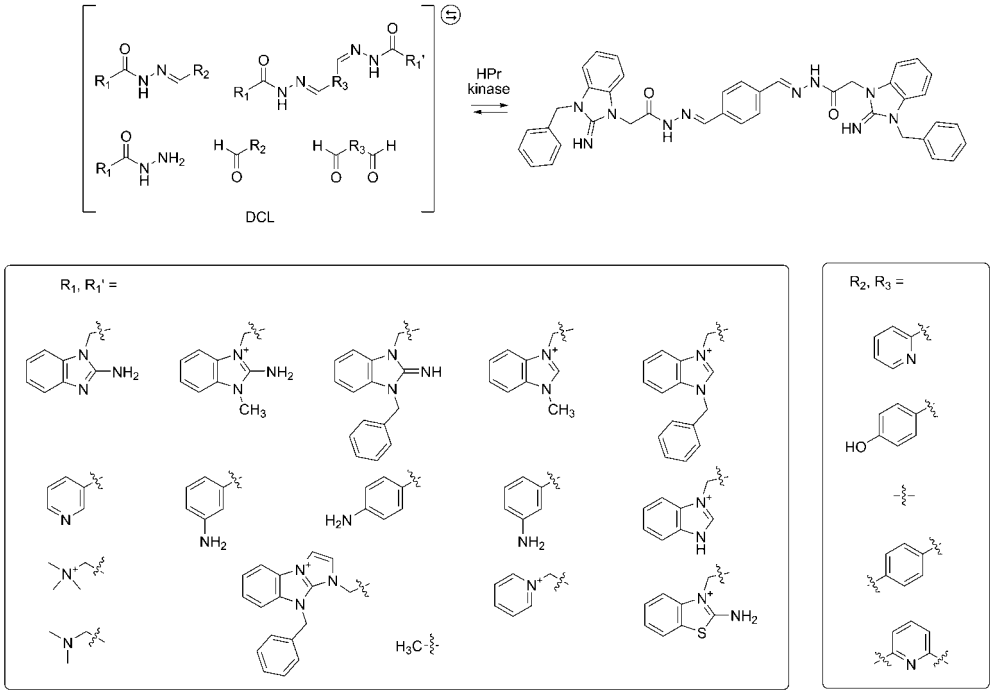


Figure 5.15 Dynamic acyl hydrazone libraries targeting AChE. A highly potent bis-pyridinium compound could be identified.

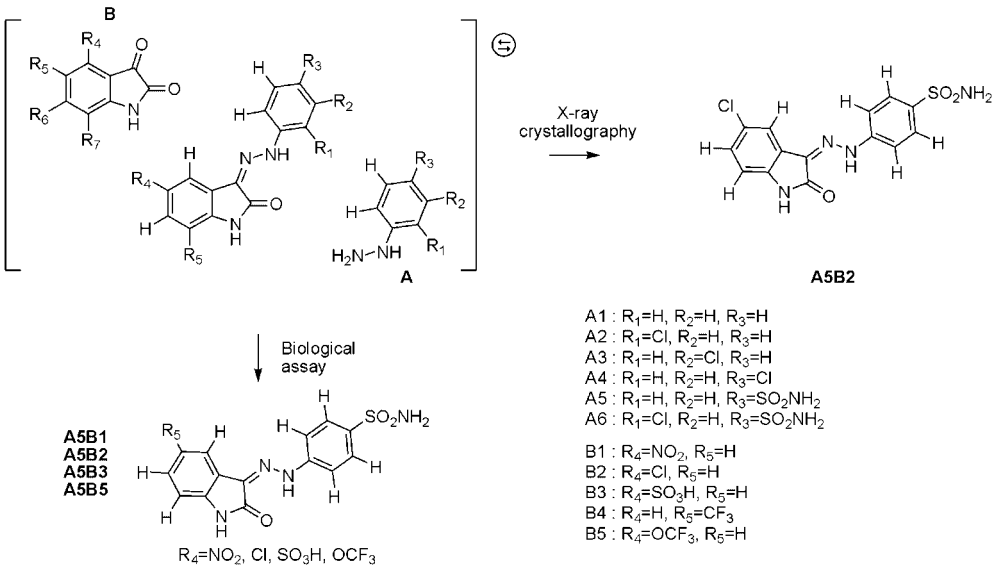
linesterase, horse liver esterase, *Candida cylindracea* lipase, β-galactosidase, trypsin, and subtilisin, where the substrate selectivities could be efficiently probed.

Kinases represent one of the most targeted classes of enzymes in drug discovery, often being key enzymes in important biochemical pathways. They are also especially challenging, owing to the fact that the phosphoryl donor site is very similar between the enzymes. Nevertheless, kinases have also been targeted with DCC protocols. In the first example, acyl hydrazone exchange was chosen to generate ditopic heterocycle DCLs against HPr kinase/phosphatase, a serine kinase/phosphatase active in bacterial carbohydrate metabolism [27]. The libraries were in this case composed of all combinations resulting from the dynamic interconversion of 21 hydrazide and aldehyde building blocks, resulting in libraries containing up to 440 different constituents together with the starting components (Scheme 5.4). Of the acyl hydrazones formed in this case, active lead compounds containing two terminal cationic heterocyclic recognition groups separated by a spacer of appropriate structure could be rapidly identified using a dynamic deconvolution procedure.



Scheme 5.4 Dynamic acyl hydrazone libraries targeting HPr kinase. A potent bis-cationic compound could be identified.

Another kinase, cyclin-dependent kinase 2, was targeted using the DCX approach (Scheme 5.5) [28]. The DCL was in this case based on phenyl hydrazone interconversion between a series of phenylhydrazines and a corresponding series of isatins.

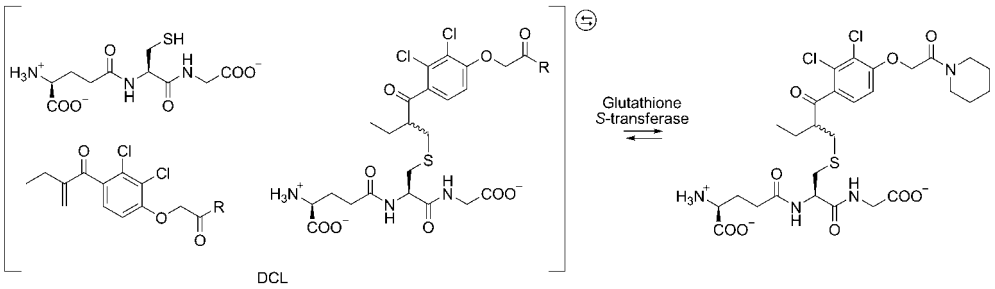


Scheme 5.5 DCX applied to cyclin-dependent kinase 2 (CDK2).

The library could in this study be generated in a dimethyl sulfoxide-rich solvent mixture, a prerequisite for not damaging the protein crystals, and the crystals were directly soaked with the DCL solution. By following the electron-density maps in the crystallography setup, the best binders could be identified and several oxindole structures with inhibitory activities in the nanomolar range were found.

A recent method relying on the "tethering" approach mentioned above targeted Aurora kinase, an enzyme that has shown significant roles in regulating mitosis [29]. An extension to the original approach was described, where the irreversible action of an extender unit is replaced with a disulfide, enabling reversible cysteine modification. A second disulfide moiety of the extender can then reversibly interact with other DCL components and combinations showing affinity to the target be detected. The DCLs were generated from a diaminopyridine-based bis-disulfide extender together with sets of 10 different disulfides and the cysteine-modified Aurora kinase. Very few hits were identified overall, owing to the simultaneous requirement for two disulfide bonds, of which a 3-amido-benzamide structure could be identified and further studied (Figure 5.16).

Glutathione *S*-transferase, a dimeric enzyme catalyzing the conjugation of glutathione to a variety of xenobiotic substances such as prostaglandins and quinones, was targeted with a stoichiometric strategy using a DCL based on conjugate addition of thiols to enones [30, 31]. The thiols were analogs of glutathione and ethacrynic acid was used as the basic structure/scaffold for the enones. DCLs were generated and screened towards the transferase, and relatively potent inhibitors could be indentified (Scheme 5.6).



Scheme 5.6 DCL from glutathione and ethacrynic acid derivatives. Species were selected from targeting the DCLs with glutathione *S*-transferase.

The most extensive testing of DCC in drug discovery was performed against the enzyme neuraminidase, one of the key enzymes in influenza virus activity (Scheme 5.7) [32, 33]. Dynamic reductive amination and the stoichiometric strategy were in this case used, generating very large imine libraries (theoretically more than 40 000 constituents) initially comprising the scaffold molecule and 20 aldehydes. Both aldehydes and ketones could, however, be used as components in the DCL. The DCL components were in part inspired by previous structure–activity relationship studies of the enzyme and in part based on a known drug against the enzyme (Tamiflu®). Highly potent neuraminidase inhibitors, showing resemblance to the Tamiflu structure, could be identified from the campaign. The best inhibitor was found to be a monosubstituted amine from an aromatic aldehyde, with a K_i value of 1.6 μM. In a second-generation library, the scaffold structure was modified to compensate for the prevalence of monosubstituted hits, by which the amplified

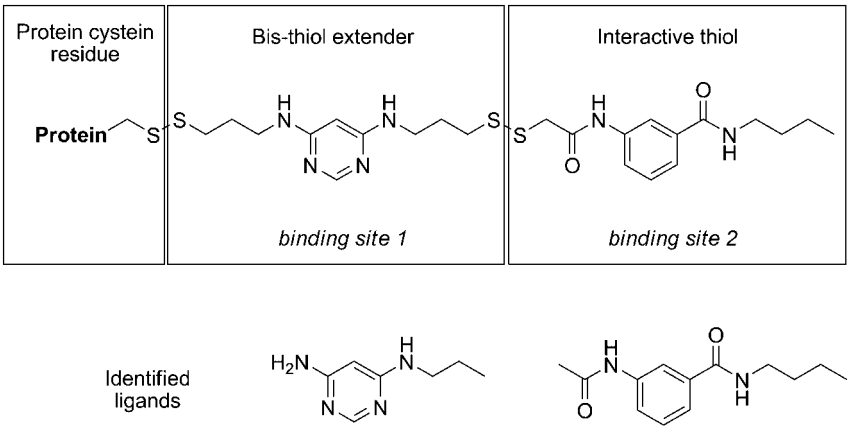
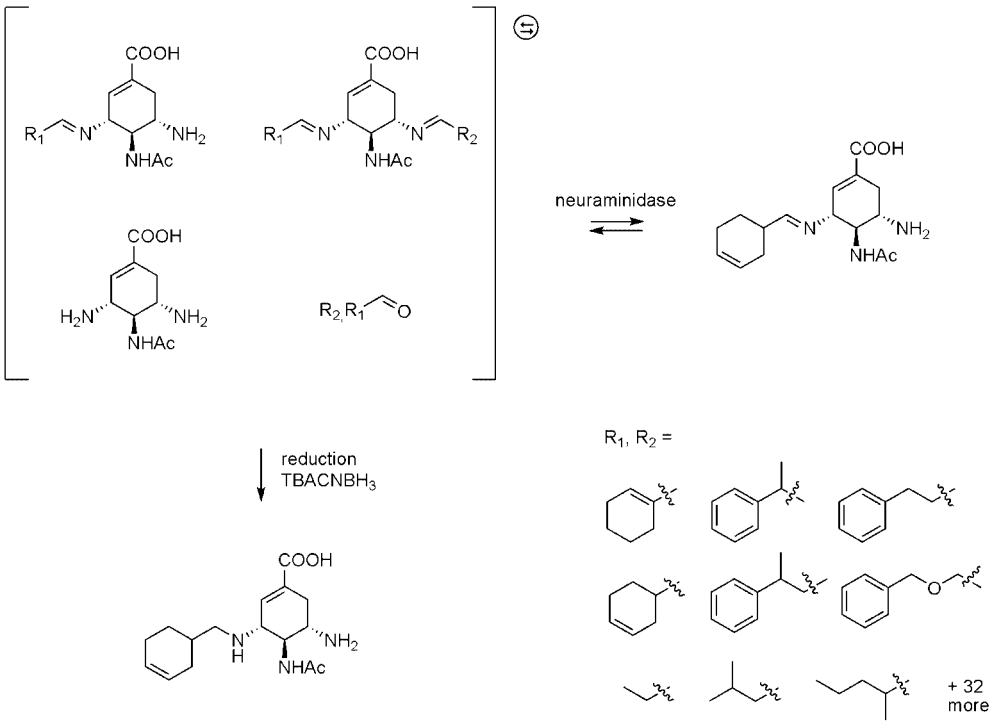


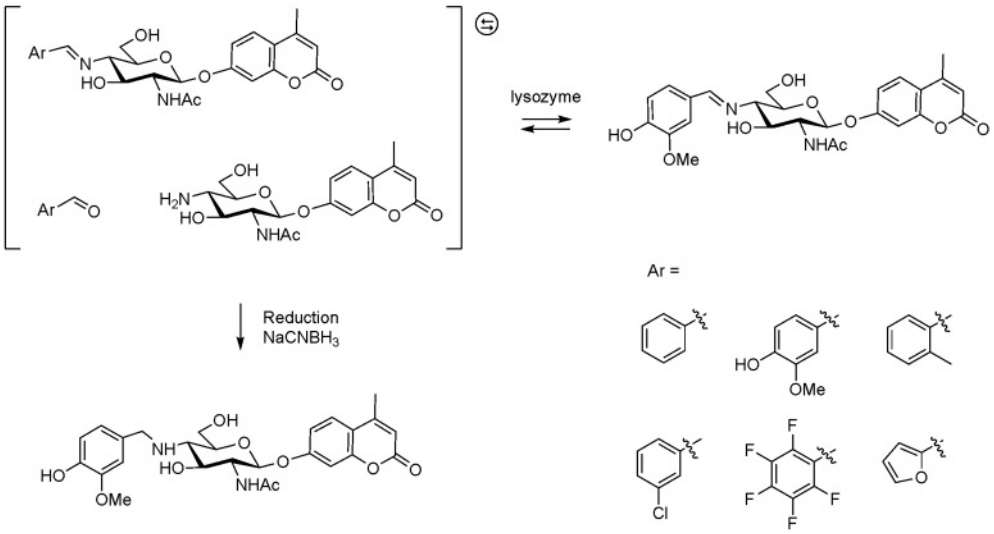
Figure 5.16 Identification of Aurora kinase inhibitors applying a bis-thiol extender to the tethering approach.



Scheme 5.7 Imine-based DCL targeting neuraminidase.

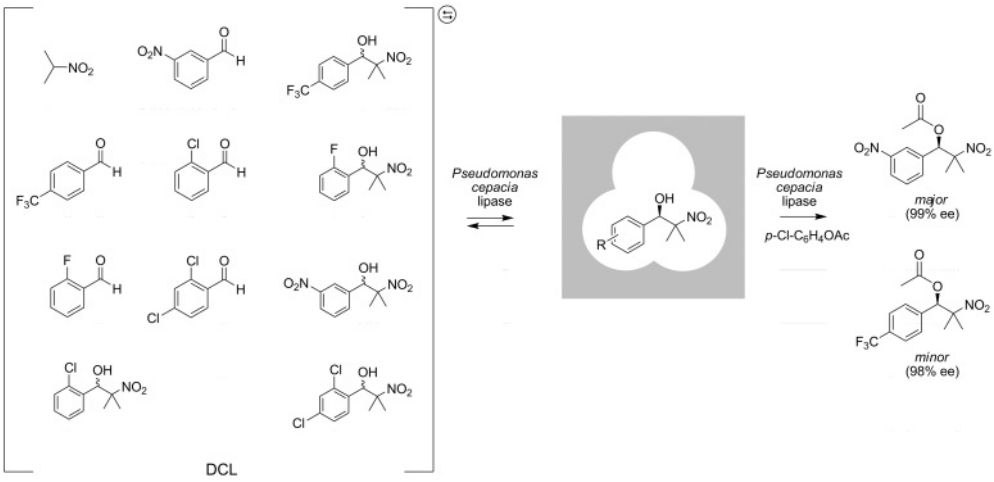
components had K_i values as low as 4.0 nM. Although the Tamiflu scaffold was used as a structural base, this study demonstrates the potential of DCC in true drug discovery. The virtual character of the library (i.e., the absence of visible constituents from the library) was in this case also very conspicuous and the best hits could only be detected upon reduction in the presence of enzyme. Very large amplification effects (above 100) were thus recorded.

Lysozyme, a group II glycosidase involved in peptidoglycan degradation in cleaving *N*-acetyl-glucosamine oligomers (chito-oligosaccharides) into their smaller units, was targeted using the postmodification strategy with transimination. 4-Methylumbelliferyl-labeled β-D-Glc or β-D-GlcNAc scaffolds carrying primary amino groups in the 4-positions, were employed to form DCLs with a series of aromatic aldehydes (Scheme 5.8) [35]. One of the reduced imine combinations showed inhibitory properties (K_i = 0.6 mM), on a par with the known inhibitor chitotriose. In an interesting follow-up study [34], the same system was used to monitor the formation of the imines rather than the amines formed after reduction. This could be achieved by freezing the equilibration by a rapid drop in pH and instantaneous high-performance liquid chromatography analysis. It could thus be shown that the parent imine was selected by the enzyme with a K_i value of approximately 0.16 mM.



Scheme 5.8 Imine-based DCL targeting lysozyme.

Lipases are a family of enzymes that, in addition to their hydrolytic activity on triglycerides, also catalyze (trans)esterification reactions. They recognize a broad range of unnatural substrates in either aqueous or nonaqueous phase, have a high commercial availability, do not require expensive cofactors, and are easily recoverable. These factors make lipases especially interesting and they have been used extensively in, for example, asymmetric synthesis. The lipase from *Pseudomonas cepacia* was also targeted in a dynamic combinatorial resolution-type protocol [36]. Based on the efficient nitroaldol (Henry) reaction, DCLs of aldehydes, nitroalkanes, and β-nitroalcohols could be easily generated (Scheme 5.9).



Scheme 5.9 Dynamic combinatorial resolution of nitroaldol DCL (ee, enantiomeric excess).

In the DCR protocol using lipase as selector, selective β-nitroalcohol acylation was subsequently performed, yielding the corresponding acetylated products. Two products from the DCL were resolved by the process, where the major product proved to be the ester produced from 3-nitrobenzaldehyde and 2-nitropropane. The nitroaldol-lipase DCR process could, however, not only amplify specific β-nitroalcohol derivatives, but also lead to their asymmetric discrimination. The enantioselectivity of the process proved very high, resulting in high enantiomeric purity of the products. The (*R*)-enantiomer of the ester was thus resolved to 99% enantiomeric excess.

Galactosyltransferases were targeted in two studies [37, 38]. The glycosyltransferase family of enzymes are active in the glycosylation of a variety of acceptors, using activated phosphate donors, and are important targets for glycobiological studies. Similar to the kinases, the enzymes are challenging targets for drug discovery, since their three-dimensional structures are often unknown and the nucleotide part of the donor is a common motif for several members of the family. In the reported studies, an imine-based system was used in combination with reductive amination. Nucleoside and carbohydrate aldehydes were combined with diamines of varying length, generating an imine DCL that was subjected to an α1–3-galactosyltransferase. Combinations from the uridine derivative with aromatic diamines were selected (Figure 5.17), which after reduction resulted in half-maximal inhibitory concentration (IC_{50}) values in the millimolar range. Substitution for the diamines by monoamines resulted in improved selection, as demonstrated for another DCL. Interestingly, the enzyme could be used in non-stoichiometric amounts, otherwise necessary in this protocol, probably due to the fact that the reduced imines were significantly less bound by the enzyme than the parent imines. The DCL could also be used to select different binders when β1–4-galactosyltransferase was targeted, where instead aliphatic amines were selected.

In a study of slightly more preliminary character, subtilisin and albumin were targeted with a disulfide-based DCL containing derivatives of thiocolchicine and podophyllotoxin [39]. Rather than targeting the binding site of the proteins, the aims of the study were to generate the disulfide library and to monitor the DCL dependence in presence of a protein. The choice of proteins were in this case based on the stability in water-poor organic solvents. From a three-membered DCL, it was also found that both subtilisin and albumin quenched the formation of the heterodimer, and that both the homodimers were favored (Figure 5.18).

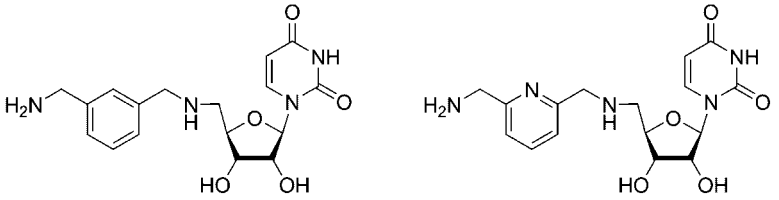


Figure 5.17 Combinations of a uridine derivative with aromatic diamines selected from an imine DCL targeting α1–3-galactosyltransferase.

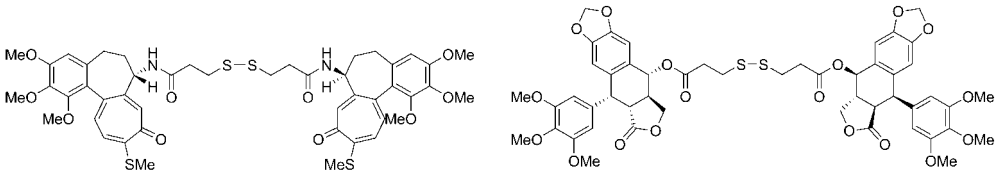


Figure 5.18 Homodimers favored by subtilisin and albumin.

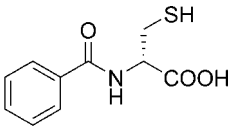


Figure 5.19 Thiol (*N*-benzoyl-D-cysteine) selected from a DCL targeting BcII metallo-β-lactamase.

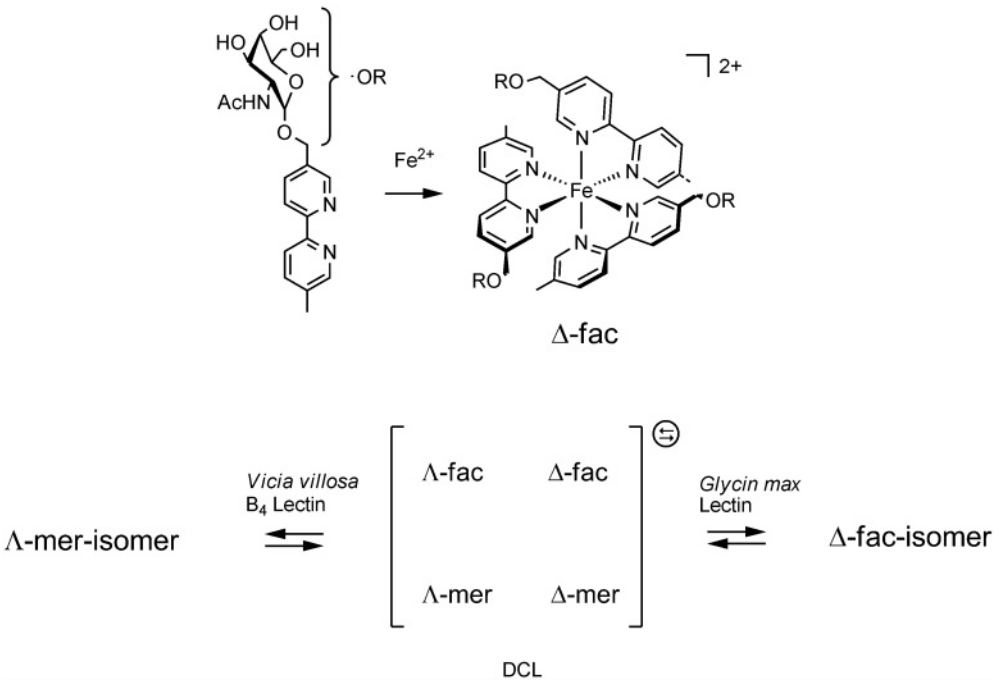
In a recent account, thiol–disulfide interchange chemistry was applied to ligand discovery with the BcII metallo-β-lactamase from *Bacillus cereus*, an enzyme that is important in β-lactam antibiotic resistance [40]. A dynamic combinatorial mass spectrometry (MS) approach was proposed, employing a generic component that was strongly complexed to the active site zinc ions and also able to form disulfides with other thiols. Electrospray ionization-MS was then used to directly analyze the enzyme-bound disulfides under nondenaturing conditions. This approach enabled the identification of an *N*-benzoyl-D-cysteine-derived inhibitor, showing a K_i value below 1 μM (Figure 5.19).

5.4.2.2 **Other Proteins**

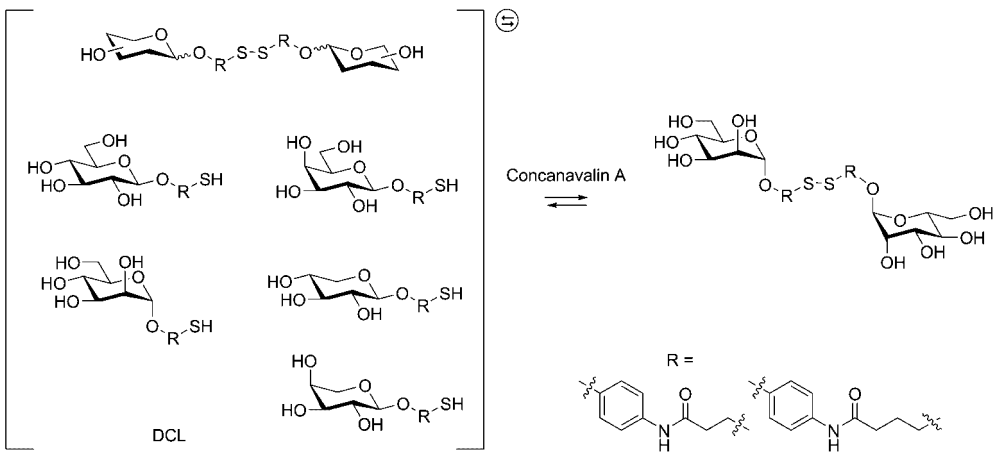
Proteins other than enzymes have been targeted with DCC systems on a number of occasions. As is the case for enzymes, mainly easily accessible (commercially available) proteins have been used, whereas in some instances proteins of more precious character have been challenged with the technique.

In addition to studies with carbohydrate-processing enzymes such as transferases and hydrolases, carbohydrate-binding proteins have also been targeted. In an early example, a prototype DCL was constructed using metal coordination and allowed to interact with different plant lectins (Scheme 5.10) [42, 43]. *N*-acetyl-D-galactopyranoside (GalNAc) groups were attached to a bipyridine unit and four different trivalent tris-bipyridine complexes were formed after addition of Fe(II) ions. The formed DCL was subsequently screened against lectins specific for GalNAc, including the *Vicia villosa* B4 (hairy vetch) and the lectin from *Glycine max* (soybean); one of the complexes was in each case selected.

In another example, disulfide interchange was used to generate a bivalent carbohydrate library (Scheme 5.11) [44]. Starting from a range of different thiol-derivatized carbohydrates, ditopic DCLs of up to 21 different disulfide species were easily generated. When screened against the plant lectin from common Jack bean,



Scheme 5.10 Dynamic generation around a metal-coordination center targeting GalNAc-specific plant lectins.



Scheme 5.11 DCL of disulfide-containing carbohydrate structures. A bis-mannoside was selected in the presence of the Jack bean lectin Con A.

Concanavalin A (Con A), immobilized on solid-support, a bis-mannoside structure was selected. Generation of the library was performed at neutral to slightly basic pH, where the redox exchange is rapid, and the exchange could efficiently be stopped by lowering the pH.

Acyl hydrazone exchange has also been probed as reversible chemistry for the generation of oligotopic and directional carbohydrate DCLs of larger size [45]. Here, carbohydrate-derivatized benzaldehyde functionalities were allowed to interact with hydrazides of different topology and structure, generating libraries composed of more than 400 different acyl hydrazones. When screened against the plant lectin Con A, used as a model carbohydrate-binding protein, a trivalent trismannoside structure could be identified (Figure 5.20). Detailed binding studies estimated a relatively strong binding affinity with an IC_{50} of 22 μM from a solid-phase assay, comparable to the natural trimannoside ligand for the protein.

Several research groups have independently probed disulfide DCLs of similar structure in part based on 1-thiosaccharides and other thiols. For example, prototype disulfide-linked glycopeptide libraries were constructed from 1-thiosaccharides and cysteine-rich oligopeptide building blocks, affording DCLs composed of various carbohydrate–peptide conjugates and cyclic peptides [46]. Preliminary surface plasmon resonance analyses suggested that a 1-thiogalactose-derived library contained divalent galactoside species capable of interacting with the peanut lectin (Figure 5.21).

Another account describes a similar system, where 1-thiosaccharides and cysteine-containing oligopeptides were allowed to form prototype disulfide glycopeptide DCLs [47]. In this case, wheat germ agglutinin was instead targeted and

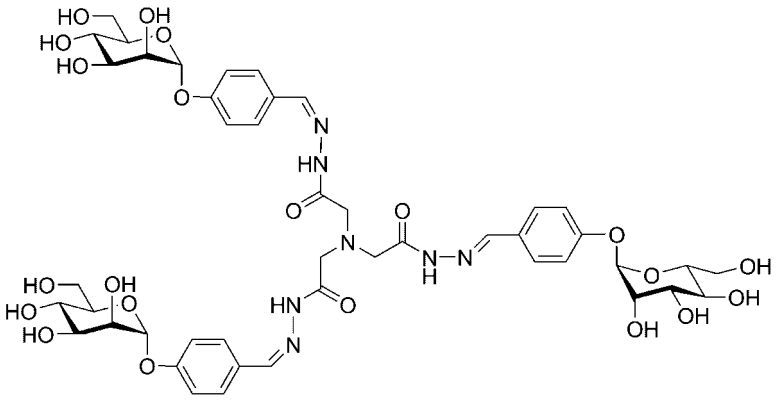


Figure 5.20 Tris-mannoside structure identified from acyl hydrazone DCL targeting Con A.

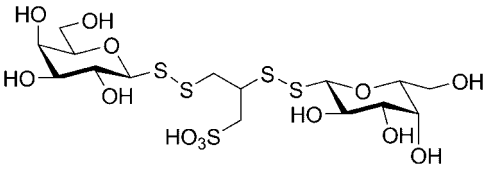


Figure 5.21 Selected divalent galactoside species capable of interacting with the peanut lectin.

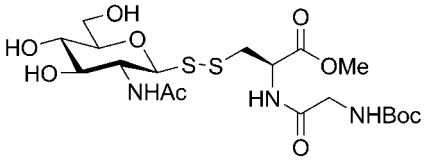


Figure 5.22 *N*-acetylglucosamine-peptide disulfide selected from a disulfide DCL targeting wheat germ agglutinin.

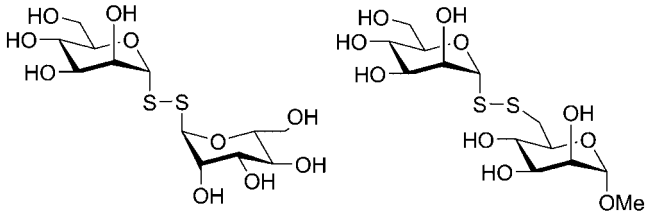


Figure 5.23 Glycosyldisulfides identified from a DCL targeting Con A.

MS analyses revealed the amplification of an *N*-acetylglucosamine-peptide disulfide (Figure 5.22), as expected for this protein.

These studies made use of 1-thiosaccharides, and in principle DCLs built from carbohydrate structures where hydroxyl functionalities are replaced by thiol group, are structurally compact and can result in disulfide adducts more closely resembling natural glycosides. This was also probed in carbohydrate–lectin interactions, using the plant lectin Con A as target species [48]. The DCLs were generated from a pool of thiosaccharides and a range of additional noncarbohydrate thiol components through reversible thiol–disulfide interchange, and screened using an efficient method based on a quartz crystal microbalance setup. It was found that dimers based on 1-thio- and 6-thio-mannose analogs were the most active inhibitors (Figure 5.23), the latter of which in part analogous to the native ligand for this lectin.

In a second study with a similar DCL [49], more physiologically relevant systems were targeted. Lectins with specificity towards galactose, fucose, or *N*-acetylgalactosamine, respectively, were tested in solid-phase assays and probed as inhibitors in lectin–tumor cell interactions. A human endogenous lectin (galectin-3), involved in tumor spread and cardiac dysfunction, was for example probed by the DCLs. *N*-acetylgalactosamine was pinpointed as the most important building block of libraries for the human lectin and the digalactoside as the most potent compound acting on the toxic viscumin. This was shown by a series of inhibitory experiments between viscumin and human B lymphoblastoid line Croco II (Figure 5.24).

Highly intriguing accounts on enzymatic DCL generation applied to a carbohydrate-binding protein have also been reported [50, 51]. Using the *N*-acetylneuraminic acid aldolase (sialic acid aldolase), reversible aldol addition of three or four different carbohydrates to pyruvate was catalyzed, yielding small DCLs of

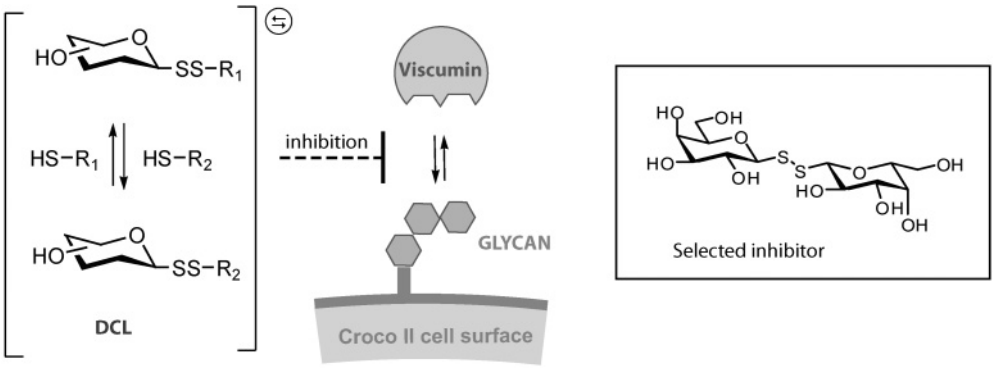
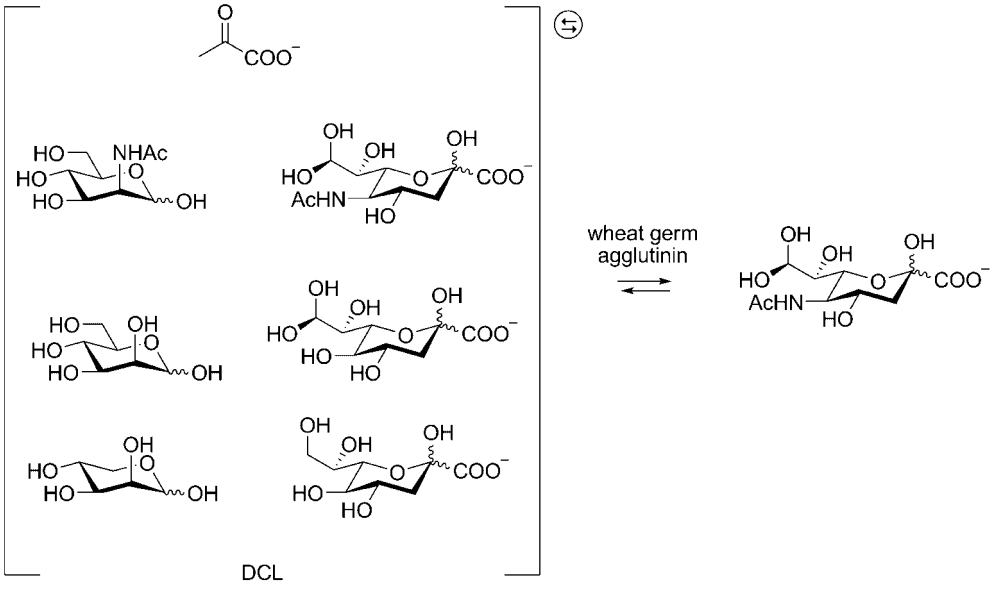


Figure 5.24 Disulfide DCL targeting the interaction between viscumin and the human B-lymphoblastoid line Croco II.

three or four new adducts of which one was sialic acid. When this library was probed against wheat germ agglutinin, the formation of sialic acid increased in comparison to the two other adducts (Scheme 5.12).



Scheme 5.12 DCL generated by sialic acid aldolase targeting wheat germ agglutinin.

Calmodulin is a calcium-binding protein that regulates several biologically important processes by interacting with various proteins. This protein was targeted with a disulfide-based DCL in a pre-equilibrated strategy [52]. A DCL generated from five cysteine-based derivatives was allowed to interact with the protein in the

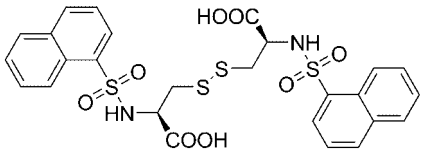


Figure 5.25 Bis-sulfonamide structure selected from a disulfide DCL targeting calmodulin.

presence and absence of calcium ions. Two of the combinations were identified in the process, of which a sulfonamide structure showed the highest affinity (Figure 5.25).

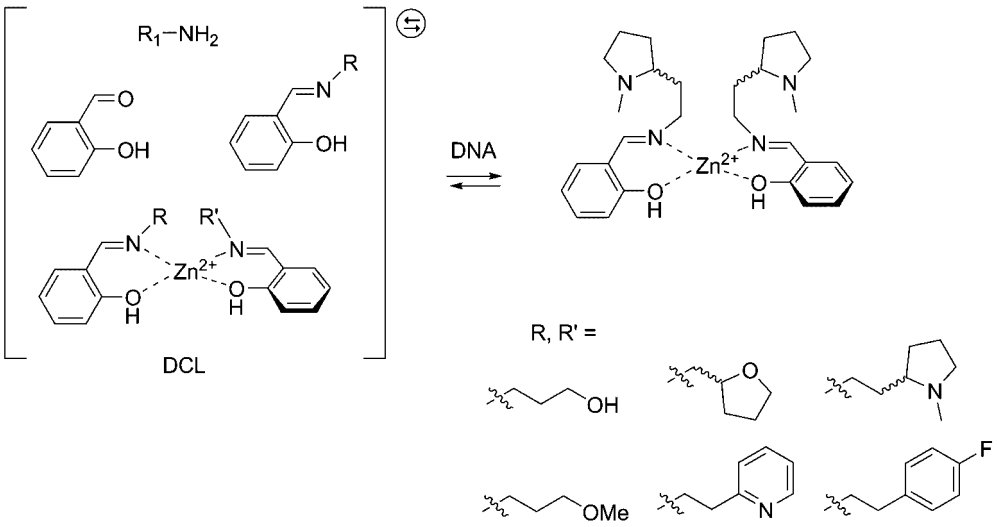
In an interesting study, gramicidin A, a bacterial channel-forming antibiotic, was used to influence the equilibrium distribution of a DCL-like system [53]. A mixture of two disulfide-based phospholipids was used to form vesicles. When the disulfide-based dimers were allowed to undergo monomer exchange in the presence of channel-forming gramicidin A, the phospholipids were left in a randomly arranged state, whereas a preference for the homodimers occurred for the non-channel-forming species. The study points to the ability of gramicidin A to communicate its conformation to the surrounding phospholipids.

5.4.3 Nucleotides

In addition to enzymes and other proteins, DNA and RNA have been successfully targeted with DCC systems in several studies. Oligonucleotide duplex formation is a highly efficient means to generate stable complexes even at very low concentration and these complexes can be easily studied by a range of different analytical methods.

However, selective and high-affinity recognition of RNA and DNA by small molecules is not a trivial task, and dynamic combinatorial approaches can in principle accelerate the discovery process. The first true examples in this area involved metal coordination as a useful means to generate diversity under controlled conditions (Scheme 5.13) [54, 55]. As mentioned, metal coordination interactions are useful for DCL formation and in some cases sufficiently prone to scrambling in aqueous media. In these examples, Zn(II) was used in combination with libraries of salicylaldimines and probed against binding to double-stranded oligo(A–T) nucleotides. Dynamic combinations of the different Zn(II) complexes could be generated and one of the species showed higher binding than all other library constituents. A binding constant in the lower micromolar range could also be recorded.

Similar DCLs were subsequently generated and screened against RNA [56]. In this case, Cu(II) was used to coordinate salicylamides, forming at least 27 square planar mono- and bis-salicylamide complexes. The DCL was probed against an RNA hairpin structure derived from the GTP-binding P7 helix from the *Pneumocystis carinii* group I intron. Equilibrium dialysis was adopted as a technique to facilitate library screening and a Cu(II)-histidine salicylamide complex could be

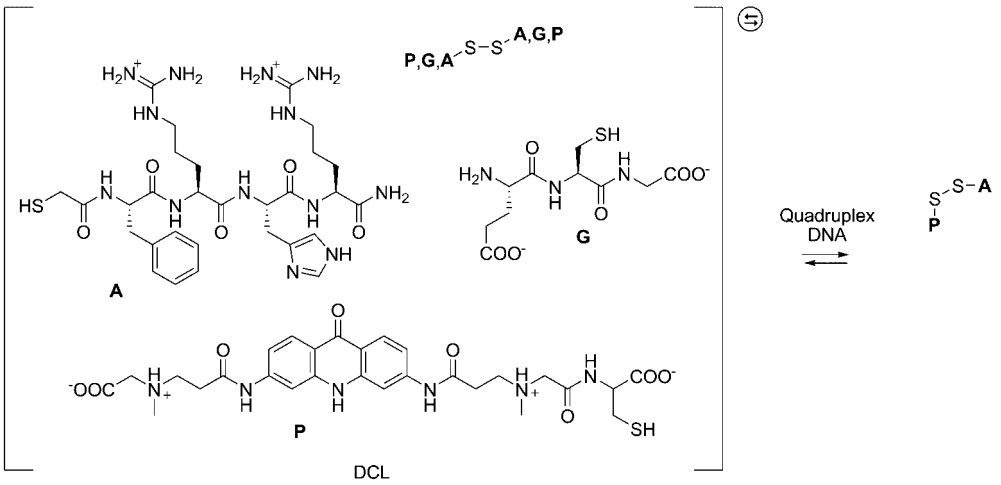


Scheme 5.13 DCL of Zn(II) complexes interacting with duplex DNA.

identified. Detailed analysis of the interaction revealed a binding constant of 50 nM.

In two interesting contributions, a resin-bound DCC method was adopted where the library components were covalently attached to a solid support and allowed to equilibrate in the presence of a labeled target. Resin beads bearing the labeled target were easily isolated and the active components could be identified [60, 61]. In the first application, selection of DNA-binding constituents was made using a thiol–disulfide approach based on the octadepsipeptide family of bis-intercalating DNA binding agents reported to bind DNA in the minor groove through bis-intercalation. Two model DNA sequences were screened, and one of the homodimers were identified to show affinity in the micromolar range to one of the DNA structures. Later, the synthesis and evaluation of a very large DCL (more than 11 000 constituents) targeting the human immunodeficiency virus (HIV)-1 frameshift regulatory stem–loop RNA critical to HIV proliferation was described using the same method [61]. Again, DNA-binding analogs of the bis-heterocyclic octadepsipeptide family of nucleotide-binding natural products were used, able to form disulfide linkages. From this library, a constituent was identified that binds the target RNA showing high selectivity and affinity in the micromolar range (Figure 5.26).

In a series of studies, quadruplex DNA was targeted [59, 63, 64]. In the first of these, a prototype disulfide DCL consisting of six possible disulfide combinations was allowed to interact with a human telomeric quadruplex DNA. Two constituents were selected by the process, of which one was an expected combination of an acridone unit, designed to interact with the terminal G tetrad of a parallel quadruplex, and a tetrapeptide sequence, known to have quadruplex recognition properties. The other hit was more surprisingly the tetrapeptide dimer (Scheme 5.14).



Scheme 5.14 Disulfide identified to interact with a human telomeric quadruplex DNA.

In a second study, selection of distamycin-type structures was performed using a similar protocol. In this case, three different thiol-derivatized oligoamides, containing one, two, or three *N*-methylpyrrole groups, respectively, were prepared and allowed to form a 10-member disulfide DCL together with glutathione. Two DNA structures were targeted: a four-stranded intramolecular G quadruplex formed by human telomeric DNA and an A/T-rich duplex DNA sequence identified in the promoter region of the oncogene c-*kit*. Three of the combinations of the longer ligands were identified from the DCC process, moderately amplified in presence of the quadruplex, but more pronounced in presence of the duplex. In the most

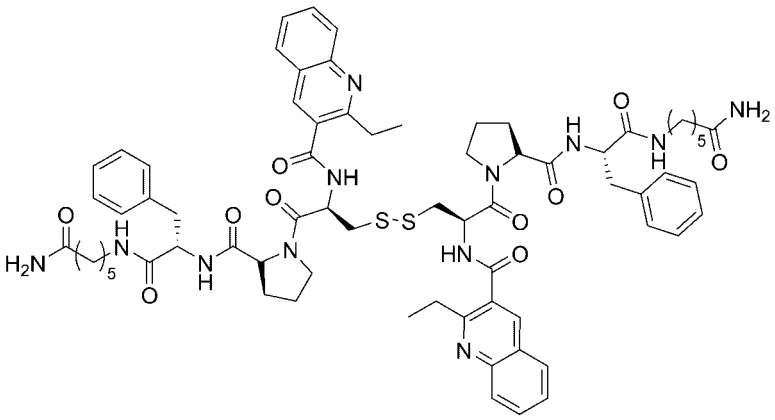


Figure 5.26 Selected disulfide from DCL targeting the HIV-1 frameshift regulatory stem–loop RNA.

recent study, two different G quadruplex structures, derived from G-rich sequences of the promoters of c-*kit* and c-*myc* proto-oncogenes, respectively, were targeted with disulfide-type DCLs. A thiol-derivatized oxazole-peptide macrocycle, known to interact with quadruplexes, was used as a generic interaction partner and a series of other thiols of various structures, with different potential for hydrogen-bonding and charge interactions, was applied. Several of the disulfides could be identified as ligands for the quadruplexes, of which combinations of the generic macrocycle with guanidine-containing thiols proved most efficient.

Following the early study discussed above [79], oligonucleotides have been derivatized as amines and the resulting diamino-duplexes allowed to form imines with a series of aldehydes [57, 58]. In one example, a loop–loop complex, also known as a "kissing" complex, displaying fundamental base-pairing complementarity and a constrained three-dimensional structure, was targeted. The DCL consisted of a 14-nucleotide aptamer in which a ribonucleotide in the loop was replaced by a 2′-amino-derivatized nucleotide together with three non-nucleotidic aldehydes. This was allowed to interact with a stoichiometric amount of a 27-nucleotide form of the transactivation-response element (TAR) RNA hairpin loop of HIV-1. Reductive amination of the DCL was applied and the increased formation of one of the different combinations could be shown. In a later proof-of-principle study [62], the concept was extended and the combinatorial SELEX (Systematic Evolution of Ligands by Exponential Enrichment) method to DNA/RNA aptamers was directly combined with the reversible imine formation. The process led to the selection of modified aptamers, conjugated to a range of non-nucleotidic aldehydes. The method proved successful for the selection of small-molecule conjugated RNA aptamers that bind tightly to the TAR element of HIV-1 (Figure 5.27).

5.4.4
Cells

The DCC concept is not restricted to discrete (macro)molecules, but DCLs can also be used to target whole cells. This was first demonstrated with a disulfide-based DCL using the pre-equilibrated approach [65]. Based on studies of the

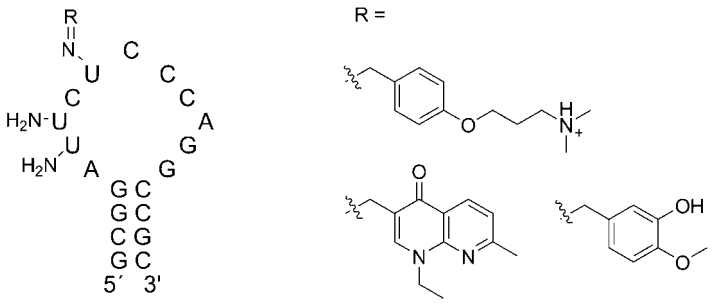


Figure 5.27 Selection of small-molecule conjugated RNA aptamers that bind tightly to the TAR element of HIV-1

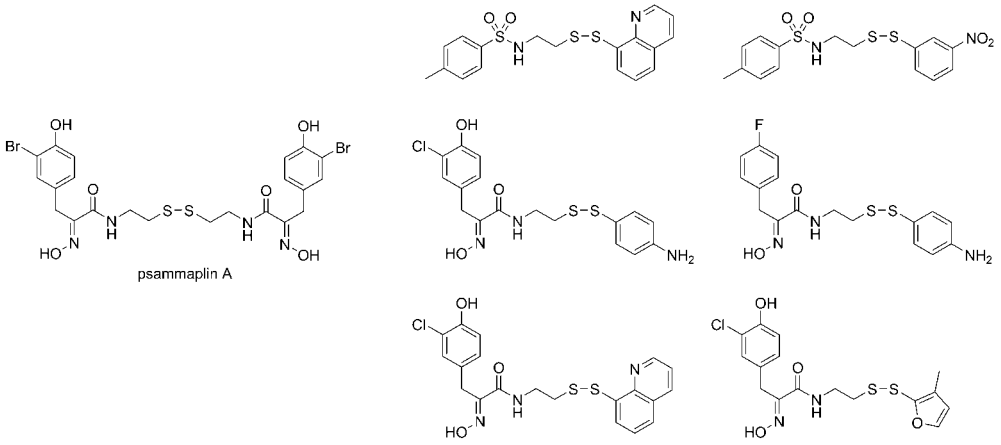


Figure 5.28 Psammaplin A and identified structures.

marine natural product psammaplin A, an antibacterial agent potent against methicillin-resistant *Staphylococcus aureus* (MRSA), a very large ditopic library composed of more than 3000 constituents was designed. Extensive screening of the DCL towards MRSA resulted in the identification of several lead compounds of higher activity than the parent psammaplin A structure (Figure 5.28).

That cell-surface determinants can be targeted with the DCC technique was shown in interesting studies where the bacterial cell wall peptide D-Ala–D-Ala was probed [66, 67]. A DCL based on vancomycin-derived components was used and the libraries were made by linking two peptide-binding units by a linker chain. The vancomycin dimer is known to bind to its ligand more efficiently than the corresponding monomer, and by varying the linker unit, an optimal distance and structure could be identified. Both alkene metathesis and disulfide interchange were used to introduce reversibility in the system, and libraries of up to 36 members were efficiently generated in a pre-equilibrated strategy. The resulting library constituents were subsequently tested for antibacterial activity against a series of vancomycin-resistant bacterial strains and several of the library components were found active (Figure 5.29).

5.5 Conclusions and Future Prospects

In conclusion, it has become evident that DCC is a useful tool for the generation of complex chemical diversity, directly able to interrogate biological target entities. Although the technique is still in its early development phase, a multitude of examples have been presented that show its potential. Thus, a variety of enzymes and other proteins have been targeted, as well as DNA/RNA and even cellular systems. The field is at present in rapid progress, with new biological species being targeted and expanding the frontier of the whole concept. Some challenges still

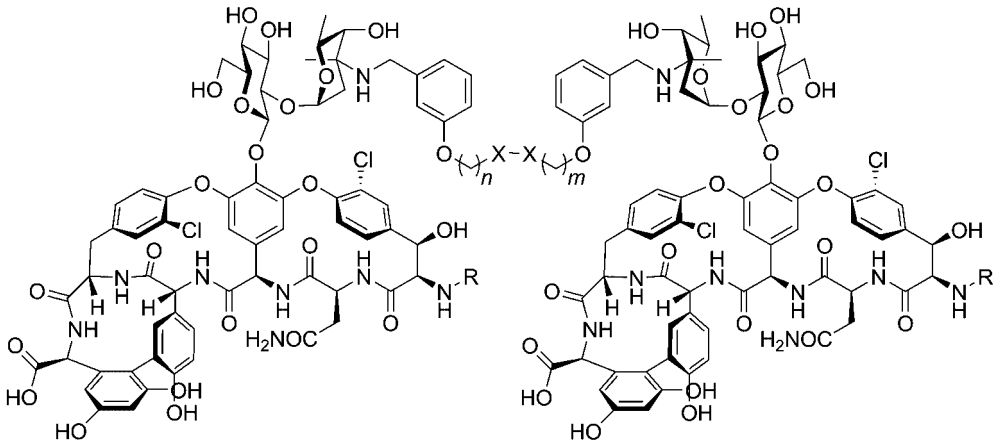


Figure 5.29 Vancomycin dimers targeting cell surface peptides. X–X = S–S, CH=CH.

persist, however, especially concerning the dynamic chemistry involved, where an improved variety of chemoselective reversible reactions are needed. Controllable and rapid chemistries that are entirely compatible with biological conditions, such as solvent, pH, temperature, and so on, and that show no interference with the targets are especially desired. Furthermore, more compact dynamic elements are desirable since biological receptor sites are often quite limited in size, capable of accommodating rather small ligands.

Still, the dynamic ligand assembly concept has indeed opened new perspectives in drug discovery processes, offering a versatile and rapid targeting method endowed with self-screening capability. It furthermore exemplifies the concept of informed molecular diversity where the chosen target entity drives the formation of its own ligand. It thus plays a major role in the emergence of adaptive chemistry [80].

Acknowledgments

Financial support from the Swedish Research Council, the Royal Institute of Technology, and the European Commission (contracts MRTN-CT-2006-35614, and MRTN-CT-2005-19561) is gratefully acknowledged.

References

1 Ugi, I. and Steinbrückner, C. (1961) *Chem. Ber.*, **94**, 734–742.
2 Furka, A. (1982) Notarized Document File 36237 [in Hungarian].
3 Fenniri, H. (ed.) (2000) *Combinatorial Chemistry: A Practical Approach*, Oxford University Press, Oxford.
4 Jung, G. (ed.) (1999) *Combinatorial Chemistry: Synthesis, Analysis, Screening*, Wiley-VCH Verlag GmbH, Weinheim.
5 Terrett, N.K. (1998) *Combinatorial Chemistry*, Oxford Chemistry Masters, Oxford University Press, Oxford.

6 Silverman, R.B. (2004) *The Organic Chemistry of Drug Design and Drug Action*, 2nd edn, Academic Press, New York.

7 Lipinski, C. and Hopkins, A. (2004) *Nature*, **432**, 855–861.

8 Dobson, C.M. (2004) *Nature*, **432**, 824–828.

9 Feher, M. and Schmidt, J.M. (2003) *J. Chem. Inf. Comput. Sci.*, **43**, 218–227.

10 Macarron, R. (2006) *Drug Discov. Today*, **11**, 277–279.

11 Furka, A., Sebestyen, F., Asgedom, M., and Dibo, G. (1988) in *Highlights of Modern Biochemistry: Proceedings of the 14th International Congress of Biochemistry*, VSP, Utrecht, vol. **5**, p.47.

12 Hajduk, P.J. and Greer, J. (2007) *Nat. Rev. Drug Discov.*, **5**, 211.

13 Rees, D.C., Congreve, M., Murray, C.W., and Carr, R. (2004) *Nat. Rev. Drug Discov.*, **3**, 660–672.

14 Erlanson, D.A., McDowell, R.S., and O'Brien, T. (2004) *J. Med. Chem.*, **47**, 3463–3482.

15 Ramström, O. (2007) *Lat. J. Chem.*, **4**, 263–269.

16 Sreenivasachary, N., Hickman, D.T., Sarazin, D., and Lehn, J.-M. (2006) *Chem. Eur. J.*, **12**, 8581–8588.

17 Ruff, Y. and Lehn, J.-M. (2008) *Biopolymers*, **89**, 486–496.

18 Ruff, Y. and Lehn, J.-M. (2008) *Angew. Chem. Int. Ed.*, **47**, 3556–3559.

19 Katz, B.A., Finer-Moore, J., Mortezaei, R., Rich, D.H., and Stroud, R.M. (1995) *Biochemistry*, **34**, 8264–8280.

20 Huc, I. and Lehn, J.-M. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 2106–2110.

21 Poulsen, S.A. and Bornaghi, L.F. (2006) *Bioorg. Med. Chem.*, **14**, 3275–3284.

22 Cheeseman, J.D., Corbett, A.D., Shu, R., Croteau, J., Gleason, J.L., and Kazlauskas, R.J. (2002) *J. Am. Chem. Soc.*, **124**, 5692–5701.

23 Corbett, A.D., Cheeseman, J.D., Kazlauskas, R.J., and Gleason, J.L. (2004) *Angew. Chem. Int. Ed.*, **43**, 2432–2436.

24 Bunyapaiboonsri, T., Ramström, O., Lohmann, S., Lehn, J.-M., Peng, L., and Goeldner, M. (2001) *ChemBioChem*, **2**, 438–444.

25 Larsson, R., Pei, Z., and Ramström, O. (2004) *Angew. Chem. Int. Ed.*, **43**, 3716–3718.

26 Larsson, R. and Ramström, O. (2005) *Eur. J. Org. Chem.*, 285–291.

27 Bunyapaiboonsri, T., Ramström, H., Ramström, O., Haiech, J., and Lehn, J.-M. (2003) *J. Med. Chem.*, **46**, 5803–5811.

28 Congreve, M.S., Davis, D.J., Devine, L., Granata, C., O'Reilly, M., Wyatt, P.G., and Jhoti, H. (2003) *Angew. Chem. Int. Ed.*, **42**, 4479–4482.

29 Cancilla, M.T., He, M.M., Viswanathan, N., Simmons, R.L., Taylor, M., Fung, A.D., Cao, K., and Erlanson, D.A. (2008) *Bioorg. Med. Chem. Lett.*, **18**, 3978–3981.

30 Shi, B. and Greaney, M.F. (2005) *Chem. Commun.*, 886–888.

31 Shi, B., Stevenson, R., Campopiano, D.J., and Greaney, M.F. (2006) *J. Am. Chem. Soc.*, **128**, 8459–8467.

32 Hochgürtel, M., Kroth, H., Piecha, D., Hofmann, M.W., Nicolau, C., Krause, S., Schaaf, O., Sonnenmoser, G., and Eliseev, A.V. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 3382–3387.

33 Hochgürtel, M., Biesinger, R., Kroth, H., Piecha, D., Hofmann, M.W., Krause, S., Schaaf, O., Nicolau, C., and Eliseev, A.V. (2003) *J. Med. Chem.*, **46**, 356–358.

34 Zameo, S., Vauzeilles, B., and Beau, J.-M. (2006) *Eur. J. Org. Chem.*, 5441–5444.

35 Zameo, S., Vauzeilles, B., and Beau, J.-M. (2005) *Angew. Chem. Int. Ed.*, **44**, 965–969.

36 Vongvilai, P., Angelin, M., Larsson, R., and Ramström, O. (2007) *Angew. Chem. Int. Ed.*, **46**, 948–950.

37 Valade, A., Urban, D., and Beau, J.-M. (2007) *J. Comb. Chem.*, **9**, 1–4.

38 Valade, A., Urban, D., and Beau, J.-M. (2006) *ChemBioChem*, **7**, 1023–1027.

39 Danieli, B., Giardini, A., Lesma, G., Passarella, D., Peretto, B., Sacchetti, A., Silvani, A., Pratesi, G., and Zunino, F. (2006) *J. Org. Chem.*, **71**, 2848–2853.

40 Lienard, B.M., Huting, R., Lassaux, P., Galleni, M., Frere, J.M., and Schofield, C.J. (2008) *J. Med. Chem.*, **51**, 684–688.

41 Swann, P.G., Casanova, R.A., Desai, A., Frauenhoff, M.M., Urbancic, M., Slomczynska, U., Hopfinger, A.J., Le Breton, G.C., and Venton, D.L. (1996) *Biopolymers*, **40**, 617–625.

42 Sakai, S., Shigemasa, Y., and Sasaki, T. (1997) *Tetrahedron Lett.*, **38**, 8145–8148.

43 Sakai, S., Shigemasa, Y., and Sasaki, T. (1999) *Bull. Chem. Soc. Jpn.*, **72**, 1313–1319.

44 Ramström, O. and Lehn, J.-M. (2000) *ChemBioChem*, **1**, 41–48.

45 Ramström, O., Lohmann, S., Bunyapaiboonsri, T., and Lehn, J.-M. (2004) *Chem. Eur. J.*, **10**, 1711–1715.

46 Sando, S., Narita, A., and Aoyama, Y. (2004) *Bioorg. Med. Chem. Lett.*, **14**, 2835–2838.

47 Hotchkiss, T., Kramer, H.B., Doores, K.J., Gamblin, D.P., Oldham, N.J., and Davis, B.G. (2005) *Chem. Commun.*, 4264–4266.

48 Pei, Z., Larsson, R., Aastrup, T., Anderson, H., Lehn, J.-M., and Ramström, O. (2006) *Biosens. Bioelectron.*, **22**, 42–48.

49 André, S., Pei, Z., Siebert, H., Ramström, O., and Gabius, H. (2006) *Bioorg. Med. Chem.*, **14**, 6314–6326.

50 Lins, R.J., Flitsch, S.L., Turner, N.J., Irving, E., and Brown, S.A. (2004) *Tetrahedron*, **60**, 771–780.

51 Lins, R.J., Flitsch, S.L., Turner, N.J., Irving, E., and Brown, S.A. (2002) *Angew. Chem. Int. Ed.*, **41**, 3405–3407.

52 Milanesi, L., Hunter, C.A., Sedelnikova, S.E., and Waltho, J.P. (2006) *Chem. Eur. J.*, **12**, 1081–1087.

53 Shibakami, M., Inagaki, M., and Regen, S.L. (1998) *J. Am. Chem. Soc.*, **120**, 3758–3761.

54 Klekota, B., Hammond, M.H., and Miller, B.L. (1997) *Tetrahedron Lett.*, **38**, 8639–8642.

55 Klekota, B. and Miller, B.L. (1999) *Tetrahedron*, **55**, 11687–11697.

56 Karan, C. and Miller, B.L. (2001) *J. Am. Chem. Soc.*, **123**, 7455–7456.

57 Bugaut, A., Toulme, J., and Rayner, B. (2004) *Angew. Chem. Int. Ed.*, **43**, 3144–3147.

58 Bugaut, A., Bathany, K., Schmitter, J., and Rayner, B. (2005) *Tetrahedron Lett.*, **46**, 687–690.

59 Whitney, A.M., Ladame, S., and Balasubramanian, S. (2004) *Angew. Chem. Int. Ed.*, **43**, 1143–1146.

60 McNaughton, B.R. and Miller, B.L. (2006) *Org. Lett.*, **8**, 1803–1806.

61 McNaughton, B.R., Gareiss, P.C., and Miller, B.L. (2007) *J. Am. Chem. Soc.*, **129**, 11306–11307.

62 Bugaut, A., Toulme, J., and Rayner, B. (2006) *Org. Biomol. Chem.*, **4**, 4082–4088.

63 Ladame, S., Whitney, A.M., and Balasubramanian, S. (2005) *Angew. Chem. Int. Ed.*, **44**, 5736–5739.

64 Bugaut, A., Jantos, K., Wietor, J., Rodriguez, R., Sanders, J.K.M., and Balasubramanian, S. (2008) *Angew. Chem. Int. Ed.*, **47**, 2677–2680.

65 Nicolaou, K.C., Hughes, R., Pfefferkorn, J.A., Barluenga, S., and Roecker, A.J. (2001) *Chem. Eur. J.*, **7**, 4280–4295.

66 Nicolaou, K.C., Hughes, R., Cho, S.Y., Winssinger, N., Labischinski, H., and Endermann, R. (2001) *Chem. Eur. J.*, **7**, 3824–3843.

67 Nicolaou, K.C., Hughes, R., Cho, S.Y., Winssinger, N., Smethurst, C., Labischinski, H., and Endermann, R. (2000) *Angew. Chem. Int. Ed.*, **39**, 3823–3828.

68 Rotenberg, S.A., Calogeropoulou, T., Jaworski, J.S., Weinstein, I.B., and Rideout, D. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 2490–2494.

69 Rideout, D., Jaworski, J., and Dagnino, R. (1988) *Biochem. Pharmacol.*, **37**, 4505–4512.

70 Rideout, D. (1986) *Science*, **233**, 561–563.

71 Rideout, D., Calogeropoulou, T., Jaworski, J., and Mccarthy, M. (1990) *Biopolymers*, **29**, 247–262.

72 Nguyen, R. and Huc, I. (2001) *Angew. Chem. Int. Ed.*, **40**, 1774–1776.

73 Yu, Y., Ye, L., Haupt, K., and Mosbach, K. (2002) *Angew. Chem. Int. Ed.*, **41**, 4459–4463.

74 Lewis, W.G., Green, L.G., Grynszpan, F., Radic, Z., Carlier, P.R., Taylor, P., Finn, M.G., and Sharpless, K.B. (2002) *Angew. Chem. Int. Ed.*, **41**, 1053–1054.

75 Manetsch, R., Krasinski, A., Radic, Z., Raushel, J., Taylor, P., Sharpless, K.B., and Kolb, H.C. (2004) *J. Am. Chem. Soc.*, **126**, 12809–12818.

76 Mocharla, V.P., Colasson, B., Lee, L.V., Roper, S., Sharpless, K.B., Wong, C.H., and Kolb, H.C. (2005) *Angew. Chem. Int. Ed.*, **44**, 116–120.

77 Krasinski, A., Radic, Z., Manetsch, R., Raushel, J., Taylor, P., Sharpless, K.B., and Kolb, H.C. (2005) *J. Am. Chem. Soc.*, **127**, 6686–6692.

78 Whiting, M., Muldoon, J., Lin, Y.C., Silverman, S.M., Lindstrom, W., Olson, A.J., Kolb, H.C., Finn, M.G., Sharpless, K.B., Elder, J.H., and Fokin, V.V. (2006) *Angew. Chem. Int. Ed.*, **45**, 1435–1439.

79 Goodwin, J.T. and Lynn, D.G. (1992) *J. Am. Chem. Soc.*, **114**, 9197–9198.

80 Lehn, J.-M. (1999) in *Supramolecular Science: Where It Is and Where It Is Going* (eds R. Ungaro and E. Dalcanale), NATO ASI series, vol. **527**, Kluwer, Dordrecht, pp. 273–286.

6
Polymers Formed by Dynamic Combinatorial Chemistry

Jean-Luc Wietor, Nicole E. Botterhuis, and Rint P. Sijbesma

6.1
Introduction

Applying the principles of dynamic combinatorial chemistry (DCC) in polymers is a highly promising approach to functional systems that takes advantage of many of the specific features of long-chain molecules. Key elements of DCC – diversity, reversibility, and the possibility for molecular recognition via multiple interactions – are either innately present or can be engineered in polymers. Polymers have intrinsic diversity because even the best polymerization techniques give rise to a significant dispersion in the degree of polymerization. Particularly when the polymer chain is formed by reversible linking of monomers, the statistics of the polymerization process lead to a "most probable distribution" of chain lengths with a high polydispersity, also known as the Flory distribution [1]. Diversity is further enhanced by the wide variety of monomers that can be used to make copolymers, giving rise to a huge number of sequences of even a single chain length. (As an example the combinatorial synthesis of a 20mer oligopeptide from a mixture of 20 natural amino acids would theoretically give rise to a library with a size of 20^{20} molecules – more molecules than can be made out of 1 kg of each of the amino acids.)

The second key element of DCC – reversibility – is a feature that is intrinsic to many polymers, specifically those that are formed by step polymerizations. Although the conditions under which bond formation takes place for the most common functional groups (esters, amides) are quite harsh, a large variety of bonding forming reactions that are reversible under mild conditions have been studied recently and have become known under the name of dynamic covalent chemistry. Reversibility is an even more prominent feature of supramolecular polymers – polymers in which the monomeric units are held together by noncovalent bonds.

Finally, the third key element of DCC – molecular recognition via multiple interactions – is a fundamental feature of polymers. Due to their size and architectural flexibility, polymers are able to combine a large number of molecular recognition sites in virtually any spatial arrangement – the abundance of nucleic acids and

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

proteins for molecular recognition in natural systems reflects the intrinsic capability of polymers for selectivity and strong binding of a wide variety of guest molecules.

Despite these favorable characteristics, relatively few systems have been described that combine all key elements of DCC in a single system and the potential of dynamic combinatorial libraries (dynamic combinatorial library DCLs) of polymers is just beginning to be explored. The current chapter outlines how reversibility and selectivity have been employed separately to develop functional polymers, and how these elements are being combined to amplify specific sequences of monomers in chains or networks. In the following section of this chapter, a closer look will be taken at the use of dynamics in polymers, discussing dynamic covalent polymers and supramolecular polymers. In the subsequent section, the focus will be on different aspects of molecular recognition in polymeric systems, both in reversible systems, where selective binding interactions lead to the preferential formation of specific compositions, and in irreversible systems, where templates have extensively been used to obtain molecular imprinted materials. Finally, the outlook discusses the challenges to develop polymeric systems that combine all essential features of DCC and the need for fundamentally new approaches to conquer entropic limitations intrinsic to DCC in polymerics.

6.2
Dynamics in Polymers

6.2.1
Exchangeable Covalent Bonds – A Familiar Feature in Polymer Chemistry

The reversible nature of some polymerization reactions has been recognized and studied for more than half a century. However, most of the time, the exchangeable character has been considered a nuisance, and only the advent of dynamic covalent chemistry [2] has promoted the reversibility of transesterifications and the like to the status of an auspicious feature. Nonetheless, most polymerization reactions only reverse under conditions too harsh for successful templating and compatibility with other functional groups.

In the following, we provide a commented overview of the reversible reactions that have found use in polymers or that we consider promising.

Table 6.1 lists exchange reactions that have been used in reversible polymers, the exchange stimuli required as well as references to key publications and reviews.

In the following, we will give a brief account of the advantages and drawbacks of the three large classes of chemistries – noncovalent, coordinative, and exchangeable covalent – used for the generation of dynamic polymers, in the order of increasing bond strengths. As a detailed overview would go beyond the scope of this chapter, we direct the interested reader to a series of review papers [10, 28–30, 39].

Table 6.1 Reversible covalent bonds used in DCC of polymers.

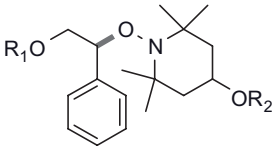
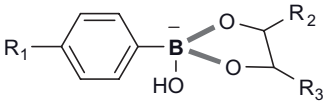
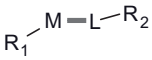
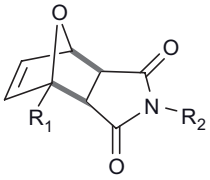
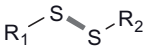
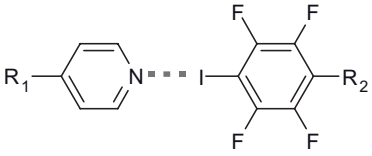
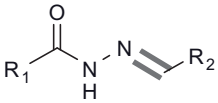
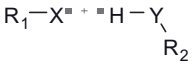
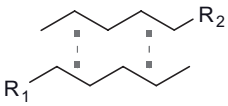
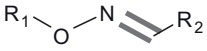
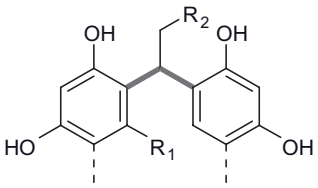
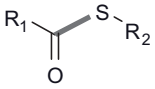
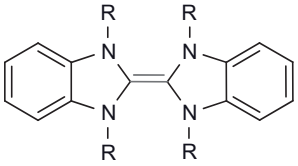
Name	Reaction scheme	Exchange stimulus	References
Alkoxyamines		heat	[3–7]
Boronic esters		base	[8]
Coordination bonds			[9–12]
Diels–Alder		heat	[13–15]
Disulfides (anionic)		thiolate	[16, 17]
Halogen bonding			[18]
Hydrazones		acid	[19–27]
Hydrogen bonding			[28–30]
Hydrophobic interactions			[31]
Imines		acid	[32–34]
Oximes		acid	[35]

Table 6.1 *Continued.*

Name	Reaction scheme	Exchange stimulus	References
Resorcinols		heat, acid	[36]
Thioesters		thiolate	[37]
Tetraaminoethylenes		heat	[38]

6.2.2
Noncovalent Chemistry – Supramolecular Polymers

Hydrogen-bonded supramolecular polymers [28–30], especially those using units with multiple donor and acceptor sites, have become the most versatile class of supramolecular polymers. A broad range of possible geometries and highly directional interactions have been among the advantages of hydrogen bonding, alongside the possibility to tune binding strength by varying the number of hydrogen bonds and the solvent. In addition to that, hydrogen bonding units can be designed to be either complementary or self-complementary, rendering a variety of geometries and structures accessible. The multiple hydrogen bond units generally rely on conjugated or polyaromatic scaffolds, which provide the required geometrical preorganization of the hydrogen bond donors and acceptors. Figure 6.1 lists some of the most widely used moieties [40–44].

Despite their undeniable advantages, one of the drawbacks of hydrogen bonds is the short timescale on which exchange takes place. For diffusion controlled reactions (which hydrogen bond formation usually is), there is a simple relation between binding constant and lifetime of the bond. Being relatively weak, hydrogen-bonded complexes, even if based on multiple hydrogen bonding units, usually break and reform at timescales well below 1 ms. This feature makes it often impossible to isolate or even identify the favored members of a library, as they will re-equilibrate as soon as separated from their template. Indeed, published examples of templating with hydrogen-bonded polymers are scarce.

Other noncovalent interactions. such as the hydrophobic effect, ionic interactions, or π–π interactions. are even weaker than hydrogen bonding, and suffer

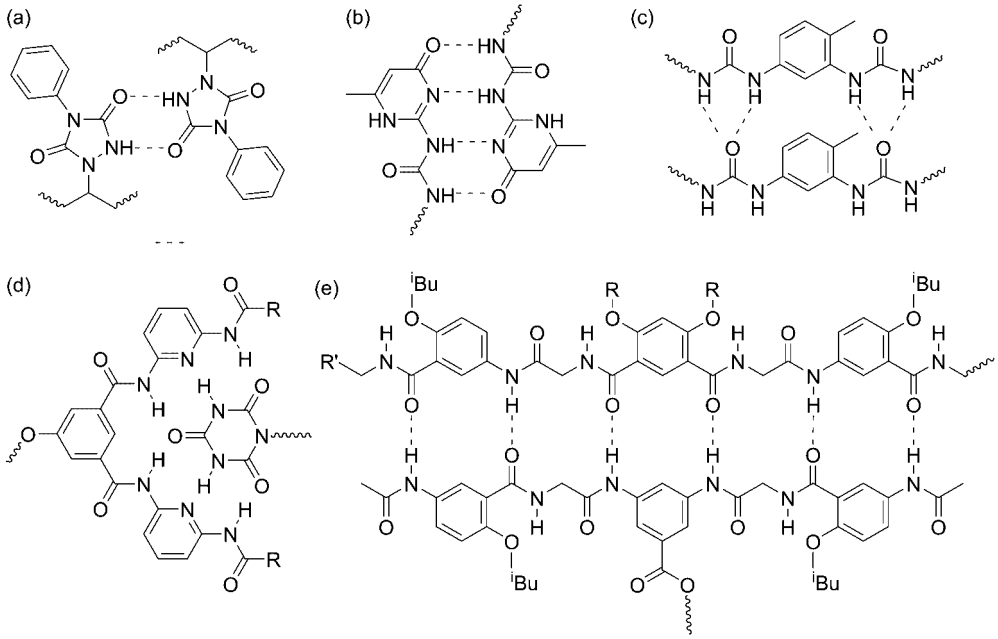


Figure 6.1 Selected examples of multiple hydrogen bonding motifs used in polymers: (a) urazole [40], (b) ureidopyrimidinone [41], (c) bis-urea [42], (d) Hamilton's wedge [43], and (e) linear arrays developed by Gong [44]

from a lack of thermodynamic and kinetic stability. Therefore, these binding interactions have hardly been used on their own.

6.2.3
Coordination Polymers

The use of coordinative bonds can lead to varied electrochemical, optical, or magnetic properties, and offers a wide range of thermodynamic and kinetic stabilities [10, 39]. Among the notable features of many systems is the possibility to radically change (e.g., completely freeze) exchange kinetics by electrochemically changing the degree of oxidation of the metal center or by substituting one metal by another, largely without changing the coordination geometry [45, 46]. Coordination polymers have recently been discussed in a tutorial review [47].

6.2.4
Exchangeable Covalent Bonds

The attraction of exchangeable covalent bonds is beyond any doubt their stability in the absence of an exchange stimulus, be it heat, acid, or base, or other catalytic species [2, 48, 49]. The still somewhat limited list of dynamic covalent bonds is

likely to be expanded over the next few years and the advent of new generations of exchange catalysts ought to allow scrambling of some functionalities under less stringent conditions. Promising reactions include transamidation, for which mild exchange catalysts are being developed [50, 51], acetal exchange [52], and ultraviolet-triggered radical disulfide exchange – a reaction proven suitable in small-molecule reactions [53] and in protein-folding studies [54–56].

6.2.5
Reversible Covalent Chemistry in Polymers

Dynamic polymers containing main-chain hydrogen bonds have been studied extensively, among others thanks to their peculiar, but outstanding material properties. However, as the ultimate goal in DCC comprises templation and isolation of particular sequences or conformations, we focus here on the kinetically more stable main-chain covalent polymers.

Using the radical alkoxyamine chemistry, the group of Takahara have developed dynamic polymers in several geometrical variations. A nitroxide radical and a styryl radical reversibly form an alkoxamine (Table 6.1), which acts as a covalent link between two polymer chains. This functionality has proven successful for (Figure 6.2a) copolymer formation [3], (Figure 6.2b) synthesis of graft polymers [4], and (Figure 6.2c) cross-linking of polymer strands [6], and tuning of molecular weight and molecular weight distribution [7]. In solution, these materials undergo ring-chain equilibria, the position of which is changed by varying the concentration [5].

The research group of Jean-Marie Lehn have also published numerous papers on dynamic polymers, both noncovalent [57, 58] and covalent. In this section, we will give a nonexhaustive list of examples of dynamic polymers.

In addition to self-assembling helical polymers based on hydrazones [19, 26], the Lehn group conducted several studies of exchange of monomers between a

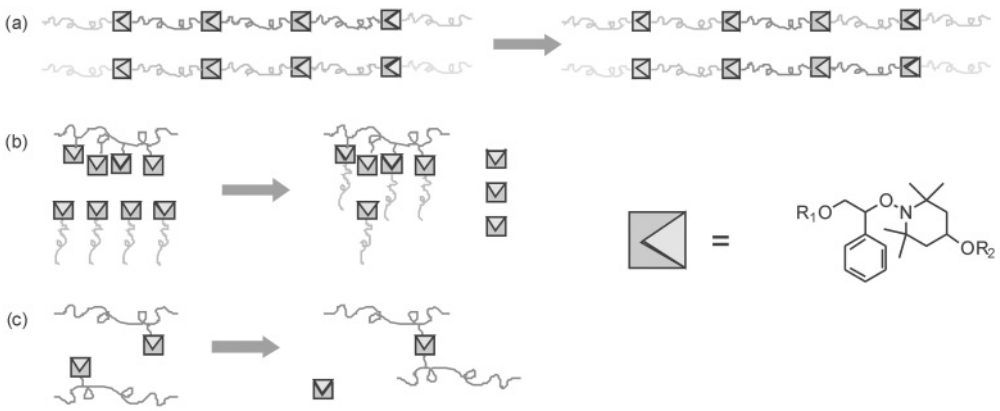


Figure 6.2 Selected geometries used in the dynamic polymers by Takahara *et al.*: (a) copolymer formation from a blend, (b) graft polymerization, and (c) cross-linking of polymer chains.

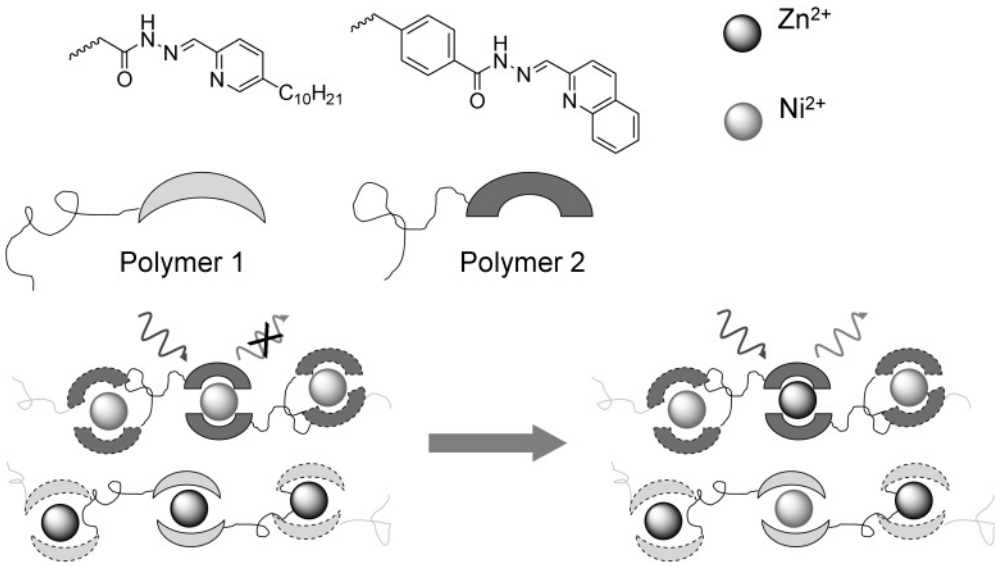


Figure 6.3 Constituents of polymers 1 and 2, and cartoon of the cation migration.

polymer and a monomer or between two polymers, exploiting several types of polymers and exchange reactions. It has been shown that such cross-over reactions occur in hydrazone polymers in the presence of some acid in solution [27] or even in the solid state at longer time-scales [24]. In one study (incidentally combining an acyl hydrazide and an aromatic hydrazide), heating induced a fluorescent response at the interface between two polymers due to copolymer formation [22]. The potential of such recombining systems for self-healing properties has been put forward; however, it remains unclear whether the mobility necessary for recombination does not jeopardize material properties, most notably retention of shape.

Imine-based dynamers of different types have also been investigated. On fluorene-based dynamers, fundamental studies were conducted on the influence of Zn^{2+} and H^{+} ions on the constitution, chain length, and, above all, optical properties [33, 34]. In a different study, two bis-aldehyde and two bis-amine oligomers were mixed to give all four copolymers in solution, as opposed to only two in the bulk material. This was ascribed to crystallization of one of the combinations, leaving only one possible combination for the two oligomers remaining in the amorphous phase [32]. If this gives an impression of self-templating, one should keep in mind that crystallization should be regarded as a kinetic trap.

The same authors used the tridentate chelating sites created by an acyl hydrazide condensed with a pyridyl aldehyde to complex metal ions, thereby combining covalent and coordination linkages within one system. As in their earlier work, luminescent properties of films were changed by simply bringing the films into contact. When two nonluminescent films (polymer 1 and 2, Figure 6.3) were stacked on each other and allowed to blend at moderate temperatures, two

nonfluorescent polymers blended into fluorescent material. Initial polymer 1 did not dispose of a fluorophore, whereas polymer 2 had a (fluorescent) quinoline unit, coordinated to a quenching nickel cation. When the metal ion quencher was allowed to diffuse away into polymer 2, fluorescent activity of the quinoline moiety was restored [9].

More recently, the dynamer concept was applied to the generation of analogs of biopolymers: hydrazone and oxime analogs of oligoglycosides and oligonucleotides. Arabinose was functionalized with either an alkoxyamine and a masked aldehyde (A–B monomer approach) to give a directional homopolymer, or with two alkoxyamines or two (masked) aldehydes (A–A + B–B approach) to give an alternating copolymer without directional orientation of the monomers [35]. Alternatively, the carbohydrates can be side-chains of the polymer if they are appended to the main-chain exchanging monomer [25].

Similarly, using hydrazone chemistry, modified nucleobases were incorporated into a dynamic analog of peptide nucleic acid [20]. Recently, the relatively weak covalent bonds of ene tetramines have been employed to develop reversible covalent polymers based on carbene dimerization [38].

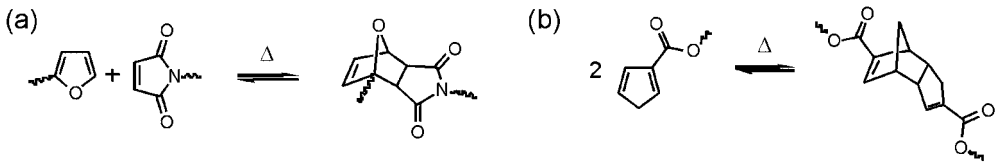
6.2.6
Dynamics Put to Use in Self-Healing Materials

Self-healing materials (SHMs) are highly desirable targets in material research that are actively sought after by polymer chemists [59, 60]. Although SHMs and DCC are clearly distinct concepts, the traits they share make a brief discussion worthwhile.

Finding a compromise between liquid (thus perfectly self-healing) and solid (shape-retaining) materials is one approach to SHMs, but reconciling these seemingly contradictory properties is not an easy task. Nevertheless, several promising SHMs reported to date rely on the dynamic character of exchangeable covalent linkages or hydrogen bonds.

6.2.6.1 **Reversible Covalent Approach**

Covalent bonds make the material strong and solid-like, but rearrangement of the covalent bonds is necessary in order to achieve mending of cracks – the fracture surfaces have to be able to flow into each other. The group of Fred Wudl has developed such materials, relying on a reversible Diels–Alder reaction that reshuffles bonds at high temperatures and allows crack repair. A first material was based on a furane–maleimide system (Scheme 6.1a), in which molecules functionalized with three diene and four dienophile moieties, respectively, were reacted to generate a highly cross-linked network. When a cracked specimen was heated to 120–150 °C for 2 h, cracks largely heal and the material regains more than half of its initial strength [13]. Further improvement of the system allowed 81% recovery after 30 min at 115 °C [15]. More recently, a system based on the formation of homodimers and homotrimers of cyclopentadiene was published [14]. The possibility of trimerizations and the identity of diene and dienophile allows for



Scheme 6.1 (a) Asymmetrical and (b) symmetrical reversible Diels–Alder reactions used in SHMs.

network formation from a single component, with a largely simplified synthesis. However, the material and self-healing properties of this material require further optimization.

6.2.6.2 **Noncovalent Approach**

Another approach relies on the dynamic nature of hydrogen-bonding interactions. The group of Ludwik Leibler recently reported a system that exhibits repeated self-healing, without requiring external stimuli such as heat or displaying liquid-like behavior [61]. In this system, dimer and trimer acids were reacted with diethylene triamine and urea to yield a complex mixture of small molecules with strong hydrogen bond donors and acceptors (Figure 6.4). When this material is plasticized with dodecane, a soft rubber is obtained, which can be cut or broken and mended by joining the cut or broken surfaces and allowing the material to heal for a few hours. The material recovers essentially all of its strength and elasticity.

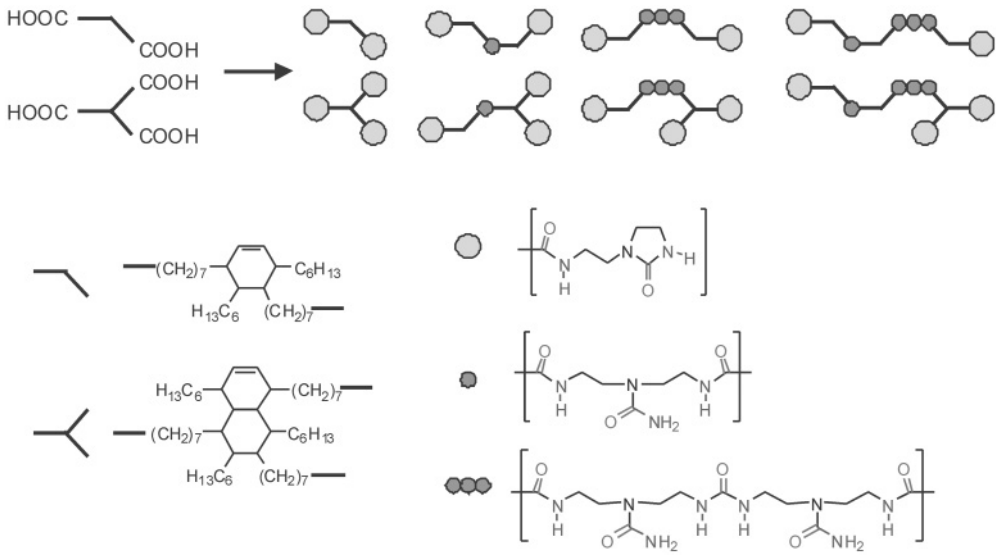


Figure 6.4 Schematic representation and molecular structures of the components of the self-healing rubber developed by Leibler.

An approach based on coordinative bonds was followed by Bielawski *et al.*, who developed electrically conductive coordination polymers that can be healed after damage by thermal treatment [62]. Incidentally, one further successful approach does not rely on dynamic features within the material. It simply carries microencapsulated monomer within its matrix, a polymer, which contains a polymerization catalyst. When a crack propagates through the polymer matrix, monomer is released and polymerized within the crack, restoring the initial material properties of the sample [63, 64].

6.3 Biasing Composition by Molecular Recognition

Complete control over the composition of and precise positioning of functional groups in polymers are essential features for the faithful replication and transcription of genetic information stored in DNA. To achieve the required level of fidelity, Nature uses a combination of molecular recognition and nonequilibrium proofreading processes. While realizing the latter in synthetic polymers will remain an extremely challenging goal for many years to come, much progress has been made in the use of recognition units to direct the composition of self-assembled polymeric systems. Self-sorting and orthogonal self-assembly are two of the concepts we will discuss here. These are self-assembly strategies to bias polymer composition and structure based on internal stimuli, instead of the templates that represent the external stimulus in DCC.

6.3.1 Self-Sorting

Control over the position of functional groups in polymers can be obtained by using self-sorting in multifunctionalized polymers [65–70]. Self-sorting is the ability of an entity to distinguish between self and nonself. This principle is found in nature in the replication of DNA and in the crystallization of racemates into conglomerates (mixtures of enantiomerically pure crystals of one enantiomer and its opposite), but even phase separation of oil and water can be considered as a self-sorting process. Isaacs *et al.* have investigated the self-sorting capacity in mixtures of several synthetic host–guest complexes in organic solvents and water [71–75]. They were able to show that each guest molecule binds its own host molecule with high fidelity in a complex mixture of hosts and guests.

In the formation of complexes via self-sorting, no external template is used. Instead, the molecular structure of the host can be considered as the template for the preparation of a host–guest complex. In this case, the amplification of the most stable complex takes place via intermolecular interactions, such as hydrogen bonding, metal coordination, or ionic interactions, between the different components in the mixture. Since the strength of these supramolecular interactions is known to be influenced by substituents, a dynamic library can be obtained – not

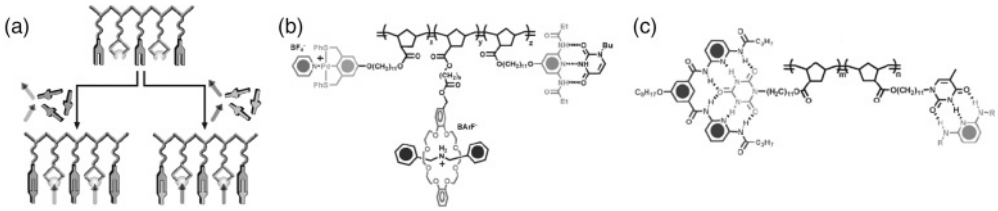


Figure 6.5 Self-sorting in polymers. (a) Schematic overview of a noncovalent approach to different copolymers from a generic polymer backbone. (b) Noncovalently multifunctionalized polynorborene terpolymer, via metal coordination (palladium pincer), pseudorotaxane formation (crown ether), and hydrogen bonding arrays (diaminopyridine–thymine). (c) Noncovalently multifunctionalized polynorborene polymer, via competitive hydrogen bonding receptors (diaminopyridine–thymine and cyanuric acid–isophthalic wedge). Reproduced from South *et al.* [66].

only by using different supramolecular interactions, but also by changing the substituents of the binding units.

Whereas the group of Isaacs has studied self-sorting in mixtures of small molecules, others have extended this approach to self-sorting in (block) copolymers [66–70, 76–81] and on surfaces [82]. Weck *et al.* synthesized polynorborenes with host groups that bind different guest molecules based on metal–ligand interactions, hydrogen bonding ion–dipole interactions, or ionic complexes (Figure 6.5). Association was observed to occur preferentially between the established host–guest complexes in a mixture of different hosts and guests. Even if a polynorbornene was functionalized with two similar hydrogen bonding hosts (Figure 6.5c), the corresponding guests were capable of selective binding to the best matching hosts [78].

Self-sorting under kinetic control has also been reported [75]. However, due to the dynamic nature of the noncovalent interactions used in polymer self-assembly, the products of self-sorting usually cannot be isolated when they form in solution.

Self-recognition of the urea group by hydrogen bonding is used in less dynamic environments such as organogels [28, 83–89] and in polymers. In the latter medium, the selectivity of the binding of guests with two urea groups has been studied extensively in the groups of Meijer and Sijbesma [90–92], because it is an easy, modular approach towards the functionalization of polymers. Matching of the spacer length between the two urea groups is a crucial factor in molecular recognition. A negligible amount of bis-urea guest was washed out from a bis-urea polymer film if the spacer lengths of the bis-urea polymer and a bis-urea guest were the same, whereas almost 70% of the bis-urea guest was lost if spacer lengths were different [92]. This type of self-sorting has been named "narcissistic" self-sorting, since the guest and host molecule have the same recognition unit, as opposed to "social" self-sorting in host–guest systems such as described by the groups of Isaacs [71–75] and Weck [66–69, 78, 79]. Orthogonality can be introduced in this bis-urea-based system by using a mixture of two bis-urea polymers with different spacer lengths [93]. It was shown that self-sorting occurs in a mixture of

a heptylene-spaced and a butylene-spaced bis-urea guest and their matching bis-urea polymers.

Molecular recognition was used in a very elegant way to direct block copolymer morphology by Shenhar *et al.* [80]. Upon addition of dendritic wedges with thymine units to a lamellar block copolymer with diaminopyridine units, a block copolymer morphology was obtained after evaporation of the solvent.

Self-sorting in polymers is not only of interest for multifunctionalization of polymers. Pollino *et al.* have shown that both functionalization and cross-linking can be done noncovalently in one step/pot by using the orthogonal assembly of palladium pincers with pyridine groups and thymine groups with diaminopyridine groups [94].

6.3.2 Templating

One important approach to using a template in order to get nonrandom, nonstatistical distributions or sequences in polymers is template polymerization. In this approach, the template is a molecule that bears different functional groups in an order and geometrical arrangement such that monomers with complementary functional groups can arrange next to each other and be polymerized in the desired sequence. Of course, such a template would much less act in the "guest" mode used in many DCLs, but rather be tailormade for a particular sequence. In an ideal case, this system has an in-built mechanism to correct sequence errors.

6.3.2.1 Templating in Biopolymers

As so often, Nature has developed a remarkably efficient system, although one which does not operate under thermodynamic conditions. DNA polymerases proceed at high speed, with impressive catalytic turnover and almost papal infallibility. When provided with a single-stranded template and suitable monomers (nucleotide triphosphates), DNA polymerases synthesize the complementary DNA polymer. Some DNA polymerases (e.g., the prokaryotic Pol I) are also equipped with a proofreading ability. If an error occurs in the polymerization, another site on the same enzyme gets activated and cuts out the erroneous base. Translation from messenger RNA to a protein on the ribozyme using transfer RNA is a similar if still more complex system! An author with an inclination towards pomposity might conclude that there would be no life without template polymerization.

The ingeniousness of Nature has been adapted to the polymerase chain reaction – a technique used in clinical and forensic science to replicate minute amounts of nucleic acids until analyzable quantities are obtained [95].

6.3.2.2 Templating in Non-Natural Polymers

The emulation of Natures' perfection in artificial systems has been pursued at different levels of sophistication. When an irreversible reaction is used to polymerize the monomers, correctness of the sequence of the polymer must be ensured by the thermodynamically controlled binding of the monomers to specific sites on the template, conditioned by the absence of any reaction in the absence of the

template. However, error correction is no longer possible when the polymerization reaction has taken place. This approach can be compared to a small-molecule DCL using reductive imination to link the building blocks in the presence of a template. This is the usual approach in templated polymerizations and several approaches have been reported [66, 96–99]. As an example, the polymer that was formed using a template approach by South and Weck is discussed here [99]. In this work, the rate of a ring-opening polymerization metathesis polymerization was enhanced via hydrogen bonding of the monomers to a polynorborene template, thereby increasing the local monomer concentration. Furthermore, the addition of the template during the polymerization reduced the polydispersity index of the resulting polymer from 1.73 to 1.19. However, the degree of polymerization of the polymers was not influenced by the degree of polymerization of the template, indicating that the method results in perfect inverse replicas from a template.

In general, this goal cannot be obtained with an approach based on equilibrium self-assembly. Even if the selectivity of binding is very high (so-called orthogonal self-assembly), finite binding energy differences between correct and incorrect replicas will result in a Boltzmann distribution of different products [100]. Although a significant yield of the best binding replica can be obtained even in complex mixtures [48], the near-exclusive formation of exact replicas requires nonequilibrium proofreading concepts like those used in Nature for the replication and transcription of the genetic information stored in DNA [101]. Such approaches have recently begun to be adapted in synthetic systems, albeit by making use of Nature's nucleic acid machinery [102, 103].

6.3.3
Molecularly Imprinted Polymers

Molecular imprinting [104–108] is one of the fields in which DCC is expected to flaunt its full power. Molecularly imprinted polymers (MIPs) are materials that have a particular affinity for certain small or larger molecules thanks to in-built binding sites. The underlying principle is that monomers containing a functionality able to bind (covalently or not) to a molecular template (or "print molecule") form a cavity around the template. The functional monomers can be copolymerized with a large excess of nonfunctional monomers and thus create a polymer matrix with embedded recognition units. When the template molecule is removed, cavities with a defined binding affinity remain; MIPs can be used for analytical, chromatographical, or catalytic purposes.

Current limitations of molecular imprinting include the frequency of errors in cavity shapes and functionality as well as other defects, due to the irreversibility of the polymerization step. As a consequence, selectivities are lower and loading capacity is diminished. If indeed a reversible step could be incorporated into the process of molecular imprinting, some of the major weaknesses of this technique could be overcome.

Among the sundry approaches to tackle this issue [108], reversible disulfide bond formation was proven efficient in a seminal publication in 2001 [109]. Building on earlier work by Wulff and Schulze on chirally selective polystyrenes cross-

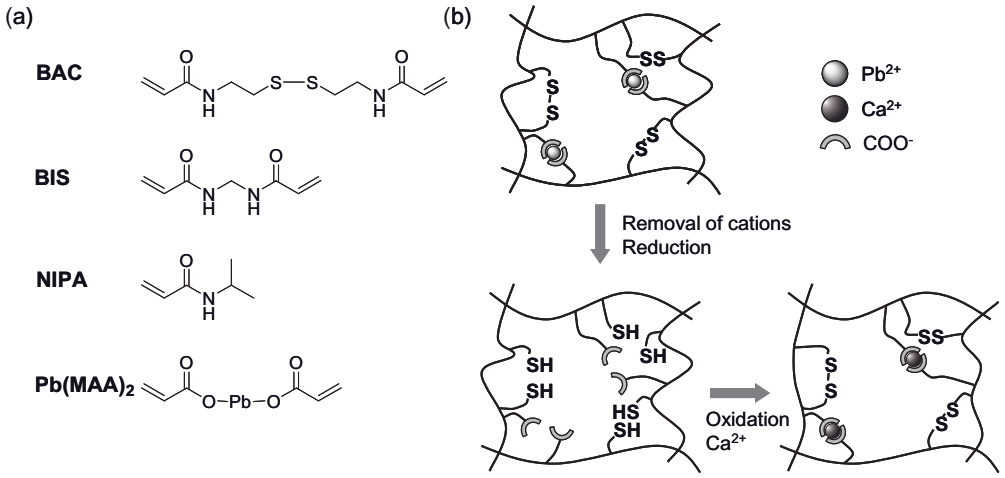


Figure 6.6 (a) Components used to make calcium-selective imprinted gels and (b) schematic of the imprinting process.

linked with disulfide bridges [110], Hiratani *et al.* designed a polymer selective for Ca^{2+} ions, using bis-chelation of the calcium ions by carboxylate residues. Polymerization of the components shown in Figure 6.6 gives a polymer gel from which the lead ions (kinetically stable cross-links under polymerization conditions) can be removed and the disulfide bridges reduced to thiols. Subsequently, the gel is soaked with calcium ions, which reversibly choose the pair of carboxylic acid residues that "fits" best. Subsequently, fixation of the network is achieved via oxidation of the thiols into disulfides.

In this strategy reminiscent of the reduction–oxidation steps in producing curls in hair by a "permanent wave", the calcium ions play the role of the curlers – they reorganize the available carboxylate groups in an optimal geometry, subsequently immobilized by reoxidation of the mercapto cross-linkers. This reduction–oxidation treatment results in a significant increase in the affinity for calcium ions. In contrast, when the gel was reoxidized in the absence of the calcium template, affinities for calcium dropped.

In our opinion, one study in template polymerization deserves special emphasis in the context of DCC. This study by Li and Lynn [111] reports on the templated oligomerization of modified dinucleotides on a complementary octanucleotide The authors use a thymidine T_1 bearing an amine and an aldehyde function or the amide-linked non-natural dinucleotide T_2 (Figure 6.7) and a solid-supported octanucleotide A_8 with eight adenine residues as the template. When either of the building blocks T_1 and T_2 and the template are mixed in aqueous sodium cyanoborohydride, species with eight thymidine bases are exclusively formed, either the octamer of T_1 or the tetramer of T_2. The solid-supported template facilitates removal of formed products and recycling of the template, and hence multiple synthesis cycles.

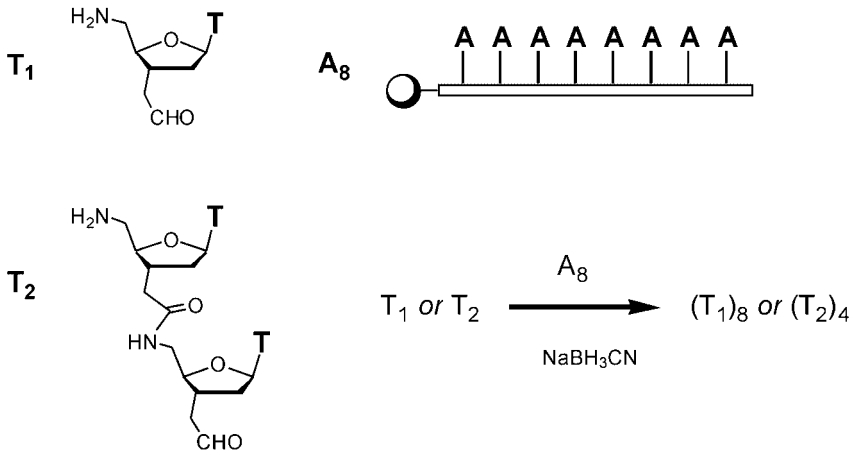


Figure 6.7 Reversible template polymerization of amine and aldehyde bearing thymidine T_1 and amide linked dinucleotide T_2.

The reaction was shown to follow step-growth kinetics in its early stages and to stop abruptly at the length defined by the template. No products were detected in the absence of any template. As the hydrogen bonding between the nucleobases and the imine formation are reversible, error correction occurs to some extent in this system. In this sense, this report matches the spirit of DCC more closely than many others. Unfortunately, no follow-up using the entire "nucleic acid alphabet" has been published hitherto.

6.4 Conclusions and Outlook

Taking into account the intrinsic potential to apply polymers for the goals of DCC, remarkably little work has been done on synthetic polymers that combine diversity, reversibility, and the possibility for molecular recognition via multiple interactions. In this chapter, we have discussed several examples of polymers that feature at least one of those functions. The task to us, as polymer scientists, is now to unite these approaches in order to create materials that match their natural counterparts in the ability to respond to external stimuli – molecular or otherwise – with the amplification of specific monomer sequences. However, we cannot beat Boltzmann when we stick to the paradigm of strict self-assembly. All equilibrium mixtures must contain a significant fraction of undesired products. This problem will only become more acute when we go to polymers in which the entropic penalty to form a single sequence becomes truly enormous as the length and monomer choice increase. Therefore, true progress in this area should draw inspiration from Nature and combine reversible self-assembly with proofreading processes that derive their selectivity from the dissipation of energy. These irreversible processes,

when part of complex networks of reactions that also include, for example, self-replication and autocatalysis, cannot only give rise to unprecedented selectivity, but also to an almost limitless variety of emergent properties (see Chapter 2) [112]. It is with these prospects that we predict a bright future for the use of polymeric systems in DCC.

References

1 Flory, P.J. (1946) *Chem. Rev.*, **39**, 137.
2 Rowan, S.J., Cantrill, S.J., Cousins, G.R.L., Sanders, J.K.M., and Stoddart, J.F. (2002) *Angew. Chem. Int. Ed.*, **41**, 898.
3 Otsuka, H., Aotani, K., Higaki, Y., and Takahara, A. (2003) *J. Am. Chem. Soc.*, **125**, 4064.
4 Higaki, Y., Otsuka, H., and Takahara, A. (2004) *Macromolecules*, **37**, 1696.
5 Yamaguchi, G., Higaki, Y., Otsuka, H., and Takahara, A. (2005) *Macromolecules*, **38**, 6316.
6 Higaki, Y., Otsuka, H., and Takahara, A. (2006) *Macromolecules*, **39**, 2121.
7 Otsuka, H., Aotani, K., Higaki, Y., Amamoto, Y., and Takahara, A. (2007) *Macromolecules*, **40**, 1429.
8 Nakazawa, I., Suda, S., Masuda, M., Asai, M., and Shimizu, T. (2000) *Chem. Commun.*, 881.
9 Chow, C.F., Fujii, S., and Lehn, J.M. (2007) *Angew. Chem. Int. Ed.*, **46**, 5007.
10 Dobrawa, R. and Würthner, F. (2005) *J. Polym. Sci. A: Polym. Chem.*, **43**, 4981.
11 Gohy, J.F., Lohmeijer, B.G.G., and Schubert, U.S. (2003) *Chem. Eur. J.*, **9**, 3472.
12 Ogawa, K. and Kobuke, Y. (2000) *Angew. Chem. Int. Ed.*, **39**, 4070.
13 Chen, X., Dam, M.A., Ono, K., Mal, A., Shen, H., Nutt, S.R., Sheran, K., and Wudl, F. (2002) *Science*, **295**, 1698.
14 Murphy, E.B., Bolanos, E., Schaffner-Hamann, C., Wudl, F., Nutt, S.R., and Auad, M.L. (2008) *Macromolecules*, **41**, 5203.
15 Chen, X., Wudl, F., Mal, A.K., Shen, H., and Nutt, S.R. (2003) *Macromolecules*, **36**, 1802.
16 Kilcher, G., Wang, L., and Tirelli, N. (2008) *J. Polym. Sci. A: Polym. Chem.*, **46**, 2233.
17 ten Cate, A.T., Dankers, P.Y.W., Sijbesma, R.P., and Meijer, E.W. (2005) *J. Org. Chem.*, **70**, 5799.
18 Xu, J., Liu, X., Lin, T., Huang, J., and He, C. (2005) *Macromolecules*, **38**, 3554.
19 Giuseppone, N., Schmitt, J.L., and Lehn, J.M. (2004) *Angew. Chem. Int. Ed.*, **43**, 4902.
20 Hickman, D.T., Sreenivasachary, N., and Lehn, J.M. (2008) *Helv. Chim. Acta*, **91**, 1.
21 Kolomiets, E. and Lehn, J.M. (2005) *Chem. Commun.*, 1519.
22 Ono, T., Fujii, S., Nobori, T., and Lehn, J.M. (2007) *Chem. Commun.*, 4360.
23 Ono, T., Fujii, S., Nobori, T., and Lehn, J.M. (2007) *Chem. Commun.*, 46.
24 Ono, T., Nobori, T., and Lehn, J.M. (2005) *Chem. Commun.*, 1522.
25 Ruff, Y. and Lehn, J.M. (2008) *Angew. Chem. Int. Ed.*, **47**, 3556.
26 Schmitt, J.L. and Lehn, J.M. (2003) *Helv. Chim. Acta*, **86**, 3417.
27 Skene, W.G. and Lehn, J.M. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 8270.
28 Bouteiller, L. (2007) *Adv. Polym. Sci.*, **207**, 79.
29 Brunsveld, L., Folmer, B.J.B., Meijer, E.W., and Sijbesma, R.P. (2001) *Chem. Rev.*, **101**, 4071.
30 Wilson, A.J. (2007) *Soft Matter*, **3**, 409.
31 Ohga, K., Takashima, Y., Takahashi, H., Kawaguchi, Y., Yamaguchi, H., and Harada, A. (2005) *Macromolecules*, **38**, 5897.
32 Chow, C.F., Fujii, S., and Lehn, J.M. (2007) *Chem. Commun.*, 4363.
33 Giuseppone, N., Fuks, G., and Lehn, J.M. (2006) *Chem. Eur. J.*, **12**, 1723.
34 Giuseppone, N. and Lehn, J.M. (2004) *J. Am. Chem. Soc.*, **126**, 11448.
35 Ruff, Y. and Lehn, J.M. (2008) *Biopolymers*, **89**, 486.

36 Kudo, H., Shigematsu, K., Mitani, K., Nishikubo, T., Kasuga, N.C., Uekusa, H., and Ohashi, Y. (2008) *Macromolecules*, **41**, 2030.

37 Kudo, H., Makino, S., and Nishikubo, T. (2007) *J. Polym. Sci. A: Polym. Chem.*, **45**, 680.

38 Kamplain, J.W. and Bielawski, C.W. (2006) *Chem. Commun.*, 1727.

39 Burnworth, M., Knapton, D., Rowan, S.J., and Weder, C. (2007) *J. Inorg. Organomet. Polym. Mater.*, **17**, 91.

40 Stadler, R. and de Lucca Freitas, L. (1986) *Colloid Polym. Sci.*, **264**, 773.

41 Sijbesma, R.P., Beijer, F.H., Brunsveld, L., Folmer, B.J.B., Hirschberg, J.H.K.K., Lange, R.F.M., Lowe, J.K.L., and Meijer, E.W. (1997) *Science*, **278**, 1601.

42 Boileau, S., Bouteiller, L., Lauprêtre, F., and Lortie, F. (2000) *New J. Chem.*, **24**, 845.

43 Berl, V., Schmutz, M., Krische, M.J., Khoury, R.G., and Lehn, J.M. (2002) *Chem. Eur. J.*, **8**, 1227.

44 Yang, X.W., Hua, F.J., Yamato, K., Ruckenstein, E., Gong, B., Kim, W., and Ryu, C.Y. (2004) *Angew. Chem. Int. Ed.*, **43**, 6471.

45 Case, M.A. and McLendon, G.L. (2000) *J. Am. Chem. Soc.*, **122**, 8089.

46 Boydston, A.J., Williams, K.A., and Bielawski, C.W. (2005) *J. Am. Chem. Soc.*, **127**, 12496.

47 Williams, K.A., Boydston, A.J., and Bielawski, C.W. (2007) *Chem. Soc. Rev.*, **36**, 729.

48 Corbett, P.T., Otto, S. and Sanders, J.K.M. (2004) *Org. Lett.*, **6**, 1825.

49 Lehn, J.M. (2005) *Prog. Polym. Sci.*, **30**, 814.

50 Bell, C.M., Kissounko, D.A., Gellman, S.H., and Stahl, S.S. (2007) *Angew. Chem. Int. Ed.*, **46**, 761.

51 Hoerter, J.M., Otte, K.M., Gellman, S.H., Cui, Q., and Stahl, S.S. (2008) *J. Am. Chem. Soc.*, **130**, 647.

52 Cacciapaglia, R., Stefano, S.D., and Mandolini, L. (2005) *J. Am. Chem. Soc.*, **127**, 13666.

53 Kleiman, M. (1949) US Patent 247484.

54 Kolano, C., Helbing, J., Bucher, G., Sander, W., and Hamm, P. (2007) *J. Phys. Chem. B*, **111**, 11297.

55 Volk, M. (2001) *Eur. J. Org. Chem.*, 2605.

56 Volk, M., Kholodenko, Y., Lu, H.S.M., Gooding, E.A., DeGrado, W.F., and Hochstrasser, R.M. (1997) *J. Phys. Chem. B*, **101**, 9607.

57 Fouquey, C., Lehn, J.M., and Levelut, A.M. (1990) *Adv. Mater.*, **2**, 254.

58 Gulik-Krczywicki, T., Fouquey, C., and Lehn, J.M. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 163.

59 Bergman, S.D. and Wudl, F. (2008) *J. Mater. Chem.*, **18**, 41.

60 Wool, R.P. (2008) *Soft Matter*, **4**, 400.

61 Cordier, P., Tournilhac, F., Soulié-Ziakovic, C., and Leibler, L. (2008) *Nature*, **451**, 977.

62 Williams, K.A., Boydston, A.J., and Bielawski, C.W. (2007) *J. R. Soc. Interface*, **4**, 359.

63 Blaiszik, B.J., Sottos, N.R., and White, S.R. (2008) *Compos. Sci. Technol.*, **68**, 978.

64 White, S.R., Sottos, N.R., Geubelle, P.H., Moore, J.S., Kessler, M.R., Sriram, S.R., Brown, E.N., and Viswanathan, S. (2001) *Nature*, **409**, 794.

65 Norsten, T.B., Jeoung, E., Thibault, R.J., and Rotello, V.M. (2003) *Langmuir*, **19**, 7089.

66 South, C.R., Burd, C., and Weck, M. (2007) *Acc. Chem. Res.*, **40**, 63.

67 Gerhardt, W., Crne, M., and Weck, M. (2004) *Chem. Eur. J.*, **10**, 6212.

68 Pollino, J.M. and Weck, M. (2005) *Chem. Soc. Rev.*, **34**, 193.

69 Weck, M. (2007) *Polym. Int.*, **56**, 453.

70 Pollino, J.M., Stubbs, L.P., and Weck, M. (2004) *J. Am. Chem. Soc.*, **126**, 563.

71 Liu, S.M., Ruspic, C., Mukhopadhyay, P., Chakrabarti, S., Zavalij, P.Y., and Isaacs, L. (2005) *J. Am. Chem. Soc.*, **127**, 15959.

72 Mukhopadhyay, P., Wu, A.X., and Isaacs, L. (2004) *J. Org. Chem.*, **69**, 6157.

73 Mukhopadhyay, P., Zavalij, P.Y., and Isaacs, L. (2006) *J. Am. Chem. Soc.*, **128**, 14093.

74 Wu, A.X., Chakraborty, A., Fettinger, J.C., Flowers, R.A., and Isaacs, L. (2002) *Angew. Chem. Int. Ed.*, **41**, 4028.

75 Wu, A.X. and Isaacs, L. (2003) *J. Am. Chem. Soc.*, **125**, 4831.

76 Ikkala, O. and ten Brinke, G. (2004) *Chem. Commun.*, 2131.

77 Ikkala, O. and ten Brinke, G. (2002) *Science*, **295**, 2407.

78 Burd, C. and Weck, M. (2005) *Macromolecules*, **38**, 7225.

79 Burd, C. and Weck, M. (2008) *J. Polym. Sci. A: Polym. Chem.*, **46**, 1936.

80 Shenhar, R., Xu, H., Frankamp, B.L., Mates, T.E., Sanyal, A., Uzun, O., and Rotello, V.M. (2005) *J. Am. Chem. Soc.*, **127**, 16318.

81 Hofmeier, H. and Schubert, U.S. (2005) *Chem. Commun.*, 2423.

82 Xu, H., Hong, R., Lu, T.X., Uzun, O., and Rotello, V.M. (2006) *J. Am. Chem. Soc.*, **128**, 3162.

83 de Loos, M., Friggeri, A., van Esch, J., Kellogg, R.M., and Feringa, B.L. (2005) *Org. Biomol. Chem.*, **3**, 1631.

84 van Esch, J., Schoonbeek, F., de Loos, M., Kooijman, H., Spek, A.L., Kellogg, R.M., and Feringa, B.L. (1999) *Chem. Eur. J.*, **5**, 937.

85 Obert, E., Bellot, M., Bouteiller, L., Andrioletti, F., Lehen-Ferrenbach, C., and Boue, F. (2007) *J. Am. Chem. Soc.*, **129**, 15601.

86 Brizard, A., Stuart, M., van Bommel, K., Friggeri, A., de Jong, M., and van Esch, J. (2008) *Angew. Chem. Int. Ed.*, **47**, 2063.

87 Mohmeyer, N. and Schmidt, H.W. (2005) *Chem. Eur. J.*, **11**, 863.

88 Mohmeyer, N. and Schmidt, H.W. (2007) *Chem. Eur. J.*, **13**, 4499.

89 Gronwald, O., Snip, E., and Shinkai, S. (2002) *Curr. Opin. Colloid Interface Sci.*, **7**, 148.

90 Koevoets, R.A., Versteegen, R.M., Kooijman, H., Spek, A.L., Sijbesma, R.P., and Meijer, E.W. (2005) *J. Am. Chem. Soc.*, **127**, 2999.

91 Wisse, E., Govaert, L.E., Meijer, H.E.H., and Meijer, E.W. (2006) *Macromolecules*, **39**, 7425.

92 Wisse, E., Spiering, A.J.H., van Leeuwen, E.N.M., Renken, R.A.E., Dankers, P.Y.W., Brouwer, L.A., van Luyn, M.J.A., Harmsen, M.C., Sommerdijk, N.A.J.M., and Meijer, E.W. (2006) *Biomacromolecules*, **7**, 3385.

93 Botterhuis, N.E., Karthikeyan, S., Spiering, A.J.H., and Sijbesma, R.P. Macromolecules, Article ASAP DOI: 10.1021/ma902585w.

94 Pollino, J.M., Nair, K.P., Stubbs, L.P., Adams, J., and Weck, M. (2004) *Tetrahedron*, **60**, 7205.

95 Mullis, K.B. (1994) *Angew. Chem. Int. Ed. Engl.*, **33**, 1209.

96 Chandra, M.S. and Radhakrishnan, T.P. (2006) *Chem. Eur. J.*, **12**, 2982.

97 Hamada, K., Serizawa, T., and Akashi, M. (2005) *Macromolecules*, **38**, 6759.

98 Janssen, P.G.A., Vandenbergh, J., van Dongen, J.L.J., Meijer, E.W., and Schenning, A.P.H.J. (2007) *J. Am. Chem. Soc.*, **129**, 6078.

99 South, C.R. and Weck, M. (2007) *Macromolecules*, **40**, 1386.

100 Moore, J.S. and Zimmerman, N.W. (2000) *Org. Lett.*, **2**, 915.

101 Cramer, F. and Freist, W. (1987) *Acc. Chem. Res.*, **20**, 79.

102 Liu, J.W., Wernette, D.P., and Lu, Y. (2005) *Angew. Chem. Int. Ed.*, **44**, 7290.

103 Lu, Y. and Liu, J.W. (2007) *Acc. Chem. Res.*, **40**, 315.

104 Ge, Y. and Turner, A.P.F. (2008) *Trends Biotechnol.*, **26**, 218.

105 Kandimalla, V.B. and Ju, H. (2004) *Anal. Bioanal. Chem.*, **380**, 587.

106 Marty, J.D., and Mauzac, M. (2005) *Adv. Polym. Sci.*, **172**, 1.

107 Ye, L. and Mosbach, K. (2008) *Chem. Mater.*, **20**, 859.

108 Zimmerman, S.C., and Lemcoff, N.G. (2004) *Chem. Commun.*, 5.

109 Hiratani, H., Alvarez-Lorenzo, C., Chuang, J., Guney, O.,, Grosberg, Yu.A., and Tanaka T. (2001) *Langmuir*, **17**, 4431.

110 Wulff, G. and Schulze, I. (1978) *Angew. Chem. Int. Ed. Engl.*, **17**, 537.

111 Li, X. and Lynn, D.G. (2002) *Angew. Chem. Int. Ed.*, **41**, 4567.

112 Corbett, P.T., Sanders, J.K.M., and Otto, S. (2007) *Angew. Chem. Int. Ed.*, **46**, 8858.

7
Analytical Applications of Dynamic Combinatorial Chemistry

Kay Severin

7.1
Introduction

The basic concepts of selection experiments with dynamic combinatorial libraries (DCLs) were articulated more than 10 years ago (see Chapter 1). Since then, a number of applications have emerged. This includes the discovery new enzyme inhibitors, receptors, and catalysts, as well as the synthesis of novel materials such as responsive gels and polymers (see Chapters 2–5). A recent addition to the list of applications is the utilization of dynamic combinatorial chemistry (DCC) for analytical purposes. This chapter summarizes the main ideas and results in this area.

The concentrations of the different members of a DCL depend on the physical and chemical environment of the respective system (pH, solvent, concentration of target molecules, etc.). The library composition is therefore a characteristic feature of the respective environment. If the DCL composition can be transduced into a signal output, it is possible to use the DCL as a sensor. Typically, DCLs are analyzed by nuclear magnetic resonance spectroscopy or high-performance liquid chromatography. For sensing purposes, however, faster and cheaper analysis methods such as fluorescence or ultraviolet-visible (UV-Vis) spectroscopy are preferred. These techniques can be used if the DCL is composed of compounds with different color or fluorescence properties (Figure 7.1).

For a DCL sensor of this kind, the information about the analyte is distributed over the entire spectrum. The spectrum therefore represents a "fingerprint" of the analyte. To correlate the spectral changes with the analyte properties of interest (identity, quantity, purity), it is advantageous to use multivariate analyses techniques. In this regard, a DCL sensor is related to sensor arrays [1, 2]. However, contrary to sensor arrays with independent sensor units, a DCL sensor is comprised of compounds that are connected by exchange reactions. Furthermore, the various sensors of an array have to be analyzed separately, whereas a single UV-Vis or fluorescence measurement is sufficient for a DCL sensor.

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

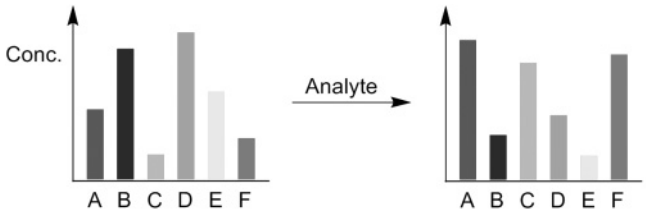
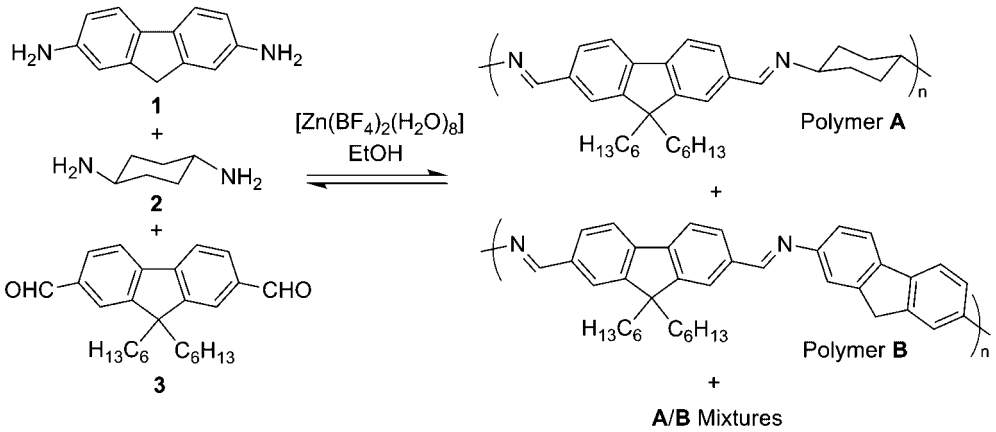


Figure 7.1 Basic principle of a DCL sensor: a dynamic mixture of colored (or fluorescent) compounds A–F undergoes an analyte-induced adaptation. The change in color (or fluorescence) can be used to obtain information about the identity, quantity, or purity of the analyte.

7.2 Fluorescent Sensors

Over recent years, the group of Lehn has studied constitutionally dynamic polymers ("dynamers"), which result from the connection of monomers via reversible covalent bonds [3]. In this context, they have investigated the condensation reaction of a 1:1:1 mixture of 2,7-diaminofluorene (**1**), *trans*-1,4-diaminocyclohexane (**2**), and 2,7-fluorene-bis-carboxaldehyde (**3**) in the presence of variable amounts of $[Zn(BF_4)_2(H_2O)_8]$ (Scheme 7.1) [4, 5].



Scheme 7.1 Condensation of the diamines **1** and **2** with the dialdehyde **3** results in the formation of a dynamic mixture of polymers. The addition of Zn^{2+} favors the formation of polymer **B**. The rearrangement is accompanied by a change of color and fluorescence.

The dominant product in the absence of zinc was polymer **A** containing the aliphatic monomer **2**. The addition of increasing amounts of zinc shifted the equilibrium in favor of polymers containing the aromatic monomer **1**. Polymer **B** was found to be the dominant species when the system was equilibrated in the presence of two equivalents of Zn^{2+}. The difference was explained by the prefer-

ential complexation of Zn^{2+} to the more nucleophilic diamine **2**. Polymers **A** and **B** show different absorption spectra. The zinc-induced change in polymer composition therefore resulted in a change in color. Similarly, the fluorescence emission maximum was shifted from 370 to 493 nm with a concomitant increase in intensity. The dynamic polymer can therefore be regarded as a system that is able to sense Zn^{2+} due to an analyte-induced constitutional rearrangement.

More recently, the Lehn group has reported the synthesis of dynamic polymers with oligosaccharides grafted to the main-chain [6]. These "glycodynamers" were found to be strongly fluorescent. The emission properties were affected by exchange processes. Since oligosaccharides are involved in numerous molecular recognition events, it is conceivable that glycodynamers may find applications in biosensing.

So far, there are only few reports about fluorescent oligomers that change their constitution upon addition of an analyte. In contrast, there are many examples of fluorescent molecules that undergo a conformational change upon addition of certain analytes (e.g., "molecular beacons" [7]). Although conformationally dynamic receptors are sometimes discussed in the context of DCC [8], these systems will not be described in more detail in the present chapter.

7.3 Colorimetric Sensors

Artificial receptors can be converted into sensors by covalent attachment of a signaling unit such as a fluorescent dye. An interesting alternative are so-called indicator displacement assays (IDAs) [9]. These assays are based on receptors that are bound to dyes (or fluorescent ligands) via noncovalent interactions. Upon addition of an analyte, the dyes are displaced, which results in a change of their optical properties. These changes can be used to identify and/or quantify the analyte.

The group of Anslyn has investigated IDAs based on the synthetic receptors **4** [10] and **5** [11] (Figure 7.2). These cationic hosts have a high affinity for oxoanions such as gallic acid, tartrate, or malate. It was found that analytical power of such IDAs can be increased if several receptors and dyes are used in parallel. This was demonstrates by making a malate and tartrate sensor composed of receptor **4**, receptor **5**, and the dyes bromopyrogallol red and pyrocatechol violet [12]. Both dyes can bind to both receptors and, consequently, four different receptor–dye complexes can be formed. Although this was not stressed in the original publication, the system can be regarded as a small DCL of receptor–dye aggregates.

When a solution containing variable amounts of malate and tartrate was added to the multicomponent sensor, a characteristic UV-Vis spectrum response was obtained. Since the different receptor–dye complexes and the free dyes all have different colors, the information about the analytes was dispersed over the entire spectrum. To analyze the spectral changes, a multilayer preceptron (MLP) artificial neural network (ANN) was employed (see Glossary in Box 7.1). To train the ANN, the UV-Vis absorption data of 45 calibration samples containing different amounts of malate and tartrate (0–1.2 mM) were used. For each sample, the absorbances at

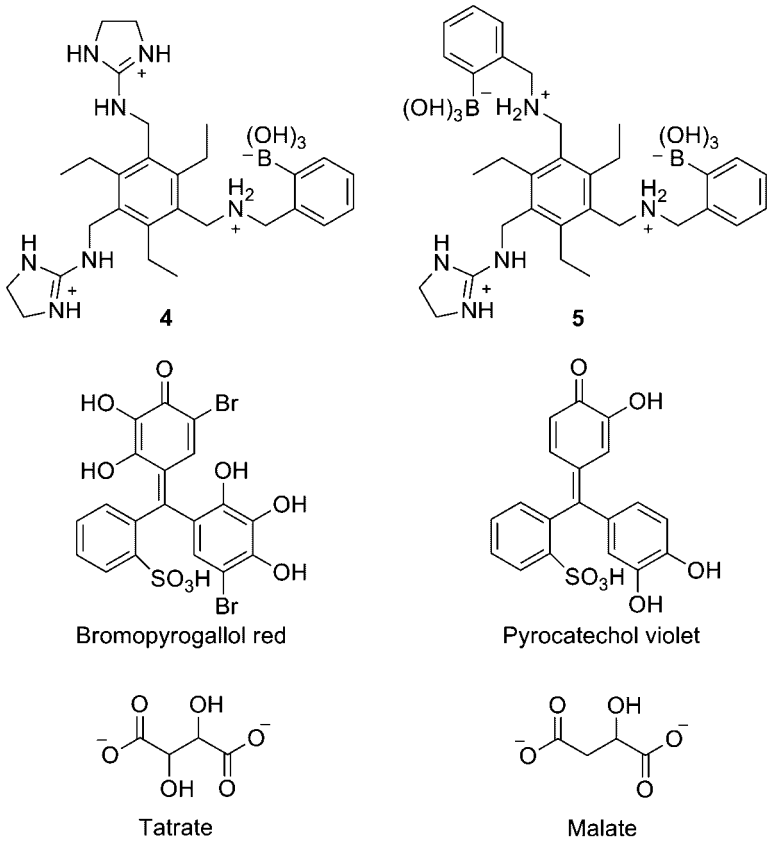


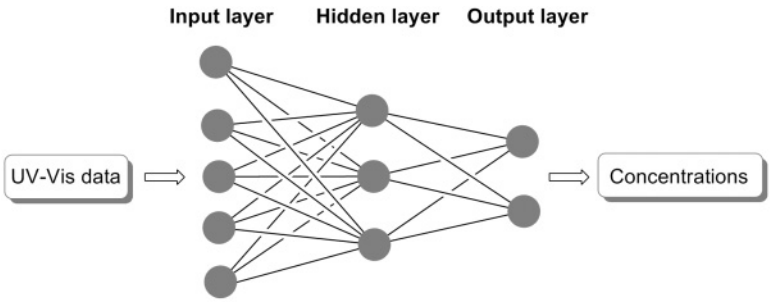
Figure 7.2 Dynamic mixture of the receptors **4**, **5**, and the dyes bromopyrogallol red and pyrocatechol violet can be used to determine simultaneously low millimolar concentration of malate and tartrate.

27 wavelengths between 375 and 675 nm were taken as the data input. The training consisted of correlating the input (UV-Vis data) with the output (concentration of the analytes malate and tartrate). The predictive power of the trained network was then evaluated with four validation samples. The absolute errors of the predictions were found to be between 1 and 6%. After further training with new calibration samples, the error was consistently lower than 2%. These results show that a dynamic mixture of receptor–dye complexes can be used as a powerful sensor.

In another study, the Anslyn group has employed receptor **6** in combination with xylenol orange to determine the concentration of Ca^{2+} and citrate in flavored vodkas (Figure 7.3) [16]. Receptor **6** can bind to the dye xylenol orange or to citrate. The dye and citrate, on the other hand, can both form complexes with Ca^{2+}. The addition of samples containing Ca^{2+} and citrate to a sensor comprised of receptor **6** and xylenol orange was therefore expected to result in a complex dynamic mixture of aggregates.

Box 7.1 Glossary: Multivariate Analysis.

Principal component analysis (PCA)	A PCA is a statistical tool that allows us to reduce a multidimensional data set to lower dimensions [13]. PCA is mathematically defined as an orthogonal linear transformation that converts the data to a new coordinate system such that the highest degree of variance of the data comes to lie on the first coordinate (called the first principal component) and other principal components follow in the order of decreasing variance. PCA is a nonsupervised method that does not use the information of what data belongs to which class (e.g., an analyte).
Linear discriminant analysis (LDA)	Similar to a PCA, a LDA is able to reduce the dimensionality of a dataset [13]. However, contrary to PCA, it is a supervised method that uses the information of which data point belongs to which class. The discriminants are linear combinations of the measured variables (e.g., sensor response). A discriminant function is found that maximizes the ratio of between-class variance to within-class variance.
Cross-validation	A cross-validation procedure partitions the dataset into training and validation data. The analysis is performed on the training data and the validation data is used to confirm the analysis. A leave-one-out cross-validation uses a single observation from the original sample as the validation data and the remaining observations as the training data. This is repeated until each observation in the sample is used once as the validation data.
Artificial neural networks (ANN)	ANNs are nonlinear computational tools suitable to model complex relationships between inputs and outputs or to find patterns in data sets [14]. The structure of ANNs is reminiscent of biological neural networks. Several simple nodes ("neurons") are connected to form a network of nodes. Algorithms define the strengths between the neurons. Multilayer preceptron (MLP) is one of the most popular and traditional models of ANNs [15]. Its structure consists of an input layer, one or more hidden layers, and an output layer. All layers contain a variable number of neurons. The network maps the input data (e.g., UV-Vis data) to a set of outputs (e.g., concentrations). A training data set can be used to optimize the strength of the connections. The trained network is then able to make predictions for a validation data set.



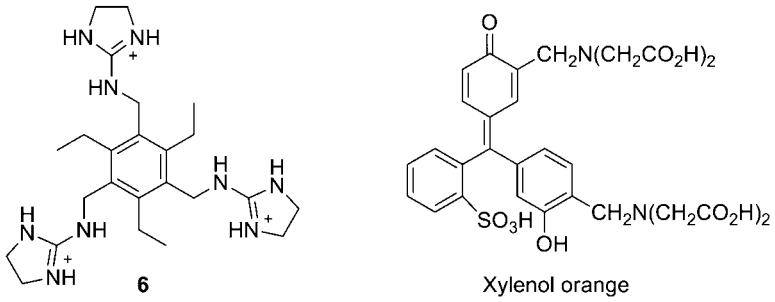
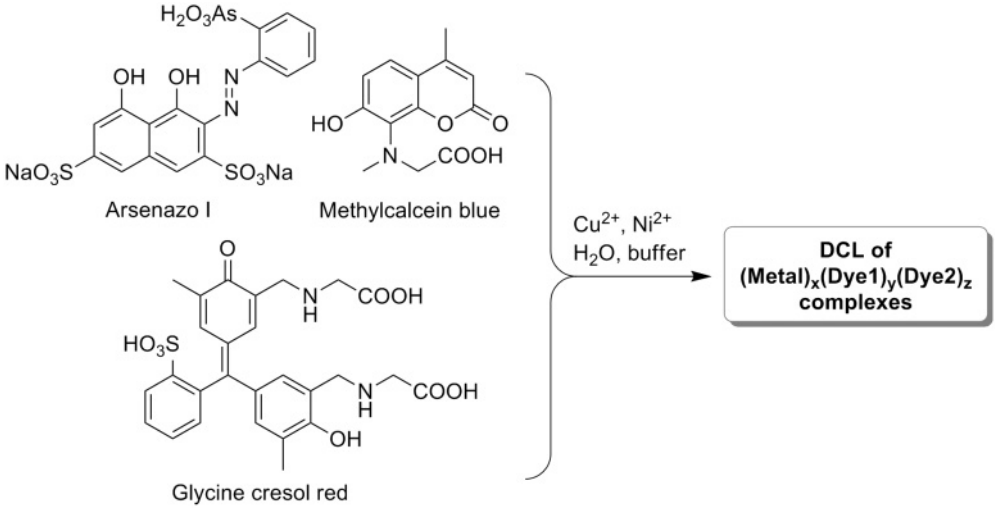


Figure 7.3 Mixture of receptor **6** and xylenol orange can be used to determine Ca^{2+} and citrate.

First, an ANN was trained with the UV-Vis data of 75 calibration samples containing 0–400 μM of Ca^{2+} and 0–800 μM of citrate. The sensor was always composed of a fixed amount of receptor **6** (240 μM) and xylenol orange (10 μM). The calibrated system was then used to predict the concentration of Ca^{2+} and citrate in five validation samples and in five flavored vodkas. Overall, the correlation between the predicted and the real values was acceptable (maximal differences of up to 33%).

A different way to construct a dynamic mixture of compounds with a characteristic color was described by our group [17]. The commercially available dyes methylcalcein blue, arsenazo I, and glycine cresol red were dissolved in aqueous buffer containing the transition metal salts $CuCl_2$ and $NiCl_2$. The resulting solution contained a mixture of $(metal)_x(dye1)_y(dye2)_z$ complexes of variable composition (Scheme 7.2). The coordination to Cu^{2+} or Ni^{2+} changes the UV-Vis absorption of the dyes. Each complex was therefore expected to display a distinct color.



Scheme 7.2 Generation of a DCL of metal–dye complexes by mixture of arsenazo I, methylcalcein blue, and glycine cresol red with $CuCl_2$ and $NiCl_2$ in buffered aqueous solution.

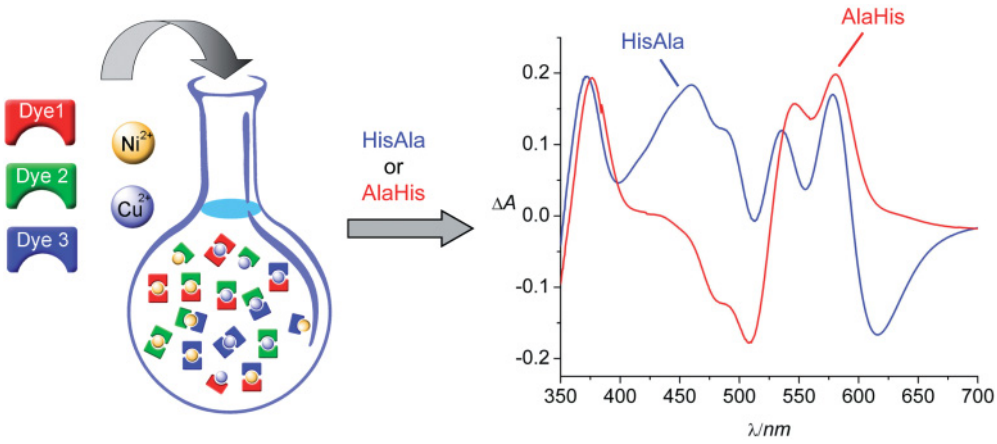


Figure 7.4 Changes in the UV-Vis spectrum upon addition of the dipeptides His–Ala or Ala–His to a DCL sensor composed of $(metal)_x(dye1)_y(dye2)_z$ complexes ([dipeptide] = 1.0 mM, $[CuCl_2] = [NiCl_2] = [dye1/2/3]$ = 75 µM, 35 mM CHES buffer, pH 8.4).

The dynamic mixture of metal–dye complexes was responsive to analytes, which are able to displace the dyes from the complexes. This was demonstrated with dipeptides. Dipeptides are known to form stable complexes with Cu^{2+} and Ni^{2+} ions. The addition of a dipeptide to the DCL sensor therefore resulted in a partial liberation of dyes, accompanied by a re-equilibration of the remaining metal–dye complexes. The results obtained for the sequence isomers His–Ala and Ala–His are shown in Figure 7.4.

The data depicted in Figure 7.4 are the UV-Vis difference spectra obtained from equilibrated solutions before and after addition of the respective dipeptide. The analytes His–Ala and Ala–His are clearly distinguishable. It is important to note that the UV-Vis spectra differ not only in terms of amplitude at a certain wavelength, but also in terms of the position of inflection points and maxima.

Peptides with His residues give rise to large changes in the UV-Vis spectrum because the side-chain of histidine has good metal-coordinating abilities. To test the scope of the DCL sensor with more challenging analytes, the structurally related dipeptides Gly–Ala, Val–Phe, Ala–Phe, Phe–Ala, and D-Phe–Ala were investigated. The UV-Vis difference spectra of these analytes were more similar to each other and chemometrics methods (for reviews, see [18]) were used to identify the analytes. For each peptide, 15 UV-Vis measurements were performed. The data of eight selected wavelengths were classified by a linear discriminant analysis (LDA) (see Glossary in Box 7.1). A 100% discrimination was achieved for a leave-one-out cross-validation, in which one measurement at a time was treated as an unknown and the rest of the data was used as the training set (see Glossary in Box 7.1). This is quite remarkable, given that closely related analytes such as the regioisomers Ala–Phe and Phe–Ala and the stereoisomers Phe–Ala and D-Phe–Ala were used, and that none of the dipeptides contained coordinating side-chains.

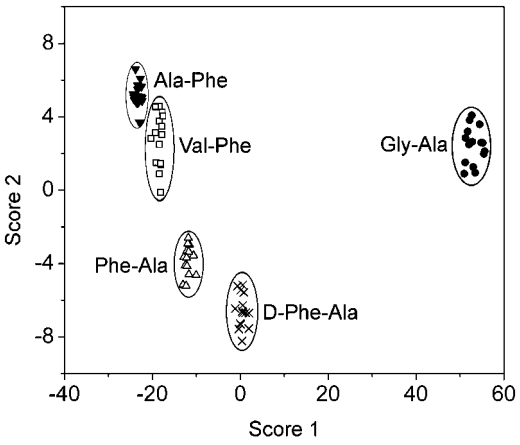


Figure 7.5 Two-dimensional LDA score plot for the dipeptide analytes Gly–Ala, Val–Phe, Ala–Phe, Phe–Ala, and D-Phe–Ala. The five different peptides can clearly be distinguished.

A graphic representation of the LDA analysis in the form of a score plot is given in Figure 7.5.

An advantage from an experimental point of view is the fact that the synthesis of the DCL sensor is very easy – it is obtained by simply dissolving $CuCl_2$ and $NiCl_2$ together with three commercially available dyes in a buffer. However, the approach has another advantage: its inherent flexibility. Since the sensor is assembled from five different building blocks, it is possible to rapidly optimize the system for a particular sensing problem by variation of the total amounts and the relative ratios in a combinatorial fashion. This was demonstrated for the analysis of sequence-isomeric tripeptides composed of one histidine and two glycines [19]. The building blocks for the DCL sensor were the same as described above. To optimize the sensor performance, 20 different DCL sensors were generated by variation of the total ($[M]_{total}$ = 80, 160, 240, and 320 μM) and relative metal concentration using constant dye concentrations ([dye1/2/3] = 75 μM). These 20 sensors were then tested for their ability to distinguish two different analyte pairs: His–Gly–Gly versus Gly–Gly–His and Gly–His–Gly versus Gly–Gly–His ([peptide] = 1.00 mM). To approximate the ability of the respective DCL sensor to discriminate between the two peptides, the area between the UV-Vis curves obtained for the two analytes was calculated. A graphic representation of the results of this screening is depicted in Figure 7.6.

For both analyte pairs, a high total metal concentration of 320 μM was found to give the best discrimination. However, interesting differences were observed for the best Cu^{2+} : Ni^{2+} ratio. For the analyte pair Gly–His–Gly and Gly–Gly–His, a mixture of 25% Cu^{2+} and 75% Ni^{2+} resulted in the largest difference in the UV-Vis spectra (Figure 7.6, left). For the analyte pair His–Gly–Gly and Gly–Gly–His, on the other hand, a sensor containing only Cu^{2+} gave the best results (Figure 7.6, right). The results show that the optimal sensor composition can

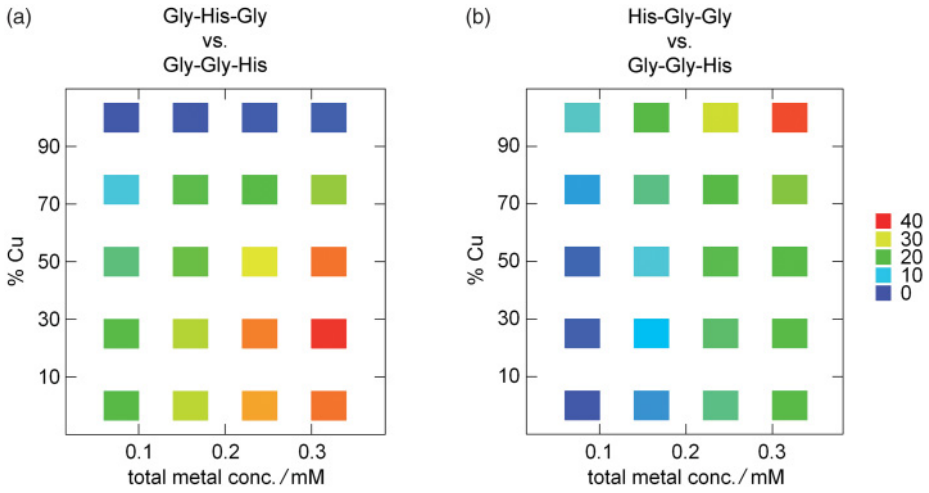


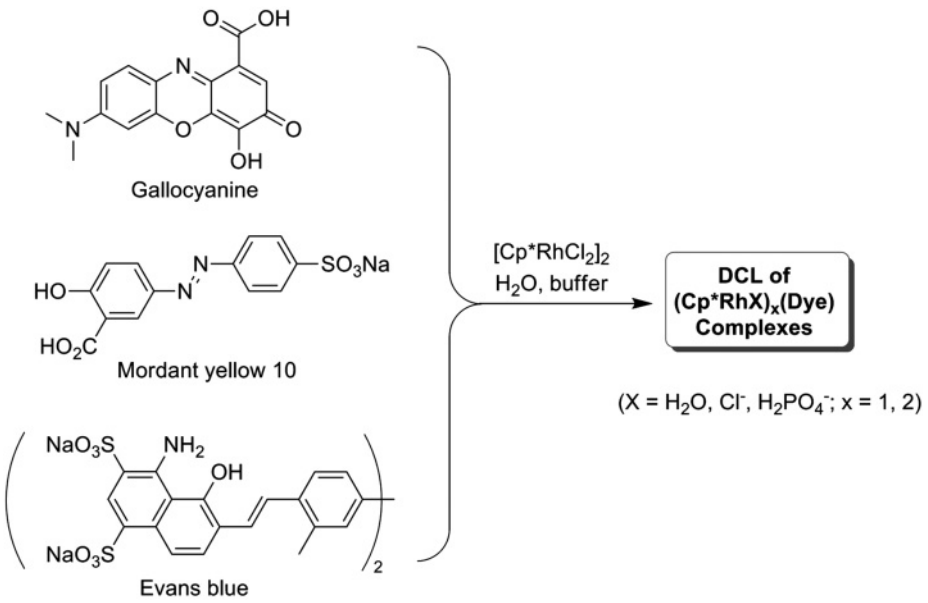
Figure 7.6 The ability of a DCL sensor to discriminate between the sequence isomers Gly–His–Gly and Gly–Gly–His (a) and His–Gly–Gly and Gly–Gly–His (b) as a function of the total metal concentration and the Cu^{2+}:Ni^{2+} ratio. The color indicates the differentiation, which was achieved for the respective sensor composition (red = good; blue = weak).

vary substantially, even for closely related analytes such as sequence-isomeric tripeptides.

In the small screening of 5×4 different sensors, only the metal concentrations were varied, but other parameters such as the relative dye concentrations were not changed. Therefore, the screening only resulted in a partial optimization. Nevertheless, the semioptimized sensors were found to show a remarkable analytical power. This was evidenced by the following experiment. Eight samples containing either Gly–His–Gly or Gly–Gly–His in concentrations of 0.25, 0.50, 0.75, or 1.00 mM were analyzed with the DCL sensor. After processing the UV-Vis data with a LDA, it was possible to identify the respective tripeptide and to determine its concentration.

More recently, a related approach was used to construct a sensor for nucleotides [20]. Instead of $CuCl_2$ and $NiCl_2$, the Rh(III) complex $[Cp^*RhCl_2]_2$ was employed. When $[Cp^*RhCl_2]_2$ is dissolved in water, air-stable Cp*Rh-aqua complexes are formed. The three coordination sites opposite to the π-ligand display a high affinity for N/O-donor ligands [21]. To obtain a small DCL with the Cp*Rh complex, it was dissolved together with the dyes gallocyanine, mordant yellow 10, and Evans blue in phosphate buffer at pH 7.4 (Scheme 7.3). Gallocyanine and mordant yellow 10 form 1 : 1 complexes with the Cp*Rh fragment, whereas 1 : 1 and 2 : 1 complexes are obtained with Evans blue.

For all three dyes, the complexation to the metal resulted in pronounced changes in the UV-Vis, but these changes occurred in different regions of the spectrum. The addition of analytes with a high affinity to the Cp*Rh complex was therefore



Scheme 7.3 Generation of a small DCL of Cp*Rh–dye complexes by mixture of gallocyanine, mordant yellow 10, and Evans blue with $[Cp^*RhCl_2]_2$ in buffered aqueous solution.

expected to give characteristic UV-Vis spectra. This was tested with nucleotides that are known to bind to the Cp*Rh fragment in aqueous solution [22]. First, the response of a sensor comprised of $[Cp^*RhCl_2]_2$ (60 μM) and the three dyes ($[dye]_{total}$ = 100 μM) was tested for the nucleotides adenosine diphosphate (ADP), adenosine triphosphate (ATP), or guanosine triphosphate (GTP) ([nucleotide] = 0.5 mM). The three nucleotides indeed gave characteristic spectra, which differed in terms of amplitude and position of the signals (Figure 7.7).

Subsequently, the range of the analytes was expanded by adding cyclic adenosine monophosphate (cAMP), adenosine monophosphate (AMP), and the pyrophosphate anion (PP_i) to the DCL sensor. Six independent measurements were performed for each analyte ([nucleotide] = $[PP_i]$ = 1.0 mM). The intensities at five different wavelengths were chosen as the input variables for a LDA (see Glossary in Box 7.1). The resulting score plot is shown in Figure 7.7. The data appear in well-separated groups, which shows that the sensor can easily discriminate the six analytes.

The Cp*Rh-based sensor can also be used for quantitative analyses. This was demonstrated with the analytes ATP, cAMP, and PP_i. The hydrolysis of ATP to give cAMP and PP_i is catalyzed by adenylate cyclase (AC). AC is an important enzyme that is involved in many signaling pathways. Its activity is commonly measured by monitoring the conversion of [α-^{32}P]ATP to [α-^{32}P]cAMP using ion-exchange chromatography to separate the nucleotides. To show that the sensor can be used to measure simultaneously the concentrations of ATP and cAMP/PP_i,

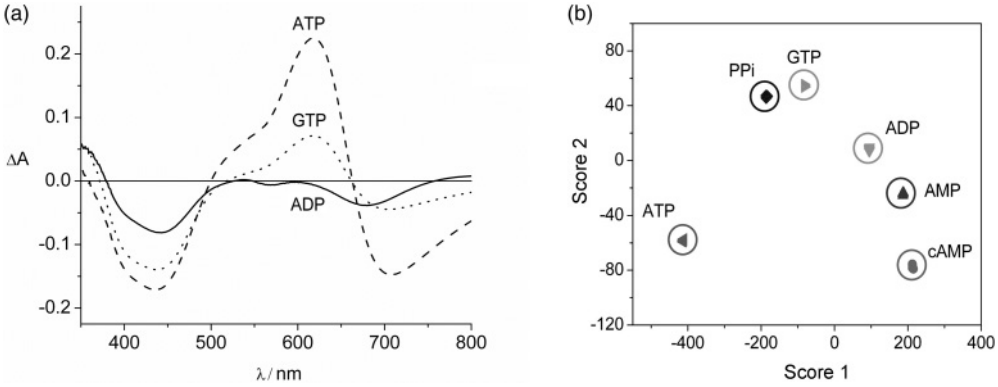


Figure 7.7 Identification of nucleotides and PP_i with a DCL sensor. (a) UV-Vis difference spectra obtained upon addition of ADP (solid line), GTP (dotted line), or ATP (dashed line) to a sensing ensemble composed of $[Cp^*RhCl_2]_2$ and the dyes gallocyanine, mordant yellow 10, and Evans blue. (b) Two-dimensional LDA score plot for the analysis of five different nucleotides and PP_i.

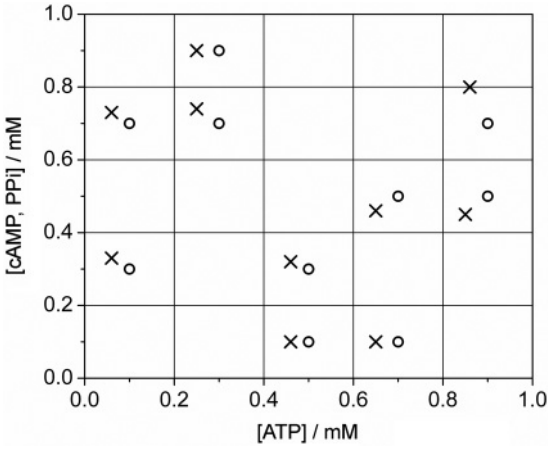


Figure 7.8 Determination of independent analyte concentrations (ATP and cAMP/PP_i) with a DCL sensor. The concentrations of the validation samples are shown as circles and the predictions of the ANN as crosses. The intersections of the grid lines represent the concentrations of the 36 calibration samples.

a multilayer ANN (see Glossary in Box 7.1) was trained with the UV-Vis data of 36 calibration samples containing variable amounts of ATP, cAMP, and PP_i ([cAMP] = [PP_i]). The predictive power of the trained ANN was evaluated with 10 validation samples with ATP and cAMP/PP_i concentrations between 0.1 and 0.9 mM. The actual and the predicted concentrations for the validation samples are shown in Figure 7.8. Overall, the predictions were quite good – the root-mean-square errors were only 45 µM for ATP and 42 µM for cAMP/PP_i.

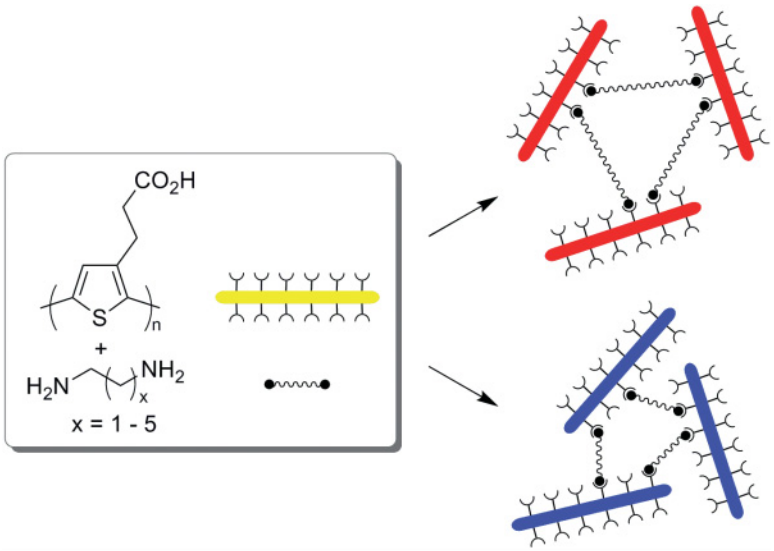


Figure 7.9 The interaction of α,ω-diamine analytes with a carboxylic acid functionalized poly(thiophene) results in an aggregation of the polymer. The chain length of the diamine affects the color of the aggregate.

The colorimetric sensors described above are based on dyes that interact with synthetic receptors or with metal complexes. The resulting dynamic mixtures of receptor–dye or metal–dye aggregates undergo competitive exchange reactions with analytes. This adaptation gives rise to characteristic changes in the UV-Vis spectrum. A different approach has been described by the group of Lavigne [23]. As the key building block of their sensor, they have employed a carboxylic acid functionalized poly(thiophene). α,ω-Diamines were chosen as analytes. The analytes bind to the polymers via electrostatic and/or hydrogen bonding interactions. As a result, the main-chains of the polymers are twisted and aggregation of the polymers occurs (Figure 7.9). Both effects influence the absorbance of the polymer in a specific way. The analytes can therefore be identified by a UV-Vis analysis. It should be noted that in this approach, the DCL of aggregates is only formed in the presence of analytes.

As a proof-of-concept, the Lavigne group has investigated the sensor response for simple α,ω-diamines with a chain lengths between two and six methylene groups. In addition, histamine was used as an analyte. The analysis was performed with the UV-Vis data obtained for a constant polymer concentration of 0.4 mM, and for five different analyte concentrations between 0.5 and 5.0 mM. The normalized absorption values at nine wavelengths were subjected to a LDA (see Glossary in Box 7.1). A "jack-knifed" classification matrix showed that the analyte can be identified with an accuracy of over 99% [23].

In a subsequent study, the Lavigne group has shown that a sensor of this kind can be used to quantify histamine in a fish sample [24]. Histamine is the most

prevalent biogenic amine found in tuna, and its amount is correlated with the "freshness" and the quality of the fish. Histamine was extracted from the tuna samples by extraction with trichloroacetic acid. Quantification of the amine was then possible using a "classical" analysis method – the ratiometric response of the polymer sensor at 420 and 530 nm. It is interesting to note that the sensitivity of the assay was better than the typical mammalian sense of smell.

The analyte-induced aggregation of functionalized poly(thiophene) is reminiscent of colorimetric sensors that are based on the aggregation or dissociation of gold nanoparticles. Gold nanoparticles display a red-shift of their plasmon band upon aggregation. This feature has been exploited extensively for (bio)analytical applications. A discussion of this work extends the scope of the present chapter and more details can be found in a number of review articles [25].

7.4 Molecular Timers

In the previous part of this chapter it was described that DCL sensors can be used to determine the identity and/or the quantity of analytes. A recent extension is the demonstration that DCL sensors can be used to obtain information about the history of analyte variations [26].

A photochemical sensor, whose optical status is determined by the time at which certain analytes are added, needs to operate under kinetic control. In this regard, Cp*Rh-based sensors are interesting because competitive dye–analyte reactions occur on the minute to hour timescale. Furthermore, it was found that the kinetics can be slowed down by using additives such as PP_i [26]. To demonstrate that DCLs of Cp*Rh(dye) complexes can be used to time molecular events, the following experiment was performed. The analytes ADP (250 μM) and ATP (250 μM) were added at times t_1 and t_2 to a freshly prepared mixture of $[Cp^*RhCl_2]_2$ and the dyes azophloxine, glycine cresol red, and naphthol blue black. The two analytes compete with the three dyes for the complexation to the metal. Since the free and the metal-bound dyes show different colors, a characteristic UV-Vis spectrum was obtained. A UV-Vis analysis at time t_3 followed by data processing with a multilayer ANN (see Glossary in Box 7.1) then allowed determining the addition times t_1 and t_2 (Figure 7.10).

For the training of the ANN, 25 calibration samples with variable ADP and ATP addition times were analyzed by measuring the absorption at five different wavelengths after t_3 = 15 min. The predictive power of the trained ANN was evaluated with 12 validation samples with randomly chosen addition times between 0 and 10 min for ADP and ATP. Analogous to the calibration samples, the validation samples were analyzed by UV-Vis spectroscopy after 15 min (five measurements each). The predictions of the ANN in comparison with the real addition times of ADP and ATP are shown in Figure 7.10. On average, the deviation of the prediction for the addition time of ADP was 36 s and for the addition time of ATP 34 s. This correlation is quite remarkable, given that the timer has to distinguish submillimolar concentrations of two structurally very similar analytes.

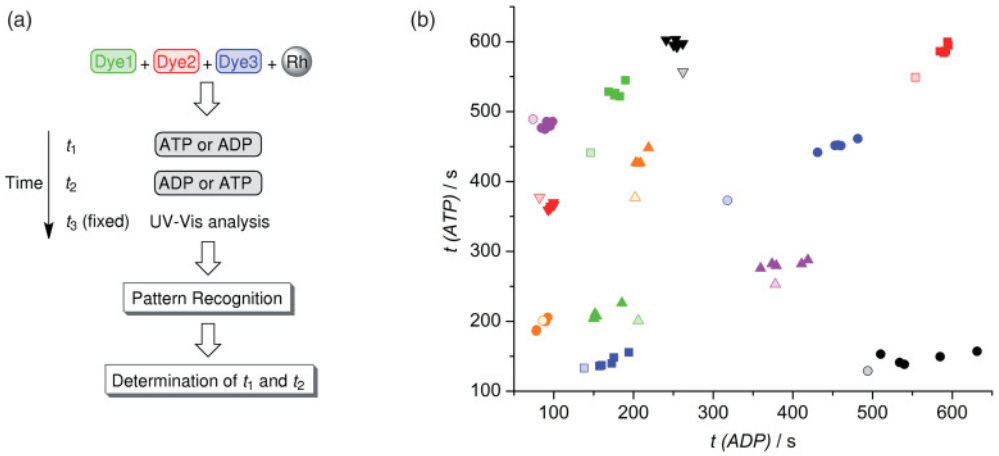


Figure 7.10 (a) UV-Vis analysis after time t_3 in combination with a multivariate analysis can be use to determine the times t_1 and t_2, when ADP and ATP were added to a solution containing a Cp*Rh(III) complex and three dyes. (b) ATP and ADP addition times determined by the molecular timer in comparison with the real addition times for 12 validation samples. The predictions are shown as filled symbols (five measurements for each sample) and the real addition times are indicated by empty symbols.

7.5 Conclusions

The examples described above are evidence that DCLs can be used as powerful analytical tools. Biologically interesting analytes such as peptides, nucleotides, or amines have been identified and/or quantified with sensors, which are cheap and easy to make. Furthermore, it has been shown that DCLs can be used to time molecular events.

For DCL sensors, the information about the analyte is represented by the library composition. Readout by fluorescence or UV-Vis spectroscopy is possible if the library is composed of compounds with a characteristic color or fluorescence. The information is then distributed across the entire spectral range of the library members. This distinguishes a DCL sensor from a classical chemosensor, for which the analysis is generally performed at either a single wavelength or by comparison of two wavelengths ("ratiometric sensor"). To process the DCL sensor data it is advantageous to use multivariate analyses methods. In this regard, DCL sensors are related to sensor arrays such as "electronic noses" or "electronic tongues."

DCLs of colored or fluorescent compounds can be used for any analyte that results in a detectable re-equilibration of the library. It is thus sufficient if the analyte displays an unspecific interaction with some library members. This is in contrast to DCL selection experiments aimed to identify a high-affinity receptor. Here, a specific and strong interaction with a particular library member is of advantage, because it results in higher amplification factors.

From an experimental point of view, DCL sensors have two key advantages. (i) The generation of a DCL sensor is very simple – all that is required is to mix the respective building blocks. This is exemplified by the colorimetric DCL sensor for dipeptides, which was obtained by mixing $CuCl_2$, $NiCl_2$, and three commercially available dyes. (ii) DCL sensors can easily be modified – and thus optimized – by variation of the nature, amounts, and relative ratios of the constituent building blocks. This flexibility is not found for classical chemosensors, which are based on receptors with covalently attached signaling units. Here, structural modifications may require substantial synthetic efforts.

The utilization of DCLs for analytical purposes is a recent addition to the constantly growing list of applications of DCC. So far, there are few publications in this area, most of which are summarized in this chapter. However, in view of the advantages outline above, it is likely that interesting new developments and applications will be reported in the near future.

References

1 Wright, A.T. and Anslyn, E.V. (2006) *Chem. Soc. Rev.*, **35**, 14–28.

2 Albert, K.L., Lewis, N.S., Schauer, C.L., Sotzing, G.A., Stizel, S.E., Vaid, T.P., and Walt, D.R. (2000) *Chem. Rev.*, **100**, 2595–2626.

3 Lehn, J.M. (2005) *Prog. Polym. Sci.*, **30**, 814–831.

4 Guiseppone, N. and Lehn, J.M. (2004) *J. Am. Chem. Soc.*, **126**, 11448–11449.

5 Guiseppone, N., Fuks, G., and Lehn, J.M. (2006) *Chem. Eur. J.*, **12**, 1723–1735.

6 Ruff, Y. and Lehn, J.M. (2008) *Angew. Chem. Int. Ed.*, **47**, 3556–3559.

7 Tan, W., Wang, K., and Drake, T.J. (2004) *Curr. Opin. Chem. Biol.*, **8**, 547–553.

8 Eliseev, A.V. and Nelen, M.I. (1997) *J. Am. Chem. Soc.*, **119**, 1147–1148.

9 (a) Nguyen, B.T. and Anslyn, E.V. (2006) *Coord. Chem. Rev.*, **250**, 3118–3127. (b) Fabbrizzi, L., Licchelli, M., and Taglietti, A. (2003) *Dalton Trans.*, 3471–3479.

10 Lavigne, J.J. and Anslyn, E.V. (1999) *Angew. Chem. Int. Ed.*, **38**, 3666–3669.

11 Wiskur, S.L. and Anslyn, E.V. (2001) *J. Am. Chem. Soc.*, **123**, 10109–10110.

12 Wiskur, S.L., Floriano, P.N., Anslyn, E.V., and McDevitt, J.T. (2003) *Angew. Chem. Int. Ed.*, **42**, 2070–2972.

13 Jurs, P.C., Bakken, G.A., and McClelland, H.E. (2000) *Chem. Rev.*, **100**, 2649–2678.

14 Marini, F., Bucci, R., Magrì, A.L., and Magrì, A.D. (2008) *Microchem. J.*, **88**, 178–185.

15 Cherkassky, V. and Mullier, F. (1998) *Learning from Data: Concepts, Theory, and Method*, John Wiley & Sons, Inc., New York.

16 McCleskey, S.C., Floriano, P.N., Wiskur, S.L., Anslyn, E.V., and McDevitt, J.T. (2003) *Tetrahedron*, **59**, 10089–10092.

17 Buryak, A. and Severin, K. (2005) *Angew. Chem. Int. Ed.*, **44**, 7935–7938.

18 (a) Lavine, B. and Workman, J. (2008) *Anal. Chem.*, **80**, 4519–4531. (b) Lavine, B. and Workman, J. (2006) *Anal. Chem.*, **78**, 4137–4145. (c) Geladi, P., Sethson, B., Nyström, J., Lillhonga, T., Lestander, T., and Burger, J. (2004) *Spectrochim. Acta*, **B 59**, 1347–1357.

19 Buryak, A. and Severin, K. (2006) *J. Comb. Chem.*, **8**, 540–543.

20 Buryak, A., Pozdnoukhov, A., and Severin, K. (2007) *Chem. Commun.*, 2366–2368.

21 (a) Zaubitzer, F., Buryak, A., and Severin, K. (2006) *Chem. Eur. J.*, **12**, 3928–3934. (b) Buryak, A. and Severin, K. (2005) *J. Am. Chem. Soc.*, **127**, 3700–3701. (c) Buryak, A. and Severin, K. (2004) *Angew. Chem. Int. Ed.*, **43**, 4771–4774.

22 Fish, R.H. (1999) *Coord. Chem. Rev.*, **185-186**, 569–584.
23 Nelson, T.L., O'Sullivan, C., Greene, N.T., Maynor, M.S., and Lavigne, J.J. (2006) *J. Am. Chem. Soc.*, **128**, 5640–5641.
24 Maynor, M.S., Nelson, T.L., O'Sullivan, C., and Lavigne, J.J. (2007) *Org. Lett.*, **9**, 3217–3220.
25 (a) Baron, R., Willner, B., and Willner, I. (2007) *Chem. Commun.*, 323–332. (b) Lu, Y. and Liu, J. (2006) *Curr. Opin. Chem. Biol.*, **17**, 580–588. (c) Rosi, N.L. and Mirkin, C.A. (2005) *Chem. Rev.*, **105**, 1547–1562.
26 Buryak, A., Zaubitzer, F., Pozdnoukhov, A., and Severin, K. (2008) *J. Am. Chem. Soc.*, **130**, 11260–11261.

8
Trends and Perspectives

Sijbren Otto and Joost N.H. Reek

8.1
Introduction

Interest in dynamic combinatorial chemistry (DCC) has increased rapidly since the first publications on the subject in the mid-1990s. Figure 8.1 shows that the annual number of publications involving DCC has risen steadily. Among the various available reversible covalent reactions, disulfide, imine, and hydrazone chemistries are the most popular (Figure 8.2). However, other reversible covalent reactions are being developed and there is definitely a need to further expand this repertoire.

As the available reversible chemistries diversify, exciting opportunities arise by combining different reversible reactions in the same system. The first such study was reported by Eliseev and Lehn, who combined hydrazone chemistry with metal–ligand coordination involving a labile Co(II) center that could be oxidized to form a stable (kinetically inert) Co(III) species [1]. Nitschke *et al.* extensively explored systems that combine imine and metal–ligand interactions [2]. Also a combination of imine, metal–ligand, and disulfide chemistry has recently been reported [3] as well as a combination of imine, metal–ligand, and boronic acid chemistry [4]. Combinations of different reversible covalent chemistries have also been reported, including imine and nitrone chemistry [5], communicating disulfide and thioester chemistry [6], and orthogonal disulfide and hydrazone chemistry [7, 8]. In the latter case, both exchange chemistries operate in different pH regions, so it should be possible to evolve such systems away from the hypothetical global thermodynamic minimum by alternating between the two exchange chemistries.

8.2
Dynamic Combinatorial Libraries as Molecular Networks

Dynamic combinatorial chemistry has grown out of an unmet demand for efficient methods for developing new synthetic receptors and new ligands for biomolecules.

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3

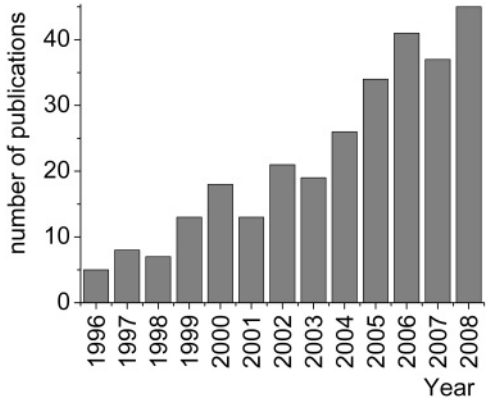


Figure 8.1 Number of publications dealing with DCC since the invention of the approach.

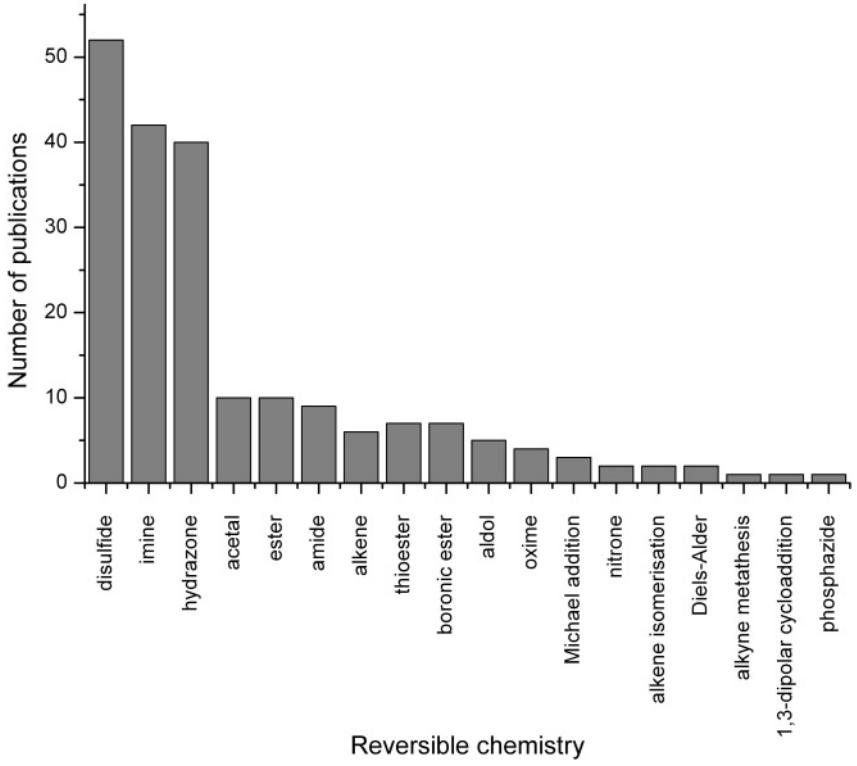


Figure 8.2 Number of publications until the end of 2008 on DCC grouped by type of reversible covalent chemistry.

It is evident from Chapters 3 and 5 that dynamic combinatorial libraries (DCLs) have indeed become popular and powerful tools for these purposes. The responsiveness of DCLs to external influences has also potential for other applications. First reports on the influence of electric fields [9], light [10–12], and temperature and pH [13] have recently appeared, and it is our expectation that there is a great deal more to be discovered by using these and other external stimuli. The same applies to the use of DCLs for identifying catalysts (see Chapter 4). The dynamic combinatorial approach has not yet seriously entered this area of science and only the first examples have appeared that show proof-of-principle. It is possible to design procedures to select a catalyst from a dynamic mixture of candidate molecules.

All of the applications mentioned above use an external molecular or physical stimulus with the aim of selecting particular molecules with special properties. While this approach is in essence reductionist (i.e., the purpose is the identification of individual molecules), the actual experiments require dealing with DCLs, which are complex mixtures. We have already seen in Chapter 2 that dynamic libraries can present counterintuitive behavior, which has important implications for the design of dynamic combinatorial experiments and may produce some pitfalls in the interpretation of the resulting data. Thus, efforts have been made to analyze the behavior of DCLs in order to come to a better understanding. While the incentive for these studies was initially to guide the design of dynamic combinatorial experiments, it gradually became apparent that the behavior of DCLs is an interesting subject in its own right. DCLs are complex networks of interacting molecules and constitute an important entry into the emerging field of systems chemistry [14].

It is clear from biology that networks of (bio)molecules can perform highly specific and advanced functions. There is currently a great drive to unravel the structure and mechanistic details of networks of biomolecules, which has given rise to the discipline of systems biology and, more recently, synthetic biology, where the insights derived from systems biology are used to engineer new systems. By analogy, it must also be possible to create networks of fully synthetic molecules that perform new and advanced functions that go well beyond a mere sum of the properties of the individual molecular constituents. In addition, fundamental understanding of biological systems properties may be facilitated by knowledge generated in the area of systems chemistry as the underlying principles are likely to be the same. The advantage of synthetic chemical systems is that we have ample tools to modify these systems such that we obtain a better understanding or an improved control over the system and its properties.

Indeed, there are already a number of examples in DCC where the focus is on the global behavior of the molecular network (i.e., the DCL) rather than on any of the individual network constituents. Examples include work on DCLs with unusual network topologies [15], where two subsets of ligands are unable to form metal–ligand complexes containing ligands from each of the two subsets. This characteristic influences the way such networks respond to the introduction of guest molecules. Another example is the formation of patterns in the amplification

behavior of DCLs of hosts that were simulated using randomly assigned host–guest binding affinities [16]. These patterns can emerge as a result of the influence of particular hosts that are already abundant in the absence of the guest and, upon addition of guest, compete strongly with the other species for the building blocks from which they are made up.

Another important trend is the use of DCLs as sensors. Here, the response of the entire library to the introduction of an analyte is recorded, and information about the nature and quantity of the analyte is obtained through multivariate analysis methods. When the kinetics of the library response are included in the analysis, the time of addition of analytes can also be determined. Chapter 7 describes these analytical applications in more detail.

A global analysis of the response of a DCL to a template molecule is determined by the affinities of the template for the various library members. Thus, the library response contains quantitative information about the template binding strengths of all library members that have significant affinities for the template. This allows us to obtain good estimates of the binding constants for the stronger hosts in a DCL directly from the library composition in the presence of different concentrations of guest [17]. This application is described in more detail in Chapter 2. This method bypasses the need for isolation of individual library members and thereby greatly increases the speed at which affinity data can be collected.

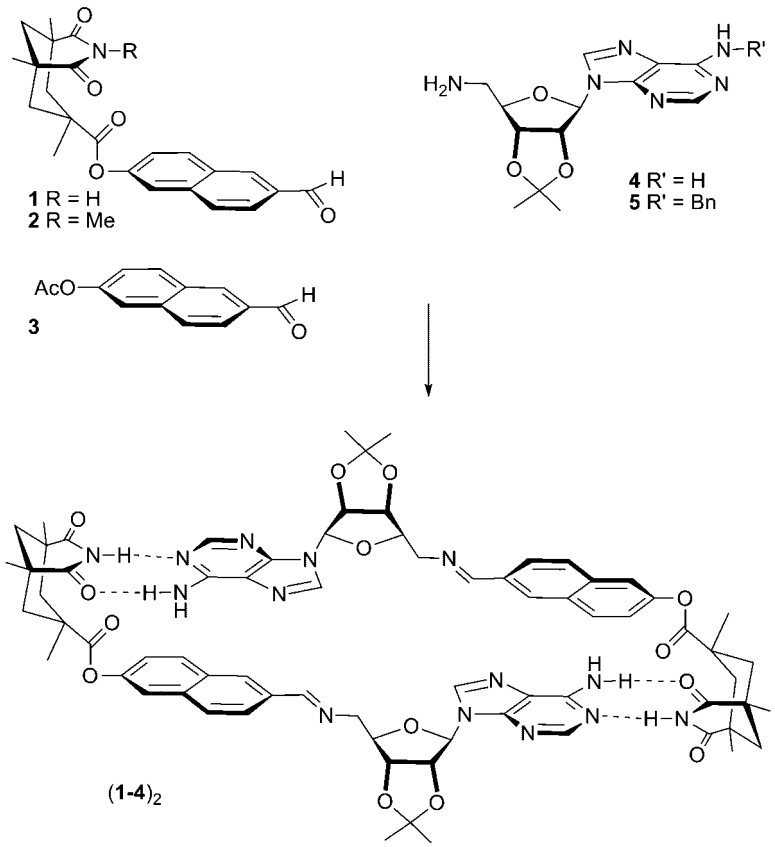
8.2.1 Self-Replication in DCLs

The above examples clearly illustrate a trend towards a more holistic approach of DCC. Further exponents of this trend are a number of recent reports that explore self-recognition in DCLs. In these systems molecular recognition between library members drives the synthesis of the very compounds that recognize each other (*cf.* Chapter 1, Figure 1.2d). Thus, the DCLs are self-screening and the selected structures can be considered to be self-replicating. We proposed this possibility back in 2006 [18], at which time the systems that came closest to showing such behavior were the disulfide-bridged lipids developed by Regen *et al.* [19]. It is gratifying to see that several systems have appeared since in which such internal templating is occurring.

Giuseppone recently reported a proof-of-principle study featuring a library member exhibiting self-complementary recognition through hydrogen bonding [20], inspired by the Rebek replicator [21]. The authors mixed aldehydes **1**–**3** with amines **4** and **5** (Scheme 8.1).

Of these building blocks, **1** has the potential to hydrogen bond to **4**. Thus, the imine formed from these two building blocks can dimerize to give the complex shown in Scheme 8.1 and this dimerization shifts the equilibrium in favor of the formation of this imine.

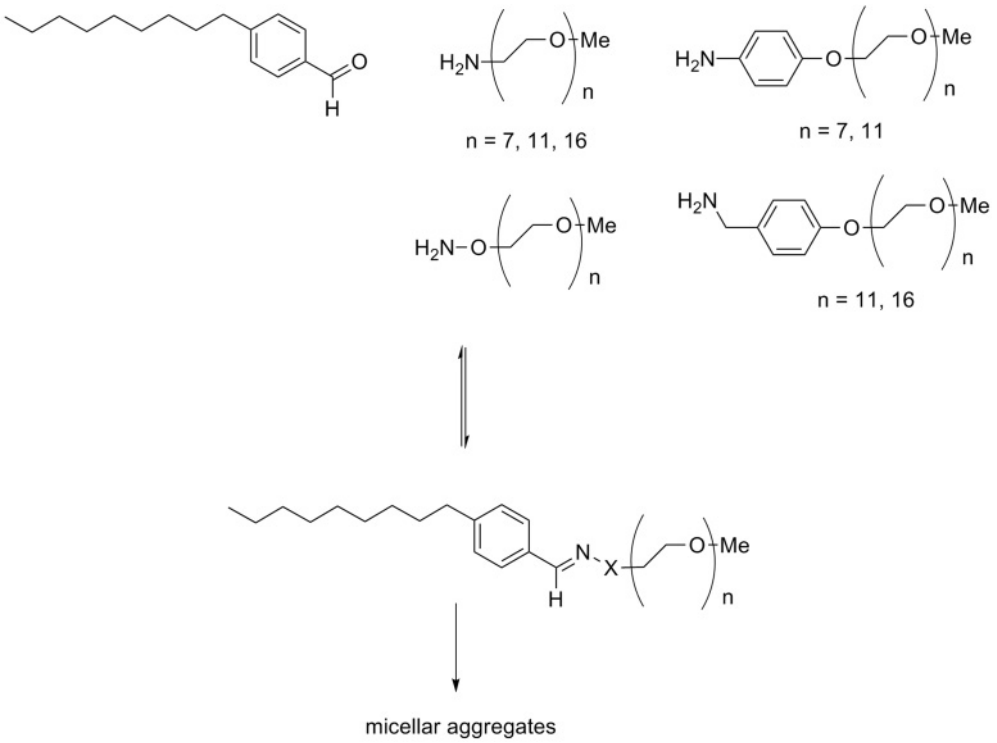
In more recent work, the same group has developed another self-replicating system based on the reversible formation of amphiphiles [22]. In this system a lipophilic aldehyde can combine reversibly with any of a series of hydrophilic



Scheme 8.1 Self-templating of an imine from a DCL of six possible imine products.

amines (Scheme 8.2). This results in the formation of amphiphilic imines that can aggregate to form micelles. These micelles promote the dissolution of the hydrophobic aldehyde and its subsequent reaction with a hydrophilic amine. Thus, the imine promotes its own formation. While this principle has been explored already in some detail by Luisi [23], the authors now introduce a new element to the system in the form of a pool of different amines. Thus, the system is able to select the "fittest" amine through reversible imine chemistry.

Another example where aggregation of a library member drives its synthesis was recently reported by Ulijn *et al.* [24, 25]. They used reversible amide bond formation, mediated by thermolysin, which is an enzyme that can catalyze both amide bond hydrolysis and formation, and is only moderately peptide-sequence-dependent. The authors reported that starting from dipeptides and fluorenyl-protected amino acids, the action of thermolysin gives rise to a dynamic mixture of peptides of different lengths (containing typically one to five amino acid residues). When using phenylalanine or leucine as the starting amino acids the

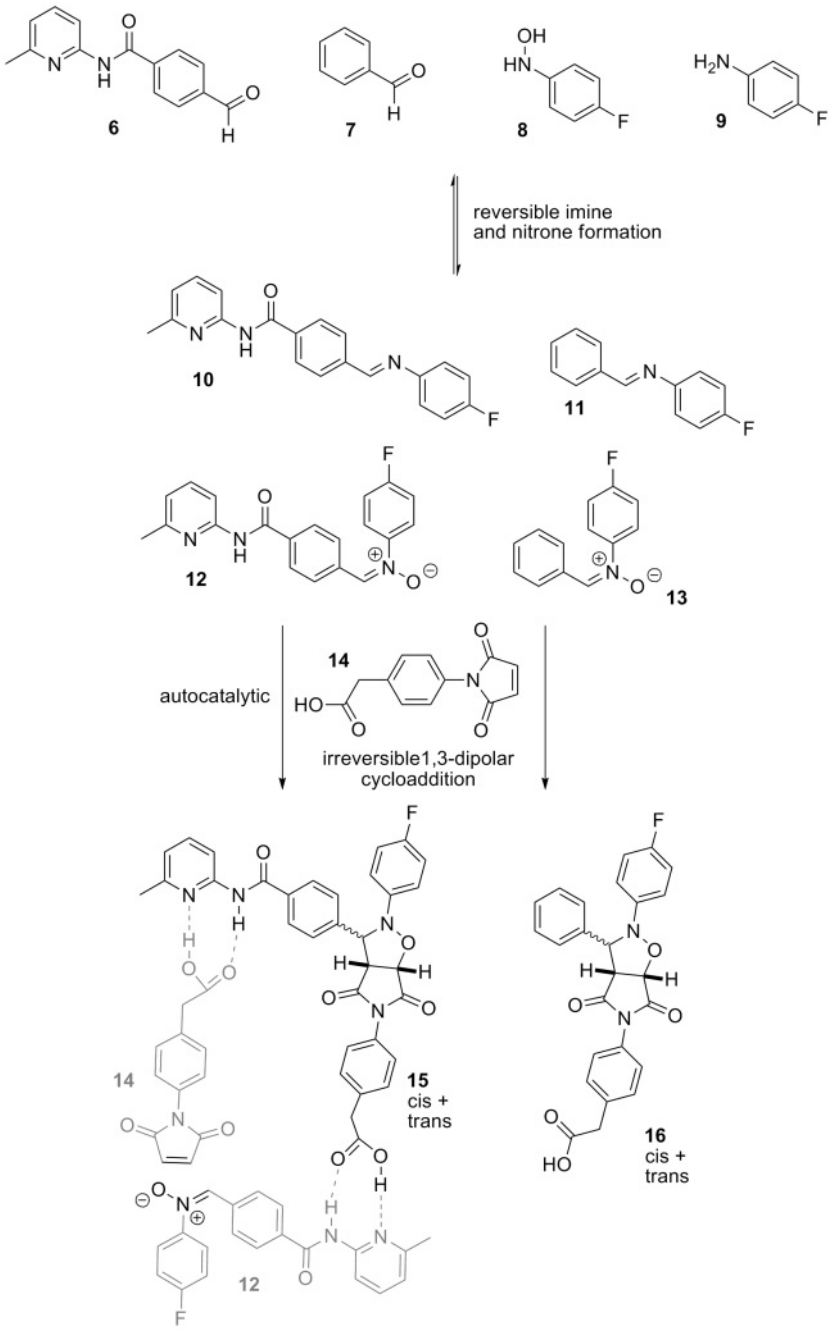


Scheme 8.2 Reversible formation of an imine surfactant in a system where surfactant aggregation shifts the equilibrium towards formation of the imine that aggregates most efficiently.

corresponding trimeric peptides were observed to self-assemble into fibers through a combination of stacking between the hydrophobic fluorenyl groups and β-sheet type main-chain hydrogen bonding. These interactions not only stabilize the nanostructures, but also cause the equilibrium between the various possible peptide products to shift in favor of the formation of the trimers. The fibers were of sufficient length and strength to give rise to entanglement and subsequent gelation of the aqueous solvent (see Figure 8.3).

Philp *et al*. have recently reported a beautiful example of a kinetically controlled self-replicating systems that feeds on a DCL, strongly biasing the final product distribution towards the autocatalytic product that had a concentration of zero at the start of the experiment [5]. The system was set up by mixing five ingredients: building blocks **6**–**9** and dipolarophile **14** (Scheme 8.3).

The four building blocks can combine through reversible imine and nitrone formation to produce a small DCL containing imines **10** and **11** and nitrones **12** and **13**. The two nitrones can undergo an irreversible 1,3-dipolar cycloaddition with maleimide derivative **14**. The *trans* product of the cycloaddition of nitrone **12** (but not of **13**) is a good catalyst for its own formation as it preorganizes **12** and



Scheme 8.3 Autocatalytic reaction that feeds on a DCL.

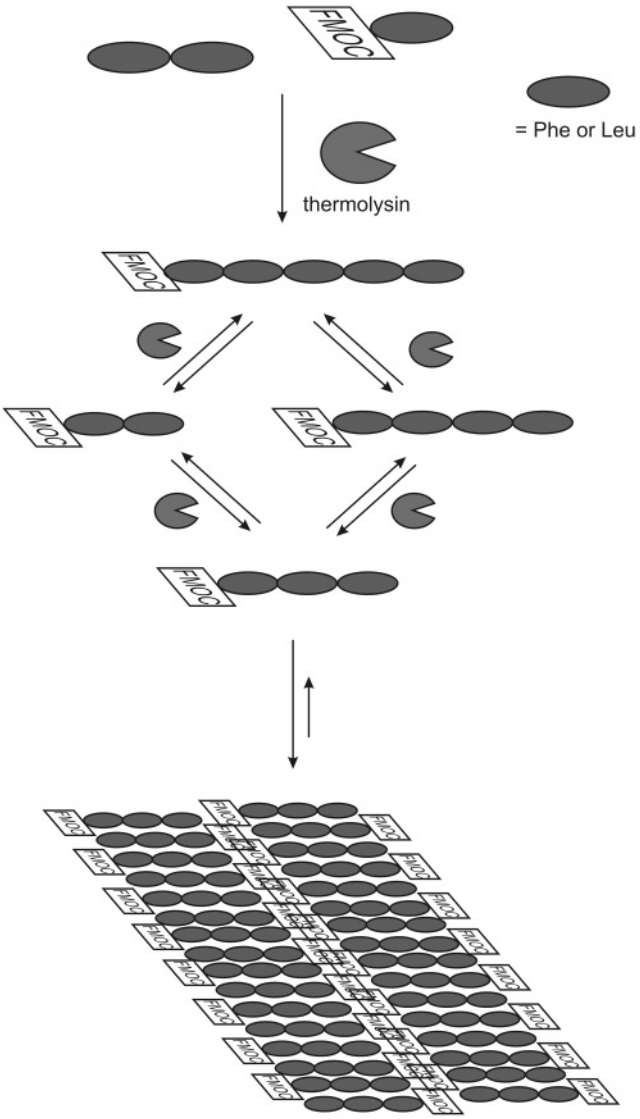


Figure 8.3 Thermolysin-mediated reversible amide bond formation gives rise to a DCL of short peptides. Aggregation of the trimeric peptide into fibers shifts the equilibrium towards this product and also causes gelation of the solvent.

14 through the formation of a ternary **12** · **14** · **15** complex as shown in Scheme 8.3. Autocatalysis by *trans*-**15** causes the cycloaddition of **12** to dominate over that of **13** causing the selective consumption building blocks **6** and **8** from the building block pool.

The above examples show that an interface is developing between DCC (i.e., dynamic molecular networks) and self-replicating systems, hinting at a possible role for DCLs in the origin of life.

8.3 Perspectives

DCC has been conceived as a technique for the discovery of new synthetic receptors and ligands for biomolecules. The results of little over a decade of research go well beyond merely establishing proof-of-principle in these areas. Lucid examples are the discovery by Miller *et al.* of lead compounds targeting myotonic dystrophy [26], and our own discovery of a synthetic receptor for spermine that is efficient enough to sequester spermine from DNA and thereby control the helicity of DNA [27].

Two potential applications of DCLs that remain underexplored are catalysis (Chapter 4) and the exploration of intramolecular recognition to give rise to molecules exhibiting stably folded structures. In both of these areas the field has not developed as rapidly and is still in the proof-of-principle phase.

However, in other areas DCC has expanded well beyond what it was originally conceived to do. When regarding DCLs not merely as discovery tools, but as molecular networks, many additional opportunities present themselves, some of which are now starting to be explored. A nice example includes the use of multiphase systems recently reported by Sanders *et al.* [28] We expect that the future will bring many new developments in new areas, which may include emergent behavior such as information processing [16], feedback behavior, and molecular communication.

References

1 Goral, V., Nelen, M.I., Eliseev, A.V., and Lehn, J.M. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 1347.
2 Hutin, M., Schultz, D., and Nitschke, J.R. (2008) *Chimia*, **62**, 198–203.
3 Sarma, R.J., Otto, S., and Nitschke, J.R. (2007) *Chem. Eur. J.*, **13**, 9542–9546.
4 Christinat, N., Scopelliti, R., and Severin, K. (2008) *Angew. Chem. Int. Ed.*, **47**, 1848–1852.
5 Sadownik, J.W. and Philp, D. (2008) *Angew. Chem. Int. Ed.*, **47**, 9965–9970.
6 Leclaire, J., Vial, L., Otto, S., and Sanders, J.K.M. (2005) *Chem. Commun.*, 1959–1961.
7 Orrillo, A.G., Escalante, A.M., and Furlan, R.L.E. (2008) *Chem. Commun.*, 5298–5300.
8 Rodriguez-Docampo, Z. and Otto, S. (2008) *Chem. Commun.*, 5301–5303.
9 Giuseppone, N. and Lehn, J.M. (2006) *Angew. Chem. Int. Ed.*, **45**, 4619–4624.
10 Eliseev, A.V. and Nelen, M.I. (1997) *J. Am. Chem. Soc.*, **119**, 1147.
11 Eliseev, A.V. and Nelen, M.I. (1998) *Chem. Eur. J.*, **4**, 825.
12 Inzgerman, L.A. and Waters, M.L. (2009) *J. Org. Chem.*, **74**, 111–117.
13 Giuseppone, N. and Lehn, J.M. (2006) *Chem. Eur. J.*, **12**, 1715–1722.

14 Ludlow, R.F. and Otto, S. (2008) *Chem. Soc. Rev.*, **37**, 101–108.

15 Saur, B., Scopelliti, R., and Severin, K. (2006) *Chem. Eur. J.*, **12**, 1058–1066.

16 Corbett, P.T., Sanders, J.K.M., and Otto, S. (2007) *Angew Chem. Int. Ed.*, **46**, 8858–8861.

17 Ludlow, R.F., Liu, J., Li, H., Roberts, S.L., Sanders, J.K.M., and Otto, S. (2007) *Angew. Chem. Int. Ed.*, **46**, 5762–5764.

18 Corbett, P.T., Leclaire, J., Vial, L., West, K.R., Wietor, J.L., Sanders, J.K.M., and Otto, S. (2006) *Chem. Rev.*, **106**, 3652–3711.

19 Davidson, S.M.K. and Regen, S.L. (1997) *Chem. Rev.*, **97**, 1269–1279.

20 Xu, S. and Giuseppone, N. (2008) *J. Am. Chem. Soc.*, **130**, 1826–1827.

21 Nowick, J.S., Feng, Q., Tjivikua, T., Ballester, P., and Rebek, J. (1991) *J. Am. Chem. Soc.*, **113**, 8831–8839.

22 Nguyen, R., Allouche, L., Buhler, E., and Giuseppone, N. (2009) *Angew. Chem. Int. Ed.*, **48**, 1093–1096.

23 Walde, P., Wick, R., Fresta, M., Mangone, A., and Luisi, P.L. (1994) *J. Am. Chem. Soc.*, **116**, 11649–11654.

24 Williams, R.J., Smith, A.M., Collins, R., Hodson, N., Das, A.K., and Ulijn, R.V. (2009) *Nat. Nanotechnol.*, **4**, 19–24.

25 Otto, S. (2009) *Nat. Nanotechnol.*, **4**, 13–14.

26 Gareiss, P.C., Sobczak, K., McNaughton, B.R., Palde, P.B., Thornton, C.A., and Miller, B.L. (2008) *J. Am. Chem. Soc.*, **130**, 16254–16261.

27 Vial, L., Ludlow, R.F., Leclaire, J., Perez-Fernandez, R., and Otto, S. (2006) *J. Am. Chem. Soc.*, **128**, 10253–10257.

28 Perez-Fernandez, R., Pittelkow, M., Belenguer, A.M., and Sanders, J.K.M. (2008) *Chem. Commun.*, 1738–1740.

Index

Dynamic Combinatorial Chemistry. Edited by Joost N. H. Reek and Sijbren Otto

ISBN: 978-3-527-32122-3