Reversible Sulfur Reactions in Pre-Equilibrated and Catalytic Self-Screening Dynamic Combinatorial Chemistry Protocols

Rikard Larsson

Licentiate Thesis
Stockholm 2006
Abstract

Dynamic Combinatorial Chemistry (DCC) is a recently introduced supramolecular approach to generate dynamically interchanging libraries of compounds. These libraries are made of different building blocks that reversibly interact with one another and spontaneously assemble to encompass all possible combinations. If a target molecule, for instance a receptor is added to the system and one or more molecules show affinity to the target species, these compounds will, according to Le Châtelier’s principle, be amplified on the expense of the other non-bonding constituents. To date, only a handful of different systems and formats have been used. Hence, to further advance the technique, especially when biological systems are targeted, new reaction types and new screening methods are necessary. This thesis describes the development of reversible sulfur reactions, thiol/disulfide interchange and transthiolesterification (the latter being a new reaction type for DCC), as means of generating reversible covalent bond reactions. Two different types of target proteins are used, enzymes belonging to the hydrolase family and the plant lectin Concanavalin A. Furthermore, two new screening/analysis methods not previously used in DCC are also presented; the quartz crystal microbalance (QCM)-technique and catalytic self-screening.

Keywords: Dynamic Combinatorial Chemistry, Reversible sulfur reactions, Catalytic self-screening, Carbohydrates, Lectins, Quartz Crystal Microbalance
List of publications

This thesis is based on the following papers, which are referred to in the text by their roman numerals:

I. Catalytic Self-Screening of Cholinesterase Substrates from a Dynamic Combinatorial Thioester Library
   Rikard Larsson, Zhichao Pei and Olof Ramström

II. Dynamic Combinatorial Thiolester Libraries for Efficient Catalytic Self-Screening of Hydrolase Substrates
    Rikard Larsson and Olof Ramström

III. Quartz Crystal Microbalance Bioaffinity Sensor for Rapid Identification of Glycodisulfide Lectin Inhibitors from a Dynamic Combinatorial Library
     Zhichao Pei, Rikard Larsson, Teodor Aastrup, Henrik Andersson, Jean-Marie Lehn and Olof Ramström

Reprints are presented with permission from the publishers: Wiley-VCH Verlag GmbH & Co KG (papers I and II) and Elsevier (paper III).
# Table of Contents

**ABSTRACT**

**LIST OF PUBLICATIONS**

**ABBREVIATIONS**

1 **INTRODUCTION**
   - 1.1 Combinatorial Chemistry
   - 1.2 Constitutional Dynamic Chemistry (CDC)
   - 1.3 Dynamic Combinatorial Chemistry (DCC)
     - 1.3.1 The dynamic process – generation of dynamic diversity
     - 1.3.2 Types of DCC formats
     - 1.3.3 Analysis and screening of DCLs
   - 1.4 Aim of the project

2 **SYNTHETIC STRATEGY OF DCLS**
   - 2.1 Synthetic strategy of thiolesters
   - 2.2 Synthetic strategy of thiols

3 **DYNAMICS OF THIOL- AND THIOLESTER REACTIONS**
   - 3.1 Introduction
   - 3.2 Transthiolesterification
     - 3.2.1 Effects of acyl groups on reaction dynamics
   - 3.3 Thiol/disulfide interchange

4 **TARGET PROTEINS**
   - 4.1 Introduction
     - 4.1.1 Hydrolases
       - 4.1.1.1 Mechanism of hydrolysis by serine hydrolases
       - 4.1.1.2 Cholinesterases
     - 4.1.2 Lectins
       - 4.1.2.1 Concanavalin A

5 **TYPES OF DYNAMIC COMBINATORIAL LIBRARIES**
   - 5.1 Catalytic self-screening from dynamic combinatorial thiolester libraries
     - 5.1.1 Introduction
     - 5.1.2 Catalytic self-screening of cholinesterase substrates
     - 5.1.3 Catalytic self-screening using different hydrolases
   - 5.2 Pre-equilibrated dynamic thiol/disulfide combinatorial libraries
     - 5.2.1 Introduction
     - 5.2.2 Screening and analysis of the DCLs using QCM biosensor

6 **FUTURE PROSPECTS**

7 **CONCLUDING REMARKS**

**ACKNOWLEDGEMENTS**

**APPENDIX**

**REFERENCES**
Abbreviations

ACh  Acetylcholine
AChE  Acetylcholinesterase
ASCh  Acetylthiocholine
Asp  Aspartic acid
BChE  Butyrylcholine esterase
β-Gal  β-Galactosidase
BSA  Bovine Serum Albumin
CA  Carbonic anhydrase
CaM  Calmodulin
CCL  *Candida cylindracea* Lipase
CDC  Constitutional Dynamic Chemistry
Con A  Concanavalin A
CRD  Carbohydrate Recognition Domain
DCC  Dynamic Combinatorial Chemistry
DCL  Dynamic Combinatorial Library
DTT  Dithiothreitol
EC  Enzyme Commission
Glu  Glutamic acid
HEWL  Hen Egg-White Lysozyme
His  Histidine
HLE  Horse Liver Esterase
NANA  *N*-acetylneuraminic acid (Sialic Acid)
NMR  Nuclear Magnetic Resonance
Phe  Phenylalanine
rt  Room Temperature
QCM  Quartz Crystal Microbalance
Ser  Serine
Sub C  Subtilisin Carlsberg
Trp  Tryptophan
WGA  Wheat Germ Agglutinin
1 Introduction

1.1 Combinatorial Chemistry

The original scientific principles of combinatorial chemistry can be traced back to the 1960’s when Bruce Merrifield started investigations on solid-phase synthesis of peptides. His Nobel Prize award-winning invention on the use of polymeric solid phase enabled simplified preparation and separation of peptides (Merrifield 1963). Since then, Merrifield’s Solid Phase synthesis concept has been refined and has spread in every field where organic synthesis is involved. Especially in the mid 1980’s several academic groups addressed different methods for parallel synthesis of peptides based on the solid-phase strategy first outlined by Merrifield. The design of the Multipin apparatus by Geysen and co-workers in 1984 allowed for individual peptides to be produced in parallel in micro plates containing 96 different wells (Geysen, Meloen et al. 1984).

![Diagram of split and mix procedure](image)

Figure 1. The split and mix procedure (Modified from Bornscheuer and Kazlauskas 1999).

Houghten’s method for multiple peptide synthesis first described in 1985, made use of polypropylene mesh packets, known as “tea-bags”, as porous capsules for solid phase resin beads (Houghten 1985). This system allows for the same peptide coupling step to be applied to many beads simultaneously irrespective of the sequence already attached to the bead. Houghten’s technique using porous polypropylene containers have since been further developed by different group using a radiofrequency encoding system (Moran, Sarshar et al. 1995; Nicolaou, Xiao et al. 1995). Despite these contributions, the split and mix procedure (Figure 1) first described by Furka in a notarized report in 1982 and published 1988 (Furka, Sebestyén et al. 1988; Furka, Sebestyén et al. 1988), and shortly after by Houghten (Houghten, Pinilla et al. 1991) and Lam (Lam, Salmon et al. 1991) marks the real birth of combinatorial chemistry.
The basic principle of combinatorial chemistry is to produce a large number of different compounds at the same time. Accelerating the process by which chemical compounds are produced, combinatorial chemistry has initiated major rethinking of how chemistry is practiced. In the past, chemists have traditionally made one compound at a time. For example compound A would have been reacted with compound B to give product AB, which would have been isolated after reaction work-up and purification through different techniques such as crystallization, distillation, or chromatography.

In contrast to this approach, combinatorial chemistry offers the potential to make every combination of compound A₁ to Aₙ with compound B₁ to Bₙ (Scheme 1). The range of combinatorial techniques is highly diverse, and these products could be made individually in parallel or in mixtures, using either solution or solid phase techniques. Whatever the technique used, the common denominator is that productivity has been amplified beyond the levels that have been routine for the last hundred years (Terrett 1998).

Scheme 1. The difference between traditional synthetic chemistry (left) and combinatorial chemistry (right).

Whilst combinatorial chemistry has helped to overcome limitations in synthesis by allowing large arrays of compounds to be synthesized in a fairly short time, much time and effort is dealt with the preparation and characterization of each compound synthesized. Although automation techniques and high-throughput methods have sped up these processes somewhat, it still remains a bottle-neck in combinatorial chemistry. Seeking to avoid synthesis and characterization entirely seems far fetched. However, if the target substance itself could be selected from a pool of compounds/constituents, and the library could adapt to the target constraints, the screening signal would be enhanced, making the screening process simplified. In addition, if the best bound species could be analyzed directly while bound to target, several synthetic steps could be avoided (Ramstrom and Lehn 2002).
1.2 Constitutional Dynamic Chemistry (CDC)

One way to overcome the limitations presented by combinatorial chemistry is to make use of dynamic chemistry. Constitutional Dynamic Chemistry (CDC) is a concept showing adaptive behavior through the generation of constitutional diversity by internal and/or external factors. These features define constitutional dynamic chemistry on both the molecular and supramolecular levels (Lehn 2002; Lehn 2002; Giuseppone and Lehn 2004; Giuseppone, Schmitt et al. 2004). The system may respond through selection of the best suitable constituents showing best adaptation to a given situation (Lehn 2005).

On the molecular level, CDC is expressed in dynamic combinatorial chemistry (DCC, Figure 2) an approach that uses self-assembly processes to generate libraries of chemical compounds.

![Figure 2. The scope of Constitutional Dynamic chemistry.](image)

1.3 Dynamic Combinatorial Chemistry (DCC)

Dynamic Combinatorial Chemistry (DCC), relies on the use of reversible processes to generate potent libraries of chemical compounds based on the continuous interchange of different building blocks within the library. One major advantage with dynamic combinatorial libraries (DCLs) over their static counterparts is their potential susceptibility to change in response to an external selection pressure. The DCC process can be divided into three different steps: first, selection of building blocks that must be able to reversibly interact with one another. These components are then spontaneously assembled to encompass all possible combinations, through the implementation of non-covalent or reversible covalent bonds, resulting in a pool of continuously interchanging library constituents, the dynamic combinatorial library. The last step includes trapping of the constituent(s). If, for instance, the dynamic library is exposed to a receptor showing affinity for one or more of the constituents formed, the dynamic system will shift in favor of the specific compounds bound (Figure 3). According to Le Châtelier’s principle (Le Chatelier 1884), the library will adapt to the selection pressure and the constituents bound to the receptor will be amplified in relation to the unbound compounds (Ganesan 1998; Klekota and Miller 1999; Lehn 1999; Cousins, Poulsen et al. 2000; Huc and Nguyen 2001; Ramstrom and Lehn 2002).

The dynamic nature of the process is a key feature of the DCC approach. Thus, the first prerequisite for a dynamic combinatorial library is that the building blocks must have functional groups that are able to undergo reversible exchange. In an ideal library this equilibrium should be sufficiently fast and the library should be put under thermodynamic control. It is...
also important for the process to be rapid enough to avoid degradation when applying DCC to sensitive biological macromolecules. Another important feature is that the connective groups should show similar or close to similar reactivities, this is to produce iso-energetic libraries and thus not creating biased libraries.

![Figure 3. Schematic representation of the concepts behind dynamic combinatorial chemistry. A dynamic library of interchanging species is formed through reversible exchange of initial building blocks. The best binder is selected by a target receptor, forcing the library to adapt and so produce more of this entity (Modified from Ramstrom and Lehn 2002).](image)

Another desirable feature is the allowance of fixation of the DCLs, i.e. freezing of the exchange process. This can be done by altering the surrounding conditions, for instance by changing pH, temperature or solvent composition. The addition of quenching agents, for example oxidation/reduction agents is other possibilities of freezing the exchange. In these cases, the reaction that is used must not disturb the equilibrium state (Ramstrom and Lehn 2002). The advantage of the freezing is that it allows for screening of the library without the screening process will alter the relative concentrations of the library components. More so, if the condition under which the interchange occurs is incompatible with the target, the equilibrium may be frozen prior to introduction of the host/guest. Freezing of the interchange will not lead to amplification, but still rapid library generation will occur.
1.3.1 The dynamic process – generation of dynamic diversity

DCLs can make use of several reversible connections and be accomplished by using any type of reversible physical or chemical mechanism. A selection of systems used and examples of systems for potential DCL are presented in Table 1.

![Diagram of various chemical reactions and interactions]

Table 1. A selection of reversible covalent bond formation and other reversible interactions that can be used in DCC. Other examples for potential DCC systems (not shown) include intramolecular processes such as: configurational (cis-trans isomerization), conformational (internal rotation; ring inversion) and structural processes (tautomerism; fluxionality)(modified from Lehn 1999).

Several of these systems have been successfully used, among others: imine formation, (Huc and Lehn 1997; Hochgur et al. Kroth et al. 2002; Hochgur et al. Biesinger et al. 2003; Godoy-Alcántar, Yatsimirsky et al. 2005); acyl hydrazone formation (Bunyapaiboonsri, Ramstrom et al. 2001; Bunyapaiboonsri, Ramstrom et al. 2003; Bornaghi, Wilkinson et al. 2004; Ramstrom, Lohmann et al. 2004); metal coordination (Klekota, Hammond et al. 1997; Buryak and Severin 2005); Diels–Alder formation (Boul, Reutenauer et al. 2005); metathesis (Cacciapaglia, Di Stefano et al. 2005); and disulfide formation (Hioki and Still 1998; Ramstrom and Lehn 2000; Erlanson, Lam et al. 2003; Sando, Narita et al. 2004), all sharing the common feature of being reversible reaction under mild conditions. These systems have
mainly been used since these formats have proven to be the most sufficient in the systems studied. This is especially the case when biological systems are targeted, since these systems require reactions that are stable under mild conditions in the aqueous phase. Overall, the DCC concept has proven to be a highly useful tool for screening and rapid identification of ligands with high affinity to target molecules such as receptors and enzymes (Ganesan 1998; Klekota and Miller 1999; Lehn 1999; Cousins, Poulsen et al. 2000; Huc and Nguyen 2001; Ramstrom and Lehn 2002).

1.3.2 Types of DCC formats

Apart from the different systems being used in DCC, different formats of the DCC concept have also been developed including the adaptive approach (Hasenkopf, Lehn et al. 1996; Huc and Lehn 1997; Otto, Furlan et al. 2000; Hochgurtel, Kroth et al. 2002; Otto, Furlan et al. 2002); the pre-equilibrated approach (Ramstrom and Lehn 2000; Bunyapaiboonsri, Ramstrom et al. 2001); the iterative approach (Elisev and Nelen 1997); and the pseudo-dynamic or deletion approach (Cheeseman, Corbett et al. 2002), all of which have their advantages and drawbacks and address specific challenges. In general they all have in common the reversible generation step, but differ in the selection step.

The adaptive approach is the first and foremost of the different approaches. Generation of constituents is carried out in the presence of the target. This results in an amplification of the best bound-species so that screening takes place simultaneously within the same compartment. All dynamic characteristics can be used in an adaptive DCL and therefore adaptation and amplification can be obtained. The second approach, termed the pre-equilibrated approach, involves generating the DCLs under reversible conditions that are equivalent to the adaptive approach. However, the screening process is performed under static conditions, resulting in no amplification taking place.

There are mainly two reasons for using the pre-equilibrated approach. The first reason is when working with sensitive biological targets that are unavailable in large amounts, and the second reason is when the connecting reaction requires conditions incompatible with the target. Although the screening process is performed according to traditional combinatorial methods, the dynamic process still rapidly generates the library.

In the iterative approach an iterative pre-equilibrated protocol is used. The DCL is generated in one compartment under defined conditions, the DCL compounds are then allowed to interact with the target, in a subsequent step, either in the same compartment or separately. In this case the unbound species must be re-transferred to the reaction chamber and be re-scrambled. After several runnings, this leads to an accumulation of the best binder, which can be analyzed (Elisev and Nelen 1997). The pseudo-dynamic or deletion approach (Cheeseman, Corbett et al. 2002) made use of a two compartment chamber model having a target in one of the chambers and an enzyme (pronase) in the other. The best target binders show a slower hydrolysis than the bad ones, rendering the best binder being the last remaining.
1.3.3 Analysis and screening of DCLs

Many different analytical techniques are possible when analyzing a DCL. For small libraries, methods such as one dimensional nuclear magnetic resonance (NMR), high-performance liquid chromatography (HPLC), high-performance capillary electrophoresis (HPCE), and gas chromatography (GC) can be used fairly easy to quantify single library constituents. Mass spectrometry (MS)-techniques is another technique which can also be used when dealing with compounds of discrete mass units. For larger libraries, evidently, the identification of constituents becomes more challenging. Amplification of the best binder of the target species in a DCL, resulting in increased concentration of that compound, simplifies the analysis. The best solution would be to instantly analyze what is bound to the target species. For this purpose MS-techniques combined with other separation techniques may be used to identify library compounds for instance time-of-flight electrospray mass spectrometry (Q-TOF-ESMS). Other protocols that may also be used for this purpose include different two dimensional nuclear magnetic resonance (NMR)-techniques.

Screening can also be performed using a dynamic deconvolution procedure. By preparing sub-libraries, in which one building block has been removed from the library, and testing these in comparison with the full library, large libraries can be screened in a short time. All species that contain this unit will be deleted from the library, and a decrease in inhibitory effect will indicate that the removed component is an important element in the generation of an active compound in the dynamic mixture.

1.4 Aim of the project

Several systems have been successfully used in DCC. However, still only a handful of different systems and formats have been used. An important challenge with DCC is the need for new means of generating reversible reactions. This is especially the case when biological systems are targeted, since these systems require reactions that are stable under mild conditions in aqueous media (pH, temperature etc). New methods are needed to advance the technique, and new reaction types are necessary for rapid, yet sufficiently stable DCL generation and screening. Furthermore, new analysis methods are needed to further advance the technique. The overall aim of this thesis is to explore reversible sulfur reactions, thiol/disulfide interchange and transthiolesterification (Scheme 2). It can be divided into two parts: first, the exploration of transthiolesterification (Papers I and II) as a way of a reversible covalent bond formation reaction having a potential for use in DCC; second, to use new techniques for screening and analysis of DCLs, namely, catalytic self-screening using different hydrolases (Papers I and II) and thiol/disulfide interchange of thiosaccharides analyzed with a Quartz Crystal Microbalance biosensor (Paper III).

![Scheme 2. Reversible sulfur reactions: a) Transthiolesterification; dynamic interaction between a thiol and a thiolester and b) thiol/disulfide interchange.](image-url)
2 Synthetic strategy of DCLs

2.1 Synthetic strategy of thiolesters

For the library using transthiolesterification, seven different thiolesters having different acyl
groups (a-g), ranging from acetyl to tert-butyl were prepared with 3-sulfanylpropionic acid
(1) from the respective acyl chloride (2a-g) to generate the thiolesters 1a-1g (Scheme 3). These
acyl groups were primarily chosen to present a homologous series of alkyl chains,
including linear and branched structures. The 3-sulfanylpropionic acid was used as thiol
counterpart to keep the acyl compounds soluble at neutral pH.

\[
\text{R} = (a) \text{Me}, (b) \text{Et}, (c) \text{n-Pr}, (d) \text{n-Bu}, (e) \text{n-Pe} (f) \text{i-Pr}, (g) \text{t-Bu}
\]

Scheme 3. Synthetic route for compounds 1a-1g: (i) R-COOH, CH\textsubscript{2}Cl\textsubscript{2}, 0 °C – rt, (76-91%).

2.2 Synthetic strategy of thiols

Several thiols used in the experiments are commercially available from different sources. Apart
from those, a range of different thiosaccharides (saccharides where a hydroxyl moiety
is replaced with sulfhydryl group) were synthesized for the thiol/disulfide interchange
experiments (Figure 4). They were made in order to generate libraries composed of structures
that mimic the natural carbohydrate ligands for a target lectin and to obtain structural
diversity. Another reason was to ensure the establishment of relatively "compact" ligands with
short spacers between the carbohydrate units, more closely resembling the oligosaccharide
glycosidic bonds.

Figure 4. Four different thiosaccharides synthesized for the thiol/disulfide interchange experiments.

Thiosaccharides, in general, represent an interesting group of compounds in glycochemistry
(Norberg 1996; Driguez 2001). The structures are often used as glycoside donors and
acceptors in oligosaccharide- and neoglycoconjugate synthesis (Crich and Li 2000; Marcaurelle
and Bertozzi 2001; Greffe, Jensen et al. 2002; Knapp and Myers 2002; Liakatos, Kiefel et al.
2003; Rich and Bundle 2004; Rye and Withers 2004; Zhu, Stolz et al. 2004). This is because of
the thiolate being a potent nucleophile and a weak base that react easily and
selectively with soft electrophiles (Norberg 1996). Another feature of thioglycosides and S-
linked conjugates is that they possess increased resistance to degradation by glycosidases.
This potentiates their use as efficient building blocks in drug design and therapeutics (Driguez
2001). Thiol-containing carbohydrates can also be easily oxidised to disulfide dimers, and
glycosyldisulfides have been identified as efficient glycosyl donors and potentially useful
glycomimetics (Davis, Ward et al. 2001; Szilagyi, Illyes et al. 2001).
In the present study, the lectin Concanavalin A (Con A) was used. It recognizes non-reducing α-D-mannose structures and for this reason the thioanalogues 1-thio-D-mannose (3) and methyl 6-thio-α-D-mannopyranoside (6) were synthesized. To obtain further structural diversity, the 1-thio analogs of N-acetyl-β-D-glucosamine (4) and N-acetyl-β-D-galactosamine (5) were also synthesized (André, Pei et al. 2006). The syntheses of compounds 3 and 6, exemplified in Scheme 4, were performed in few steps from the suitably derivatised carbohydrates by reaction with thioacetate, and subsequent deprotection.

![Scheme 4](image)

**Scheme 4. Synthetic routes.** (a) compound 3: (i) HSAc, BF₃·Et₂O, CH₂Cl₂, 0 °C – rt, (65%); (ii) NaOMe, MeOH, rt, (90%). (b) compound 6: (i) TsCl, pyridine, 0 °C – rt, (70%); (ii) KSAc, DMF, 60 °C, (60%); (iii) NaOMe, MeOH, H₂O, rt (90%)
3 Dynamics of thiol- and thiolester reactions

3.1 Introduction

The thiol/disulfide interchange reaction as well as transthiolesterification are biologically fundamental processes. The thiolester acetyl-Coenzyme A (acetyl-CoA), formed in a transthiolesterification reaction in the oxidative decarboxylation of pyruvate (Lehninger, Nelson et al. 1993), and the glutathione/oxidized glutathione pair (G–SH/G–SS–G), which forms the major intracellular redox buffer, responsible of maintaining the redox state of cells and for protecting the organism from oxidative stress (Gilbert 1997), are two such examples. Many enzymes such as: thiol proteases, enolase, β-ketoacylthiolase and thioredoxin, require a cysteine in the active site and are rendered inactive upon oxidation of the thiol group. Protein folding and cleavage of DNA by calichemicin and esperamicin are other biological processes where thiol/disulfide interchange plays an important role (Fernandes and Ramos 2004).

3.2 Transthiolesterification

Transthiolesterification, the reaction between a thiolester and a thiol, is a reversible reaction type that is rapid and sufficiently stable under mild conditions in aqueous media to be used in DCC. In order to test the dynamics of transthiolesterification a range of thiols were investigated (Figure 5). These thiols were of different nature and differed with respect to structure and functionality. Thus, thiols possessing an adjacent amine/ammonium group (12, 13), hydroxy group (14, 15, 21, 22), sulfonate group (16), carboxylate group (1, 17), carboxylic ester group (18), carboxylic amide group (19, 20), or trifluoromethyl group (23) were tested for their performances. An aromatic thiol (20) and two cyclic thiols (21, 22) were furthermore studied. All thiols were chosen so as to be soluble in aqueous media at neutral pH.

Figure 5. Thiol structures used in testing of dynamics of transthiolesterification.
The dynamic features of the thiols were subsequently evaluated in transthiolesterification reactions in D$_2$O with the acetylcholine (ACh) analog acetylthiocholine (ASch, 12a), where the formation/thiolysis of each thiolester was followed (Scheme 5).

![Scheme 5. DCL generation by transthiolesterification.](image)

By mixing equivalents of the thiols together with acetylthiocholine in NaOD/D$_3$PO$_4$ buffer at pD 7.0, the exchange rate and equilibrium composition of each combination was determined. The exchange was measured by $^1$H-NMR at different time intervals. For the rapidly reacting thiols ($t_{1/2} < 15$ min) only one measurement was possible. The rate of exchange directly correlates to the pK$_a$ of the thiols (Table 2).

<table>
<thead>
<tr>
<th>Thiol</th>
<th>pK$_a$</th>
<th>Exchange with acetylthiocholine</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ratio</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>7.7</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>7.7</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>9.7</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>9.5</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>9.1</td>
</tr>
<tr>
<td>6</td>
<td>17</td>
<td>9.8</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>10.3</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>7.8</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>8.5</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>6.1</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>7.7</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>7.7</td>
</tr>
<tr>
<td>13</td>
<td>23</td>
<td>7.3</td>
</tr>
</tbody>
</table>

The lower the pH, the faster the exchange reaction. Thiols having pH values lower than 8.5 all reached equilibrium very rapidly ($t_{1/2} < 15$ min). The results also indicate that the majority of thiols produce equilibrium concentrations that are close to the concentration of acetylthiocholine, thus showing near isoenergetic behavior. Thiolesters from secondary thiols were however considerably less stabilized compared to acetylthiocholine. For the aromatic thiol 20, and the 1-thio-β-D-glycopyranosides 21 and 22, the ratio was clearly shifted in favor of the reactants, and only about 10-20% of the thiols were present as the corresponding thiolesters at equilibrium. When comparing all thiols, these components showed the largest differences in reactivity.

3.2.1 Effects of acyl groups on reaction dynamics

The effects of the acyl groups were estimated from two different dynamic libraries. In the first of these, DCL-A, five of the thiolesters 1a-1e were treated with thiocholine (12), and the exchange was monitored by $^1$H-NMR at pD 7.0 at room temperature. The exchange rate was in this case relatively rapid and equilibration of the library was reached with a $t_{1/2}$ value of 50 min (Figure 6). The second dynamic library DCL-B was instead composed of thiolesters 1a-1c and If-Ig and subjected to the same conditions as DCL-A. Library generation was in this case less efficient, and equilibrium from the reaction with thiol 12 was attained considerably slower ($t_{1/2} = 110$ min, Figure 6). As expected, the results from these libraries implied that the branched acyl groups f and g reduced the exchange rate of the libraries, the carbonyl group being more sterically hindered than for the linear acyl groups.

Figure 6. Equilibration of dynamic thiolester libraries: (■) DCL-A from thiol 12 and thiolesters 1a-1e; (▼) DCL-B from thiol 12 and thiolesters 1a-1c, If-Ig. Both DCLs were made using equimolar amounts of substrates.

As previously described, in order to avoid creating biased libraries, it is desirable to include components that show comparable reactivities, and for this reason some components were excluded from the subsequent libraries. Due to the slower exchange rate for some of the thiols and the branched acyl components, these were generally excluded. In addition, the secondary thiols were excluded because of their unfavorable equilibria. In a broader perspective, however, the excluded thiol or acyl components tested are not generally disqualified for transthiolesterification libraries, but may well be part of dynamic libraries for other purposes. All of these components are still taking part in the dynamic exchange, albeit showing slower kinetics or unfavorable equilibria. A mean to increase the rate of transthiolesterification for the less reactive thiols is also to increase the basicity of the solution, and the dynamic exchange was also efficient at higher pH. However, increasing the basicity not only
accelerates thiolyses, but also the competing and unproductive hydrolyses of the thiolesters. In the present study, a neutral pD was generally chosen, because of negligible hydrolysis at the time scale used. Hydrolysis was thus considerably less pronounced at pD 7.0 (3.7%) compared to pD 8.0 (13.5%) and pD 9.0 (17.9%) after three days. However, it is still possible for a higher pD-value to be chosen. Apart from working under nitrogen atmosphere and using degassed solvents, in order to prevent oxidation, and thus disulfide formation, reducing agents such as dithiothreitol (DTT) or phosphines (Tam, Lu et al. 1999) can be added to the media.

3.3 Thiol/disulfide interchange

To be able to predict the reactivity of the thiol/disulfide interchange, intense research on structure-activity relationships have been established for different kinds of thiols. Factors that have to be considered are steric hindrance and what solvent is used. Thiol/disulfide interchange (Scheme 6), occurs by an S$_{N}$2 displacement mechanism in which the reactive form of the thiol, the thiolate anion, approaches the disulfide bond along its S–S axis (Rabenstein and Weaver 1996).

\begin{align*}
(1) & \quad R_1\text{-SH} + H_2O & \rightleftharpoons & \quad R_1\text{-S}^{\ominus} + H_3O^{\ominus} \\
(2) & \quad R_1\text{-S}^{\ominus} + R_2\text{-S} - R_2 & \rightarrow & \quad R_1\text{-S-S-R}_2 + R_2^{\ominus} \\
(3) & \quad R_1\text{-S-S-R}_2 + R_1\text{-S}^{\ominus} & \rightleftharpoons & \quad R_1\text{-S-S-R}_1 + R_2^{\ominus} \\
(4) & \quad R_2\text{-S}^{\ominus} + H_3O^{\ominus} & \rightleftharpoons & \quad R_2\text{-SH} + H_2O
\end{align*}

**Scheme 6.** The thiol/disulfide interchange reaction represented schematically by chemical equations 1-4.

The interchange proceeding through an S$_{N}$2 transition state is the most favorable mechanism from a kinetic point of view and the kinetics of disulfide reduction is mostly determined by two factors: the basicity and the nucleophilicity of the thiol group (Whitesides, Lilburn et al. 1977; DeCollo and Lees 2001). Thiols with high pK$_a$-values (pH ~ 9.0) would be expected to be strong nucleophiles and thiols with low pK$_a$-values (pH ~ 4-5) would be expected to be weak ones. However, the reacting species is believed to be the thiolate anion, therefore, low pK$_a$ thiols would be favored at physiological pH. Balancing these two factors it emerges that the most reactive thiols are those with a pK$_a$ close to the pH of the solution (Whitesides, Lilburn et al. 1977; DeCollo and Lees 2001; Fernandes and Ramos 2004).
4 Target Proteins

4.1 Introduction

Several protocols that target biological macromolecules or biogenic ligands have been developed for DCC. Biological molecules are the most interesting, but also the most challenging targets to use. They are often found in only small amounts and are often unstable when used in harsh environment for a long period of time.

As for protein targets, mainly enzymes and receptor proteins have been used. Enzymes that have been used include; acetylcholine esterase (AchE, Bunyapaiboonsri, Ramstrom et al. 2001) and carbonic anhydrase (CA, Huc and Lehn 1997). CA has also been used together with another enzyme, pronase, which was used to hydrolyze constituents not binding the target (Cheeseman, Corbett et al. 2002). The enzyme N-acetylneuraminic acid (NANA) aldolase has been used together with a receptor protein, wheat germ agglutinin (WGA, Lins, Flitsch et al. 2002) known to specifically bind NANA (Sialic acid). In addition to WGA, other receptor proteins that have been used include; Concanavalin A (Ramstrom and Lehn 2000; Ramstrom, Lohmann et al. 2004), hen egg-white lysozyme (HEWL, Zameo, Vauzeilles et al. 2005) and calmodulin (CaM, Milanesi, Hunter et al. 2006). The small bacterial cell wall binding-block peptide D-Ala-D-Ala have also been probed with vancomycin derived elements (Nicolaou, Hughes et al. 2000). Non-protein targets such as nucleotides have also been probed (Klekota, Hammond et al. 1997; Karan and Miller 2001; Bugaut, Toulme et al. 2004; Bugaut, Bathany et al. 2005).

4.1.1 Hydrolases

Enzymes are divided into six different categories depending on what type of reaction is catalyzed. For identification purposes, every enzyme has a four digit Enzyme Commision (EC)-number. The EC classification for these enzymes generally classifies them firstly by the main type of reaction, secondly by the nature of the bond hydrolysed, then by the nature of the substrate, and lastly by the individual enzyme (Faber 2000). The hydrolases, EC 3, catalyze the hydrolytic cleavage of C–O, C–N, C–C and some other bonds, including phosphoric anhydride bonds. The systematic name always includes hydrolase, however, the recommended name is in many cases formed by the name of the substrate with the suffix -ase. It is understood that the name of the substrate with this suffix means a hydrolytic enzyme.

Several features have helped making hydrolases popular for biotransformations and have made hydrolases very useful to organic chemists. Hydrolases often show broad substrate specificity, high stereoselectivity and they also catalyze several related reactions, such as condensations and alcoholysis. A large number of hydrolases are commercially available, they do not require cofactors and they tolerate, to some extent, the addition of water miscible solvents (e.g. DMF, DMSO). Furthermore, lipases, esterases and some proteases are also stable and active in neat organic solvents (Bornscheuer and Kazlauskas 1999).
4.1.1.1 Mechanism of hydrolysis by serine hydrolases

The mechanism of amide- and ester-hydrolyses (Scheme 7) is very similar to that of chemical hydrolysis by base. A nucleophilic group, situated in the active site of the enzyme, attacks the carbonyl group of the substrate. This nucleophilic operator can be either the hydroxy group of a serine, a carboxy group of a aspartic acid or a thiol functionality of a cysteine. The mechanism of the serine hydrolases is the one that has been most elucidated in detail (Faber 2000).

![Scheme 7. Mechanism of serine hydrolases.](image)

The main players of the serine hydrolases are the three amino acids forming the catalytic triad (Brady, Brzozowski et al. 1990). This structure is preserved in all serine hydrolases and consists of three essential amino acids; serine (Ser), histidine (His) and aspartic acid (Asp). Each amino acid in the triad plays an essential role and performs a specific task in the catalytic process. The serine has an -OH group that act as a nucleophile, attacking the carbonyl carbon in the substrate. The pair of electrons on the nitrogen histidine has the ability to accept the hydrogen from the serine -OH group, thus coordinating the attack of the substrate. The carboxylic group on the aspartic acid turns hydrogen bonds with the histidine, making the pair of electrons more electronegative. When, for example, a peptide enters the active site, the serine -OH attacks the carbonyl carbon, the nitrogen of the histidine accepts the hydrogen from the -OH of the serine and a pair of electrons from the double bond of the carbonyl oxygen moves to the oxygen. As a result, a tetrahedral intermediate is generated. The bond joining the carbon and the nitrogen in the peptide bond is broken. The electrons creating this bond attack the hydrogen of the histidine, breaking the connection. The electrons that previously moved from the carbonyl oxygen double bond move back from the negative oxygen to recreate the bond, generating an acyl-enzyme intermediate. Now, water can act as a nucleophile and attack this acyl-enzyme intermediate. The water replaces the N-terminus of the cleaved peptide, and attacks the carbonyl carbon. In case of the enzyme working in an environment of low water activity, other nucleophiles can compete with the water, leading to number of useful transformations (Faber 2000). The electrons from the double bond once again move to the oxygen making it negative, as the bond between the carbon and the oxygen of the water is formed. This is all coordinated by the nitrogen of the histidine, which accepts a proton from the water. Overall, this generates another tetrahedral intermediate. Now, the bond...
formed in the first step between the serine and the carbonyl carbon moves to attack the hydrogen that the histidine just acquired. The now electron-deficient carbonyl carbon reforms the double bond with the oxygen. As a result, the C-terminus of the peptide is ejected.

4.1.1.2 Cholinesterases

The cholinesterases, acetylcholinesterase and butyrylcholinesterase, are serine hydrolase enzymes that belong to the α/β-protein family within higher eukaryotes. Whereas the biological role of acetylcholinesterase (AChE, EC 3.1.1.7) is the termination of impulse transmissions at cholinergic synapses within the nervous system (Figure 7), where it catalyzes the hydrolysis of the neurotransmitter acetylcholine (ACh) to acetate and choline (Scheme 8) (Schumacher, Camp et al. 1986; Sussman, Harel et al. 1991), no physiological function has yet been ascribed for butyrylcholinesterase (EC 3.1.1.8). BChE tolerates a large variety of esters and it is more active with butyryl and propionyl choline than with acetyl choline (Main, Soucie et al. 1974). Structure-activity relationship studies have shown that different steric restrictions in the acyl pockets of AChE and BChE cause the difference in their specificity with respect to the acyl moiety of the substrate (Järvi 1984). Apart from its catalytic function, AChE also affects cell proliferation, differentiation and responses to various insults (Venekei, Szilagy et al. 1996).

Scheme 8. Hydrolysis of acetylcholine by acetylcholinesterase, AChE.

AChE hydrolyzes ACh at a very high rate. The maximal rate for hydrolysis of AChE and its thio analog acetylthiocholine (ASCh) are around $10^9$ M$^{-1}$s$^{-1}$ (Nolte, Rosenberry et al. 1980), approaching the upper limit allowed by diffusion of the substrate. This has caused a lot of debate, since the active site is localized at the bottom of a narrow 20 Å deep catalytic gorge (Sussman, Harel et al. 1991). It is evident that this structure restricts the “traffic” in and out of the active site. Mechanisms to overcome these restrictions have been widely discussed in the literature and the mechanism of aromatic or electrostatic guidance, electrostatic trapping and the presence of a “back door” have all been discussed (Tougu 2001).

Figure 7. The mechanism of release of ACh from the synapse by exocytosis into the synaptic cleft, where it is rapidly hydrolyzed by acetylcholinesterase.
The three-dimensional structure of AChE from the Pacific electric ray (*Torpedo californica*) was resolved in 1991 (Sussman, Harel et al. 1991). The active site is composed of two subsites: the esteratic subsite which contains the catalytic triad, and the anionic subsite that accommodates the positive quaternary pole of acetylcholine. The esteratic subsite contains the catalytic machinery of the enzyme: a catalytic triad of Ser200, Glu327, and His440 (Figure 8). Glu appears as the third residue instead of Asp, typically present in serine proteases. In addition, the triad is a mirror image to that of other serine proteases and AChE also possesses a three-pronged oxyanion hole (Ordentlich, Barak et al. 1993) as opposed to the corresponding two-pronged structures in other serine proteases. The anionic subsite is defined by Trp84, Phe330, and Phe331 (Figure 8). Its role is to orient the charged part of the substrate that enters the active site (Sussman, Harel et al. 1991).

![Figure 8. Acetylcholine (grey) bound to the active site of acetylcholinesterase. The active site contains two subsites: the esteratic subsite (green) defined by Ser200, Glu327 and His440; and the anionic subsite (purple) defined by Trp84, Phe330, and Phe331.](image)

The peripheral anionic site (PAS), is situated on the surface of the enzyme at or near the rim of the aromatic gorge (Barak, Kronman et al. 1994). This site has the ability to bind to many different types of ligands, and by doing so effects the conformation of the active center, thus changing the functionality of the enzyme. It has been suggested that this is a result of evolutionary design aimed to confer optimal catalytic activity under a wide variety of conditions that are characteristic for the operation of acetylcholinesterase in the synaptic cleft (Ordentlich, Barak et al. 1993). However, the potent allosteric inhibitor, fasciculin, blocks the gorge entrance completely (Bourne, Taylor et al. 1995), hence the inhibition should occur without involvement of additional complicated mechanisms. Still, it remains obscure what evolutionary advantage could be achieved by the allosteric regulation of AChE activity (Tougu 2001).

### 4.1.2 Lectins

Lectins are a class of proteins which occur widely in nature and are capable of specific recognition of, and binding to, carbohydrate ligands. They represent a family of biological targets with strong potential bearing on drug discovery as well as various biotechnological applications (Karlsson 1991; Yarema and Bertozzi 1998; Wong 1999; Bertozzi and Kiessling 2001; Gabius, Andre et al. 2002; Gabius, Siebert et al. 2004). Lectins not only distinguish between different monosaccharides, but also show a remarkable specificity in their
recognition of highly branched complex carbohydrates. The binding occurs in the carbohydrate recognition domain (CRD) and there is usually only one way that a certain branched sugar can enter the CRD, rendering the lectins specificity. The binding of the lectin is largely due to hydrogen bonding between the backbone and side chain carbonyl groups in the protein and the hydroxyl groups of the sugars. Protein bound Ca$^{2+}$ and a transition metal ion, predominantly Mn$^{2+}$, can also chelate vicinal hydroxyl carbohydrate groups. Van der Waals’ interactions between hydrophobic lectin protein residue side chains and hydrophobic “patches” also increase the binding affinity. The binding between a single sugar and the CRD is usually very weak. However, when more than one sugar of the right type form clusters and accommodate the right orientation there is a rapid increase in both the affinity and specificity of the corresponding lectin. This additive effect resulting from the bond carbohydrate structures have been termed the multivalent or cluster effect (Davis and Fairbanks 2002; Lundquist and Toone 2002).

4.1.2.1 Concanavalin A

Concanavalin A (Con A) is a lectin protein isolated from the Jack bean (Canavalia ensiformis). The specificity of Con A is non-reducing and α-D-mannose structures and has its optimal activity at neutral pH. Con A is composed of four identical subunits of 237 amino acid residues with no covalently bound carbohydrate or prosthetic group. The amino acid chain consists of two antiparallel beta sheets, the first one has six strands and the other one has seven strands (Figure 9). Above pH 7, Con A exists mainly as a tetramer: Two six-stranded sheets are joined together to form a twelve-stranded dimer, which in turn form a functional complex by layering side to side with another dimer, thus resulting in four binding sites for the saccharides (Wang, Cunningham et al. 1975).

![Figure 9. Monomeric subunit of Concanavalin A, a plant lectin from the Jack bean (Canavalia ensiformis). The carbohydrate recognition domain, CRD, is adjacent to the two metal ions; calcium (black) and manganese (grey)](image)

In its native state one monomer of Con A binds two metal atoms; one Ca$^{2+}$ ion and one ion of a transition metal, usually Mn$^{2+}$ that must be present for saccharide binding. The metals are bound by amino acids that point away from the seven strand curved sheet. The concavity of this face provides a shallow carbohydrate binding site that is easily accessible not just to monosaccharides, but to oligo- and polysaccharides as well (Lis and Sharon 1998).
5 Types of Dynamic Combinatorial Libraries

5.1 Catalytic self-screening from dynamic combinatorial thiolester libraries

The dynamic features of the transthiolesterification reaction have been probed for a range of components of different character, resulting in potent dynamic thiolester libraries. These libraries were also exposed to a series of different hydrolases, where the performance and selectivity of the self-screening process were investigated.

5.1.1 Introduction

DCC offers the potential susceptibility for compounds to change in response to an external selection pressure. If the binding event is coupled to a secondary process, the selection may be enhanced by allowing the bound species to be removed from the equilibrating pool. Thus, coupled to for instance a biocatalyst, the dynamic process would generate more of the best recognized constituent of the library. Following the catalytic action, the products would be expelled from the active site, thus rendering the site free to host more of the DCL constituents and forcing the dynamic system to run to completion. Substrates to the biocatalyst would therefore be selectively produced and easily identified using this self-screening dynamic system. This process generates more of the best bound species by re-equilibration of the DCL (Scheme 9).

[Diagram: Scheme 9. Schematic representation of the DCL catalytic self-screening process with thiolesters.]

Generation of the dynamics in the DCC systems made use of transthiolesterification (Larsson, Pei et al. 2004), which had not previously been used in DCC systems. It was shown that this reversible reaction type was rapid and sufficiently stable under mild conditions in aqueous media. Since our report, transthiolesterification have been used in other protocols (Woll and Gellman 2004; Leclaire, Vial et al. 2005).
5.1.2 Catalytic self-screening of cholinesterase substrates

The dynamic combinatorial libraries were generated from a series of thiolesters and thiols, respectively. The reaction between the five thiolesters 1a-1e and the four thiols 12, 13, 16 and 19 generated library DCL-C (Scheme 10). For every thiol added, five additional thiolesters are formed. Thus, during formation of the dynamic libraries these components form 25 different thiolesters, all of which being in exchange with all others during the whole process.

The transthiolesterification reaction took place effectively under mild conditions in aqueous media by simple mixing all the components. The library generation process was initiated with equimolar amounts of all acyl components and five equivalents of each of the thiols. Since thiol component 1 is connected to the five different acyl functionalities at $t_0$, this ensured equal quantities of all thiol components in the system. Thus, the resulting concentrations of the formed constituents were relatively comparable and the libraries showed close to isoenergetic behavior. Experiments where thiol 1 was left in excess to the other thiols yielded the same final results albeit showing slightly longer reaction time.

Treatment of the thiolester library with acetylcholinesterase resulted in the best substrate being immediately recognized by the enzyme and thus, hydrolyzed. This resulted in loss of the acyl component from the library, which forced the library to reconstitute to accommodate the increasing amounts of free thiol and to generate more of the hydrolyzed species. The catalytic action of the enzyme acetylcholinesterase, (AChE, EC 3.1.1.7), a serine hydrolase that catalyzes the hydrolysis of the neurotransmitter acetylcholine to acetate and choline at neuromuscular synapses (Sussman, Harel et al. 1991) was relatively slow under the same set of conditions. This lower rate was chosen to ensure pseudoequilibrium conditions of the libraries during the enzymatic action, with scrambling being faster than enzyme catalysis. Over time, two of the acyl functionalities, the acetyl and propionyl groups, proved to be mainly acted upon by the enzyme, with the acetyl species being more rapidly hydrolyzed than the propionyl counterpart (Figure 10d). The $^1$H-NMR spectrum of a mixture of the five original acyl components and five original thiol components is displayed in Figure 10a and the spectrum after formation of the full library is shown in Figure 10b. Figure 10c shows the library in presence of acetylcholinesterase after complete hydrolysis of the acetate and propionate functionalities. The overall rates of formation proved to be: for acetic acid, $t_{1/2} = 260$ min and for propionic acid, $t_{1/2} = 310$ min. Only after substantial hydrolysis of the two
main substrates (~500 min) did the enzyme start hydrolyzing the butyrate ester, and even
then, the reaction was considerably less efficient (t_{1/2} > 1800 min).

Figure 10. *H-NMR spectra of libraries/components: a) Components for DCL-C before library generation; b) DCL-C in the absence of AChE c) DCL-C in the presence of AChE, *, #, and × indicate the signals for acetate, propionate, and butyrate groups, respectively d) formation of acetate (■), propionate (▲) and butyrate (▼) hydrolysis products in DCL-C.

This long lag phase may be caused by inhibitory activities of the present esters (Cho, Garsia et al. 1996). All the other acyl groups remained untouched by the enzyme, a result which is in accordance with the known specificity of acetylcholinesterase. Addition of the enzyme after complete equilibration of the library did not result in any major differences and the same distribution of compounds was recorded, as had been anticipated. In both cases, the combined library interchange and catalytic processes resulted in the complete hydrolysis of certain acyl thioesters to form the corresponding acids and thiols. Thus, this catalysis assistance could be used to fully amplify the selected acyl groups. Removal of the thiols failed to generate any hydrolytic products under the same conditions, thus demonstrating that they were essential for the reaction to occur.

The different thiols, 12, 13, 16 and 19, were also analyzed alone with the five thiolesters, 1a-1e. During formation of the dynamic libraries these components form ten different thiolesters, which are all in exchange with the others during the whole process. When thiocholine was used as the thiol, DCL-A, again the two acyl functionalities, acetyl and propionyl groups, proved to be mainly acted upon by the enzyme, with the acetyl species being more rapidly
hydrolyzed than the propionyl counterpart (Figure 11c-d). The $^1$H-NMR spectrum of a mixture of the five original acyl components is displayed in Figure 11a and the spectrum after formation of the full library is shown in Figure 11b.

![Figure 11. $^1$H-NMR spectra of DCL-A components/constituents: a) Before library generation, b) DCL in the absence of AChE, c) DCL in the presence of AChE, d) enlarged section of (c). *#, and × indicate the signals for acetate, propionate, and butyrate groups, respectively.](image)

The overall rate of formation of acetic acid ($t_{1/2} = 210$ min) proved to be around 20% faster than the rate of formation of propionic acid ($t_{1/2} = 270$ min, Figure 12. Similar to the full library, these products were formed at a significantly faster rate than butyric acid ($t_{1/2} = \text{ca. 1500 min}$).

The screening of the DCL thus proved straightforward, and the selection effect of the combined process can be estimated as the ratio of the amount of hydrolytic products formed versus the amount of substrate fragment added. In the present case, 100% selection/amplification was achieved for the acetate and propionate groups, respectively.

The additional libraries using one of the thiols 13, 16 or 19 together with acyl constituents 1a-1e, yielding libraries DCL-D, DCL-E, and DCL-F, respectively. The half-lives of formation of the different hydrolysis products, acetate, propionate and butyrate, in the five different DCLs, are summarized in Table 3. For DCL-A, using only thiocholine (12) together with each thiolester 1a-1e, the results were very similar to DCL-C, and the acetyl and propionyl acyl groups were instantly recognized by the enzyme. The slightly shorter time required is a consequence of the smaller library size. For every other 10-compound library, the overall rates of formation were considerably less sufficient, well in accordance with the known substrate specificity of acetylcholinesterase.
It was anticipated that the strong resemblance between compounds 12 and 13 could lead to the latter substrate(s) also being hydrolyzed by the enzyme (compare DCL-A and DCL-D). The acetate constituent (13a) was in this case also the best substrate for the enzyme. However, despite the similarity, the overall formation of acetate was in this case about 12 times slower for 13a (t½ ≈ 2500 min) compared to 12a (t½ = 210 min).

For propionyl and butyryl, the only other two acyl functionalities being hydrolyzed, even larger differences were recorded. The estimated half-times of formation are well beyond 4000 min for both substrates. In the libraries containing 12 (DCL-A, DCL-C), formation of acetate was around 20% faster than propionate, with butyrate forming only after substantial hydrolysis of the two main substrates (~ 600 min).

<table>
<thead>
<tr>
<th>Library</th>
<th>Thiol components</th>
<th>t½ (min)</th>
<th>Acetate</th>
<th>Propionate</th>
<th>Butyrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCL-C</td>
<td>1,12,13,16,19</td>
<td>260</td>
<td>310</td>
<td>&gt;1800</td>
<td></td>
</tr>
<tr>
<td>DCL-A</td>
<td>1,12</td>
<td>210</td>
<td>270</td>
<td>&gt;1500</td>
<td></td>
</tr>
<tr>
<td>DCL-D</td>
<td>1,13</td>
<td>~2500</td>
<td>&gt;4000</td>
<td>&gt;&gt;4000</td>
<td></td>
</tr>
<tr>
<td>DCL-E</td>
<td>1,16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DCL-F</td>
<td>1,19</td>
<td>~4000</td>
<td>&gt;4000</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Summary of the rates of formation of the different hydrolysis products in the five DCLs using transthiolesterification.

In the experiments containing 2-diethylaminoethanethiol (13, DCL-D), the difference between propionate and butyrate formation, compared to acetate formation, was around 40% and 55% slower, respectively, at 3000 min.

In contrast to both DCL-A and DCL-D, the DCL in which N-(methyl)mercaptoacetamide (19) was used together with 1a-1e (DCL-F), did not show any hydrolysis of the butyryl functionality. As with DCL-D, the t½ values recorded for the formation of acetate and propionate, respectively, were well beyond that of DCL-A. The library with mercaptoethanesulfonate (16, DCL-G) failed to give any hydrolysis products within the time measured (~ 4500 min), identical to the effects of thiol 1.
5.1.3 Catalytic self-screening using different hydrolases

To test the selectivity of the self-screening process, six other enzymes belonging to the hydrolase family were tested under the same set of conditions as in DCL-A. These enzymes were: butyrylcholinesterase (BChE, EC 3.1.1.8), horse liver esterase (HLE, EC 3.1.1.1), Candida cylindracea lipase (CCL, EC 3.1.1.3), β-galactosidase (β-Gal, EC 3.2.1.23), trypsin (EC 3.4.21.4), and subtilisin Carlsberg (Sub C, EC 3.4.21.62). The dynamic libraries were thus exposed to each of the enzymes and the formation of the hydrolysis products analyzed. The results are summarized in Table 4, recorded as percent product formation after 210 min (t½ for acetate in presence of AChE). All hydrolases acting on carboxylic ester bonds (EC 3.1.1.X) showed some hydrolysis, although the lipase from Candida cylindracea (CCL) only very modestly. In contrast to acetylcholinesterase, butyrylcholinesterase (BChE) acted on all acyl groups and hydrolyzed all groups in roughly the same time. This result is well in accordance with the known substrate pattern for this enzyme. The esterase from horse liver (HLE) show a pattern in which the longer acyl chains being slightly faster hydrolyzed than its shorter counterparts. For the two proteases trypsin and subtilisin, only the latter shows some activity under these conditions, also with some selectivity for the longer acyl chains. The hydrolase belonging to the glycosylases, β-galactosidase (β-Gal), did not show any activity. Control experiments with bovine serum albumin (BSA) also failed to give any hydrolysis products, as expected.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Butyrate</th>
<th>Valerate</th>
<th>Caproate</th>
</tr>
</thead>
<tbody>
<tr>
<td>AChE</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BChE</td>
<td>44</td>
<td>44</td>
<td>43</td>
</tr>
<tr>
<td>HLE</td>
<td>16</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>CCL</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td>&lt;5</td>
</tr>
<tr>
<td>β-Gal</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td>&lt;5</td>
</tr>
<tr>
<td>Trypsin</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sub C</td>
<td>&lt;5</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>BSA</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4. Acyl product formation for seven hydrolases and bovine serum albumin (BSA), with DCL-A.

The results clearly indicate that the catalytic self-screening process is efficient in identifying enzyme substrates. Of all the constituents formed in the DCLs, two of the thiolesters (12a, 12b) were easily recognized by the method as essentially being the best substrates for acetylcholinesterase, well in accordance to the recorded selectivity of the enzyme. The technique could however also distinguish constituents with considerably lower substrate activities, such as 12c and 13a. For acyl functionalities longer than butyryl, all constituents failed to give any hydrolysis products with acetylcholinesterase, and thiolesters having a negative charge at neutral pH were inefficient as substrates in all cases. The selectivity of the method was furthermore clearly demonstrated in self-screening with a range of different hydrolases. Under the same set of conditions, these hydrolases showed selective activities for the library constituents, differing from the performance of acetylcholinesterase. Butyrylcholinesterase and horse liver esterase both showed broader specificities than acetylcholinesterase, and horse liver esterase as well as subtilisin displayed more selectivity towards longer acyl groups.
5.2 Pre-equilibrated dynamic thiol/disulfide combinatorial libraries

The pre-equilibrated approach involves generating DCLs under reversible conditions, whereas the screening process is performed under static conditions. In this study, thiol/disulfide interchange was used as a way of generating reversible bonds. The target molecule used was Concanavalin A and the produced DCLs were evaluated using Quartz Crystal Microbalance (QCM)-technique.

5.2.1 Introduction

The synthesis of carbohydrate libraries is a challenging task (Schweizer and Hindsgaul 1999; St. Hilaire and Meldal 2000; Barkley and Arya 2001). In this respect, DCC offers a complementary route in forming libraries of dynamically interchanging carbohydrate species (Sakai, Shigemasa et al. 1999; Ramström and Lehn 2000; Ramström, Lohmann et al. 2004). In the present study, different thiosaccharides and other thiol-components were chosen as building blocks for the generation of the DCLs (Figure 13). Six different thiosaccharides (3–6, 24, 25), based on the common carbohydrate structures α-D-mannose, α-D-glucose, β-D-galactose, N-acetyl-β-D-glucosamine, N-acetyl-β-D-galactosamine and eight non-carbohydrate building blocks were tested (1, 13–15, 17–19, 26). These were chosen to contain different functionalities, where carboxylic (1, 17), hydroxyl (14, 15), amine (13, 26), amide (19) and ester (18) groups were probed for their potential interactions with the lectin binding site.

Figure 13. Thiol components used in DCL generation through thiol/disulfide interchange.
The dynamic combinatorial libraries were prepared by mixing the thiol components in neutral phosphate buffer. The interconversion of the library was initiated by oxidation with hydrogen peroxide, added in small aliquots. This process was found to be both mild and rapid, with DCL equilibration within hours and no overoxidation of the disulfide products. Once the libraries were fully oxidised to disulfides, the interconversion was essentially blocked and no further scrambling could be recorded.

All 14 thiol-components were used as components for the first library (DCL-G), giving rise to 105 different disulfide dimers in the resulting dynamic combinatorial library. In order to screen the library for individual component activity, a deconvolution strategy was employed (Bunyapaiboonsri, Ramström et al. 2001). Thus, in addition to the full DCL-G, based on all fourteen components and composed of 105 constituents, fourteen sublibraries were prepared in the same way where one of the components was substituted for buffer solution. These sublibraries were composed of 91 different constituents, respectively.

5.2.2 Screening and analysis of the DCLs using QCM biosensor

The libraries were screened in presence of excess calcium ions, since the dimers from 4, 24 and 25 interfere with the Ca\(^{2+}\) binding in Con A (Pei, Aastrup et al. 2005), resulting in a destroyed carbohydrate recognition domain. The screening protocol made use of QCM analysis, modified from a mannan-based protocol commonly used in Con A analysis (Pei, Anderson et al. 2004). QCM uses gold-plated quartz crystals that were first coated with polystyrene, and then coated with mannan while mounted online. Screening can subsequently be performed on the same crystal using competitive binding between the mannan and the tested ligands in the DCLs. Samples were prepared with Con A and the different DCLs (full DCL and sublibraries, respectively) and injected consecutively. By applying an electric current to the quartz crystal a frequency shift can be recorded. This frequency shift of the crystal can be altered by adding or removing molecules at the surface of the crystal. The frequency shift was recorded after each injection, corresponding to Con A binding to the surface. The bound Con A can be released from the mannan-surface by injection of buffer at lower pH.

All libraries were of similar high inhibitory activity, except for sublibraries A3 and A6, which both showed reduced activity (Figure 14). This indicates that these two components, 3 and 6, are important in constituents competing with mannan for binding to the lectin. Since none of the other sublibraries showed any diminished action, it was concluded that constituents based on combinations of 3 and/or 6 were more active.

![Figure 14. Dynamic deconvolution of glycosyl disulfide DCL-G. Concanavalin A binding to mannan in absence of DCLs (Ref), in presence of complete library A0, and sublibraries of compounds shown in Figure 13. Sublibraries A3 and A6 indicate strong dependence from compounds 3 and 6, respectively.](image-url)
Because of the non-carbohydrate components not showing any apparent activity, a focused dynamic combinatorial library (DCL-H) was subsequently synthesized. This library was composed of the thiosaccharide components 3-6, 24 and 25. These six components resulted in the generation of 21 different dimers using the same method as DCL-C. At the same time, six sublibraries were prepared from employing five of the components to form 15-member sublibraries. The QCM screening results (Figure 15), clearly indicates that components 3 and 6 are important in generating constituents able to compete with mannan for binding to Con A.

![Figure 15. Dynamic deconvolution of glycosyldisulfide DCL-H. Concanavalin A binding to mannan in absence of DCLs (Ref), in presence of complete library B0, and sublibraries B3-B6, B24 and B25. Sublibraries B3 and B6 indicate strong dependence from compounds 3 and 6 respectively.]

To further evaluate the active constituents of the libraries, the apparently efficient components 3 and 6, as well as their dimers: 3–3, 6–6, and 3–6 were tested. The mannoside dimers were synthesised from their parent monomers using the same method as for the library generation. The homodimers 3–3 and 6–6 could be directly used in individual testing after oxidation without further purification, while heterodimer 3–6 was purified by chromatography after oxidation.

<table>
<thead>
<tr>
<th>Compound</th>
<th>EC\textsubscript{50}/mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>&gt;5.0</td>
</tr>
<tr>
<td>6</td>
<td>&gt;&gt;20</td>
</tr>
<tr>
<td>3-3</td>
<td>1.2</td>
</tr>
<tr>
<td>6-6</td>
<td>&gt;20</td>
</tr>
<tr>
<td>3-6</td>
<td>1.2</td>
</tr>
</tbody>
</table>

*Table 5. Estimation of EC\textsubscript{50} values (50% inhibition of Concanavalin A binding) for tested carbohydrate structures.*

Competition assays using the QCM setup used in the library screening resulted in estimates of the individual compounds’ inhibitory performance. The resulting competition curves are presented in Figure 16, and the calculated EC\textsubscript{50} values (50% competition of ConA binding) are displayed in Table 5. Compounds 3–3 and 3–6 are the most active of the library constituents with EC\textsubscript{50} values of 1.2 mM. Compound 6–6 proved essentially inefficient with an EC\textsubscript{50} value exceeding 20 mM. These values are still relatively high compared to known inhibitors of ConA interactions, but in comparison to their parent components, the effects are dramatic. Compound 3, and especially compound 6, both showed lower inhibitory effects.
Both components 3 and 6 show much lower activity compared to methyl α-D-mannopyranoside (EC$_{50}$: 1.1 mM) (Pei, Aastrup et al. 2005). The exchange of a methoxy group for a thiol group at the anomeric position thus reduces the Con A binding activity, indicating that the sulfhydryl group is largely unstabilised by the Con A aglycon binding pocket. The activity is in addition slightly lower than that of D-mannose (EC$_{50}$: 5.3 mM, mixture of α/β anomers) (Pei, Anderson et al. 2004), however less significantly.

Component 6 proved much less active than its analogue methyl α-D-mannopyranoside, with an EC$_{50}$-value more than 20 times higher. Since the only difference between these two compounds lies in the exchange of an oxygen for a sulfur, and because SH is both larger and a poorer hydrogen-bond donor than OH, this result suggests that this hydroxyl group is involved in pronounced interactions with the Con A lectin in accordance with known binding patterns (Naismith and Field 1996).

As expected, the combination of two units of the inefficient component 6 results in an almost equally inactive dimer 6–6. On the other hand, compound 3–3 proved more active than its corresponding monomer 3. This result can be rationalised in two ways; either the binding enhancement is due to added interactions from each monomer unit while bound in the site, or it is due to a concentration effect arising from two equal units tied together. Since the increase in binding is only slightly higher than can be expected from the concentration, leading to a three-fold decrease in the EC$_{50}$-value, the concentration effect is likely the main reason. The results thus suggest that that both monomer units of the homoditopic constituent 3–3 bind to the same site.

Compound 3–6 displays the same activity as 3–3. In spite of the inefficiency of monomer 6, its combination with 3 yielded an efficient inhibitor. The same arguments as for compound 3–3 might be applied to the effect of constituent 3–6, which could arise from a concentration factor where each unit interacts with essentially the same site. However, in this case it is more probable that both units interact with different parts of the Con A binding site. Not only proved compound 6–6 to be inefficient, leading to the conclusion that the 6-unit is unable to bind to the 3-binding site, but compound 3–6 also mimics the natural trimannoside ligand, inasmuch as it resembles the α-D-Man-(1,6)-D-Man-part of the trimannoside.
6 Future prospects

Dynamic combinatorial chemistry is an interesting way of producing molecular diversity for use with biological systems. However, to advance the technique new methods and analysis protocols have to be explored. We have established the technique using a self-screening protocol in catalytic amount, catalytic self-screening, and are currently continuing to evaluate the scope of this concept. One new method that makes use of the catalytic self-screening concept has recently been developed. It was anticipated that using the thiolester, in which the carbonyl group and the sulfur atom switch places (31) compared to propionylthiocholine (12b), would generate 35 upon hydrolysis with AChE. Starting with the three thioesters 1b, 27 and 28 and the two thiols 12 and 29, this would generate, upon mixing, a library composed of nine thioesters (Scheme 11), all interchanging with all others at physiological pH.

Upon addition of the serine hydrolase acetylcholinesterase, the compounds that best fit into the active site of the enzyme are being hydrolyzed. By following the scrambling and analyzing the library by $^1$H-NMR, the two compounds that are being hydrolyzed, 34 and 35, are possible to distinguish. Hence, on addition of the enzyme, it is possible to screen the whole library without the necessity of making any additional time consuming sublibraries. By using this “combo-catalysis” it is possible to tell what moieties are necessary for hydrolysis by the enzyme.
7 Concluding remarks

New means of generating reversible covalent bond reactions using transthiolesterification have been explored. Further, two new screening/analysis methods not previously used, namely quartz crystal microbalance (QCM) and catalytic self-screening, have been probed.

- It has been demonstrated that transthiolesterification is a highly useful way to generate DCLs under mild conditions in aqueous media, and the resulting system can be subjected to self-screening in the presence of enzymes. The outcome of the self-screening system relying on the selectivity of the specific enzyme used. This approach also enables screening of complex DCLs without the necessity of using equimolar amounts of targets. Furthermore, the present study has demonstrated that catalysis can be used as a way to self-screen substrates from dynamic combinatorial libraries. It has been demonstrated that enzymes can be used as efficient catalysts for targeting and rapid identification of the best substrates formed in a dynamic combinatorial library. This system is, however, not restrained to enzyme catalysis; it may be extended to any catalytic system, including organic and inorganic catalysts, and may be employed to rapidly screen reactions of catalysts for new substrates.

- Highly compact dynamic combinatorial carbohydrate libraries was constructed and screened against the Con A lectin, based mainly on thiosaccharide units and thiol-disulfide interchange. These libraries could be efficiently screened using a QCM method by monitoring the competition of the libraries with Con A interactions to mannan. Two active dimeric carbohydrate structures could be identified from the screening process, resulting from combinations of 1-thio-α-D-mannopyranose and methyl 6-thio-α-D-mannopyranoside. The results indicate that the heterodimer 3–6 is active due to its resemblance to the natural trimannoside ligand for the lectin. The results also indicate that the hydroxyl group in the 6-position of methyl-α-D-mannopyranoside is crucial for efficient binding to Con A, compared to the analogous sulfhydryl group, whereas the anomeric hydroxyl in D-mannose is less important in this respect.

- New dynamic combinatorial chemistry formats are currently being developed within the group. We are presently working on, and continuing to, evaluate the scope of catalytic self-screening as a method of screening and analyzing dynamic combinatorial libraries.
Acknowledgements

I wish to express my sincere gratitude to:

Doc Olof Ramström for accepting me as a PhD-student, being an excellent supervisor and for your inspiring guidance and support throughout this study.

My co-workers in the Ramström group: Zhichao Pei, Hai Dong, Marcus Angelin, Pornrapee “Jom” Vongvilai, Dr Yuxin Pei and Rémi Caraballo for being good friends, for pleasant times in the lab and for interesting discussions about chemistry and life. A special thanks to Zhichao for showing me around and helping me get started back in the early days.

My co-authors: Zhichao Pei, Dr Teodor Aastrup, Henrik Andersson and Professor Jean-Marie Lehn.

Past and present post docs, Erasmus- and Master students for helpful assistance and for spreading a nice atmosphere throughout the group. A special thanks to Magnus Hermansson for helping me during my first time at the department.

Everyone at the chemistry department for a friendly and pleasant working atmosphere, for fruitful discussions and for helping out when needed.

Dr Ulrika Rosenström for skilful and inspiring guidance during my Master project at Uppsala University.

The Swedish Research Council (VR), the Swedish Foundation for International Cooperation in Research and Higher Graduation (STINT) and the Carl Trygger foundation for financial support.

Jag vill också tacka familj och vänner för att ni ställer upp och för att ni får mig att koppla bort kemin för ett ögonblick.

Mamma och pappa för att ni alltid stöttat mig och låtit mig gå mina egna vägar.

Kattis, den bästa syrra man kan ha!

Brigitte, för ditt stöd, din kärlek och för att du står ut med alla sena arbetsdagar.

Rikard
Appendix

The following is a description of my contributions to papers I-III.

Paper I: I performed a majority of the lab work and wrote parts of the article.

Paper II: I performed all labwork and wrote the article.

Paper III: I performed some of the lab work and wrote parts of the article.
André, S., Z. Pei, et al. (2006). "Glycosyldisulfides from dynamic combinatorial libraries as O-glycoside mimetics for plant and endogenous lectins: Their reactivities in solid-phase and cell assays and conformational analysis by molecular dynamics simulations." Submitted.


