Disulfide-based dynamic combinatorial libraries of macrocyclic pseudopeptides as bio-inspired complex chemical systems

Joan Atcher Ubiergo

DEPARTAMENT DE QUÍMICA ORGÀNICA

FACULTAT DE QUÍMICA
UNIVERSITAT DE BARCELONA

Programa de doctorat de Química Orgànica

Disulfide-based dynamic combinatorial libraries of macrocyclic pseudopeptides as bio-inspired complex chemical systems

Memòria presentada per **Joan Atcher Ubiergo**
per optar al títol de Doctor per la Universitat de Barcelona

Tesi realitzada al Departament de Química Biològica i Modelització Molecular de l'Institut de Química Avançada de Catalunya (IQAC-CSIC)

Director:

Dr. Ignacio Alfonso Rodríguez
Departament de Química Biològica i
Modelització Molecular (IQAC-CSIC)

Tutor:

Dr. Pedro Romea García
Departament de Química Orgànica
Facultat de Química (UB)

Doctorand:

Joan Atcher Ubiergo
Departament de Química Biològica i
Modelització Molecular (IQAC-CSIC)

ACKNOWLEDGMENTS

This doctoral thesis would not have been possible without the help, support, guidance and patience of many people. Here I would like to express my sincere gratitude to all of them.

Primerament m'agradaria donar les gràcies al meu director de tesi, el Dr. Ignacio Alfonso, per haver-me ajudat i guiat durant el transcurs d'aquest treball. Els seus consells i optimisme han estat molt valuosos. També voldria agrair-li el fet d'haver confiat en mi. Valoro positivament la llibertat que he tingut ja que m'ha permès aprendre de la millor manera: provant (i equivocant-me).

Gràcies a tots els companys del lab 309. Amb ells hem compartit un munt d'hores, cafès, problemes, discussions, dinars, seminaris, pastissos, cerveses, nits... El magnífic ambient viscut dins i fora del laboratori ha estat el millor aliat a l'hora d'afrontar els moments difícils.

També voldria expressar el meu agraïment a tothom qui ha contribuït amb els experiments i càlculs d'aquesta tesi. En particular m'agradaria donar les gràcies al Dr. Jordi Bujons per la seva col·laboració amb els estudis computacionals i a la Dra. Yolanda Pérez pel seu ajut amb els experiments de RMN.

I would like to thank Prof. Sijbren Otto (Stratingh Institute for Chemistry, University of Groningen) for giving me the opportunity to work in his research group (Feb.-Apr. 2014). It was really encouraging for me to work in one of the world's leading research groups in the field of DCC. I would also like to thank all the members of the Otto group for making my short stay in the Netherlands unforgettable.

Finalment voldria donar les gràcies a la meva família i a "aquella gent" per haver-me fet costat durant tot aquest temps. Sense el seu suport anímic incondicional aquest treball no hauria estat possible.

This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO, CTQ2009-14366-C02-02 and CTQ2012-38543-C03-03 projects), la Generalitat de Catalunya (AGAUR, 2014SGR231), and the European Union (COST action CM1304). I am also grateful to CSIC and the European Social Fund for personal financial support (JAE-predoc fellowship).

ABBREVIATIONS

AA	amino acid
AF	amplification factor
AF_n	normalized amplification factor
Alloc	allyloxycarbonyl
BB	building block
bis-Tris	bis(2-hydroxyethyl)aminotris(hydroxymethyl)methane
Boc	*tert*-butyloxycarbonyl
Cbz	benzyloxycarbonyl
CD	circular dichroism
DCC	dynamic combinatorial chemistry
DCCD	*N,N′*-dicyclohexylcarbodiimide
DCL	dynamic combinatorial library
DCM	dichloromethane
DCvC	dynamic covalent chemistry
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
DN	dialkoxynaphthalene
DNA	deoxyribonucleic acid
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTNB	5,5'-dithiobis(2-nitrobenzoic acid)
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	1,2-ethanedithiol
ESI	electrospray ionization
FIA	flow injection analysis
Fmoc	fluorenylmethyloxycarbonyl
GAP	gemini amphiphilic pseudopeptide
Gnd	guanidine
HBTU	*o*-(benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate
HDS	hydrodesulfurization
HIV	human immunodeficiency virus
HOBt	1-hydroxybenzotriazole

HRMS	high resolution mass spectrometry
LC	liquid chromatography
MD	molecular dynamics
MS	mass spectrometry
NCL	native chemical ligation
NDI	naphthalene diimide
NMR	nuclear magnetic resonance
PG	protecting group
PNA	peptide nucleic acid
QD	quantum dot
RD	relative difference
RMSD	root mean square deviation
RNA	ribonucleic acid
RP-HPLC	reversed phase high performance liquid chromatography
TCEP	tris(2-carboxyethyl)phosphine
Tf	trifluoromethanesulfonate
TFA	trifluoroacetic acid
TIS	triisobutylsilane
TLC	thin layer chromatography
TOF	time-of-flight
Trt	triphenylmethyl
UPLC	ultra performance liquid chromatography
UV-Vis	ultraviolet-visible

TABLE OF CONTENTS

GENERAL INTRODUCTION

Supramolecular chemistry

Supramolecular chemistry[1-2] is defined[3-4] as the chemistry that goes beyond the covalent bond and beyond the individual molecule. Thus, it focuses on the study of non-covalent interactions in and between molecules, and the resulting multimolecular complexes, *i.e.* the supramolecular species. These non-covalent interactions, including electrostatic effects, metal coordination, H-bonding, π-π stacking, van der Waals forces, solvophobic effects and halogen bonding, provide molecular recognition and self-organization. Thereby, supramolecular chemistry plays a key role in a variety of fields such as materials science and catalysis, and is fundamental for the understanding of many biological processes.[5]

The nature of supramolecular chemistry is dynamic due to the lability of the non-covalent interactions connecting the molecular components. The resulting ability of supramolecular species to reversibly dissociate and associate allows them to rearrange their building blocks. Similarly, the "molecular chemistry", *i.e.* the chemistry of the covalent bond, can also exhibit dynamic features by virtue of reversible unions. In this regard, dynamic combinatorial chemistry exploits the dynamic nature of both supramolecular and covalent linkages for generating complexity within the field of systems chemistry.

Systems chemistry

Nature is composed of intrinsically complex networks. A network is an interconnected group or system, and consists of a series of nodes, which represent single entities, such as living species, people, computers or molecules. The properties of these nodes, together with their connections, *i.e.* the components of the system and the relationship between them, define the structure and behavior of the whole network. Examples can be found in biological ecosystems,[6-7] sociology[8] and economic markets.[9-10] However, parallels in chemistry are still scarce,[11] despite all the natural world can be ultimately reduced to the chemical level. Commonly, chemists have been guided by a methodological reductionism in which complex dynamic mixtures have been deliberately avoided. This was driven by necessity as, until recently, few analytical tools existed that allowed to study the interplay between molecules in complex mixtures. This situation has now changed and recently systems chemistry has irrupted as an emerging discipline at the interphase of chemical, physical and biological sciences.

The term "systems chemistry" was originally coined by von Kiedrowski in 2005 when describing the behavior of a nearly exponential organic replicator.[12] In the same year, a more detailed definition was given describing the systems chemistry as a "new field of chemistry seen as the offspring of prebiotic and supramolecular chemistry on the one hand, and theoretical biology and complex systems research on the other".[13] Somehow, systems chemistry attempts to capture the complexity and emergent phenomena prevalent in other disciplines like biology. In the systems chemistry approach, a group of synthetic chemical entities is designed to interact and react with many partners within a network. In this manner, the synthetic chemical system can potentially show interesting emerging properties, *i.e.* properties that are not simply the linear sum of the attributes of the individual components.

"Complexity" is a concept highly dependent on the context, and chemistry has its own understanding of it. Whitesides in 1999 defined a complex chemical system as "the one whose evolution is very sensitive to initial conditions or to small perturbations, one in which the number of independent interacting components is large, or one in which there are multiple pathways by which the system can evolve".[14] More recently, in 2013 Lehn defined the complexity within the frame of dynamic chemistry as the result from a combination of three features: multiplicity (chemical diversity in constitution and function), interconnection (non-covalent and covalent interactions, as well as their dynamics), and integration (combination of all features through networks with feedback and regulation).[11]

Although systems chemistry is a quite recent term, it is not a totally new field of chemistry but a title under which a few already quite well established research areas can cooperate. In this regard, dynamic combinatorial chemistry has been central to developing complex systems.[15-17] Other areas of development include self-replication (section 3.1.1.1), prebiotic chemistry,[18] molecular logics[19-21] and oscillating reactions,[22-23] among others.

Dynamic combinatorial chemistry

Dynamic combinatorial chemistry (DCC)[24-26] can be defined as combinatorial chemistry where the library members interconvert continuously by exchanging building blocks with each other. The objects of investigation of DCC are dynamic combinatorial libraries (DCLs), which are currently considered a starting point for the development of

systems chemistry. DCLs are formed upon mixing building blocks which can form reversible covalent or non-covalent interactions with one another to generate a mixture of oligomeric species. The distribution of all molecules in such a network is typically, but not necessarily,[a,15] governed by thermodynamics. Changing the experimental conditions may alter the stability of the library members and thereby alter the composition of the entire molecular network. This introduces a characteristic property, which is the possibility of the chemical system to adapt to changes in its environment according to Le Châtelier's principle, thus enabling adaptive chemistry.[27]

Another commonly used term, constitutional dynamic chemistry (CDC),[27-29] was defined by Lehn as a synonym of DCC. According to him, "DCC denomination stresses the combinatorial character, whereas CDC highlights the basic feature of chemical entities: their constitution".[27]

DCC was pioneered by the groups of Sanders and Lehn in the mid-1990s.[b,30] The basic principles arose from the realization that the task of constructing effective and specific molecular receptors is very difficult to address using a straightforward synthetic design approach. A new, more efficient, and general approach was conceived, encompassing combinatorial, selection, proofreading, self-correction and amplification elements. This new methodology proposes the creation of a library of compounds that continuously interchange their subunits, resulting in a dynamic mixture of products, which is thermodynamically controlled. The template molecule (guest) would be introduced into this library, allowing it to select its own most effective binder(s) from among the mixture of possible (virtual)[c,31] hosts. The member(s) of the library most efficiently interacting with the template would be selected and amplified at the expense of the other members,[32] minimizing the whole energy of the system (Figure I.1). After freezing the chemical equilibrium, the detection, identification and isolation of the best host molecule(s) would be possible.

[a] DCC was initially defined as combinatorial chemistry under thermodynamic control. Recently, however, extensions of the principles of DCC to systems that are not at equilibrium have been explored. Examples include self-replication and molecular machines.

[b] The first paper clearly articulating the concept of DCC was published by the group of Sanders and appeared in 1996.

[c] According to Lehn, all combinations of interchanging compounds are possible even if some are not present in the pool, making the assembly "virtual".

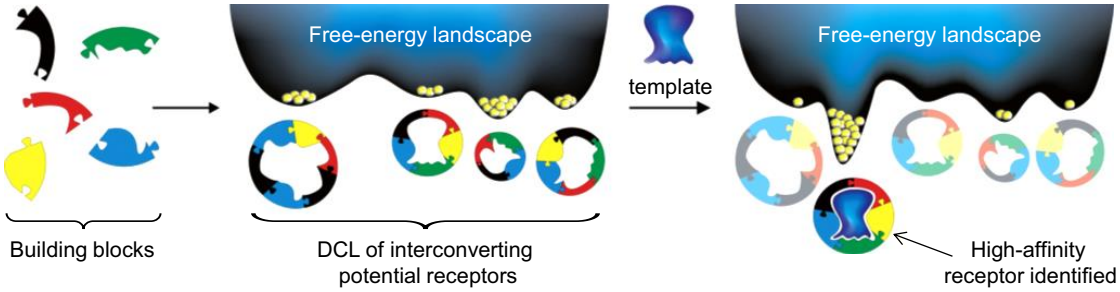


Figure I.1. Schematic representation of a DCL and its free energy landscape, showing how the introduction of a template can lead to the amplification of the receptor with the highest affinity for the guest (Figure modified from reference[33]).

Thus, the first application that stood behind the development of DCC was the identification and preparation of synthetic molecular receptors.[33] Soon after, dynamic combinatorial approaches were implemented in the synthesis of ligands for biomolecules,[34-36] molecular cages[37-38] and sensors;[39-40] and DCC started to have an impact on some adjacent areas like catalysis,[41-42] multiphase systems,[43-44] surface chemistry,[45-46] dynamic combinatorial materials[47-48] and interlocked structures.[49-52] In nearly all of these examples the DCLs are under thermodynamic control. More recently, however, DCC is expanding into the rich field of out-of-equilibrium systems, including self-replication[53] and molecular machines.[54-57]

Dynamic covalent chemistry

As already stated, in DCC the bonds formed and broken can be both of covalent and non-covalent nature. Dynamic covalent chemistry (DCvC)[d,58-60] is the term often used to describe the first case. Thus, in contrast with the weak non-covalent interactions of supramolecular chemistry, DCvC deals with more robust covalent bonds with generally slower kinetics of bond formation and cleavage. Supramolecular chemistry is dynamic in the domain beyond molecules through intermolecular non-covalent interactions, whereas dynamic covalent chemistry is dynamic within molecules through reversible formation and breaking of covalent bonds.

Several factors have to be considered for evaluating whether a reversible covalent reaction is suitable for DCvC or not.[59] Firstly, the reversible linkage must fulfill certain thermodynamic and kinetic requirements. From the thermodynamic point of view, the bonds should be stable enough to hold molecular structures detectable and even

[d] DCC and DCvC have been oftentimes mistakenly considered to be synonymous, although the first concerns both dynamic covalent chemistry and dynamic non-covalent (supramolecular) chemistry.

isolable, yet still have a dynamic behavior. From the kinetic point of view, generally short equilibration times are preferred for practical purposes. Secondly, mild reaction conditions that are compatible with a broad range of functional groups present in building blocks and templates are preferred. Bio-compatible aqueous reaction conditions are highly desired for the design of bio-inspired DCLs. Finally, the exchange process should be easily turned off.[e] The possibility of "freezing" the exchange allows the analysis, separation and handling of the library members without fear of their decomposition. Satisfying these requirements, there are many reversible reactions suitable for DCvC.[25-26,59] Some of the most common ones are listed below in Figure I.2.

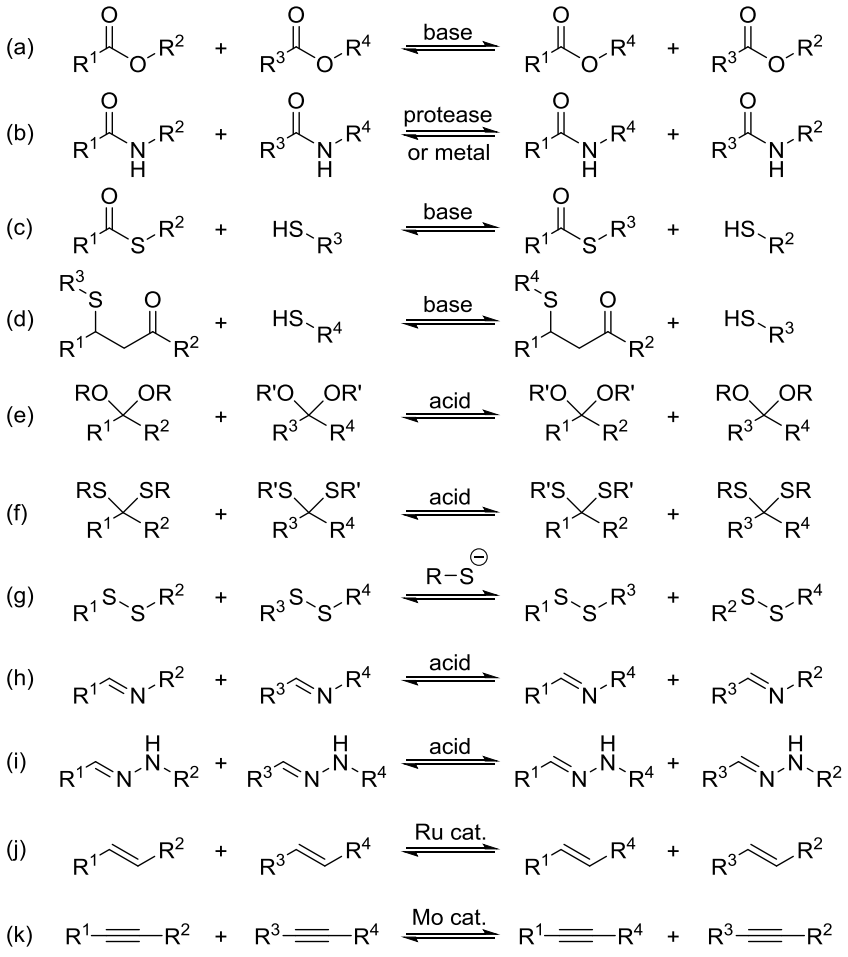


Figure I.2. Most common reversible reactions used in DCvC. Examples of acyl transfer: (a) transesterification, (b) transamidation[f,61] (c) transthioesterification and (d) Michael/retro-Michael reaction. Acetal exchange and related: (e) acetal exchange and (f) thioacetal exchange. Example of nucleophilic substitution: (g) disulfide exchange. Examples of C=N exchange: (h) imine exchange and (i) hydrazone exchange. Other reversible covalent bonds: (j) alkene metathesis and (k) alkyne metathesis.

[e] Common methods to stop the exchange process include temperature control, pH control, removal of catalyst and kinetic trapping through oxidation/reduction.

[f] Very recently, a promising new methodology based on native chemical ligation (NCL) has been proposed as a mild reaction to prepare dynamic covalent peptides at physiological conditions.

An important property of the reversible bonds used in DCvC is their symmetry. Some of the dynamic bonds are directional and unsymmetrical, and require the combination of two different components (*e.g.* imine bonds made from condensation of aldehydes and amines, Figures I.2h and I.3a). On the other hand, symmetrical bonds allow self-exchange (*e.g.* disulfide bonds, Figures I.2g and I.3b). Thus, the symmetry of the bond dictates whether the building blocks will form alternated sequences or self-oligomerize. The unsymmetrical bonds offer greater control over the structure of the library members, whereas the symmetrical bonds generate dynamic libraries of larger diversity.

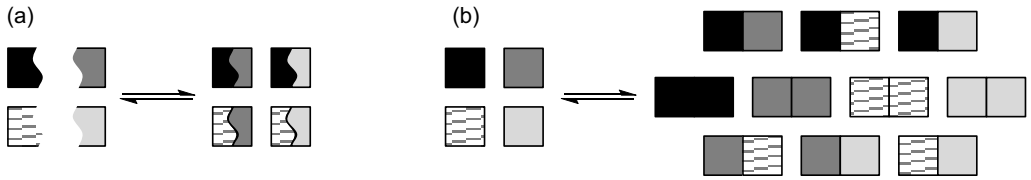


Figure I.3. Schematic representation of DCLs created using (a) an unsymmetrical reversible bond, and (b) a symmetrical reversible bond.

Context and perspectives

The present doctoral thesis has been carried out in the context of the research group led by Dr. Ignacio Alfonso, in the Department of Biological Chemistry and Molecular Modeling (IQAC-CSIC). The group has expertise in the use of peptidic and pseudopeptidic compounds for different applications in supramolecular chemistry and DCvC, and this work represent the starting point of a new research line dedicated to the study of disulfide-based dynamic combinatorial libraries.

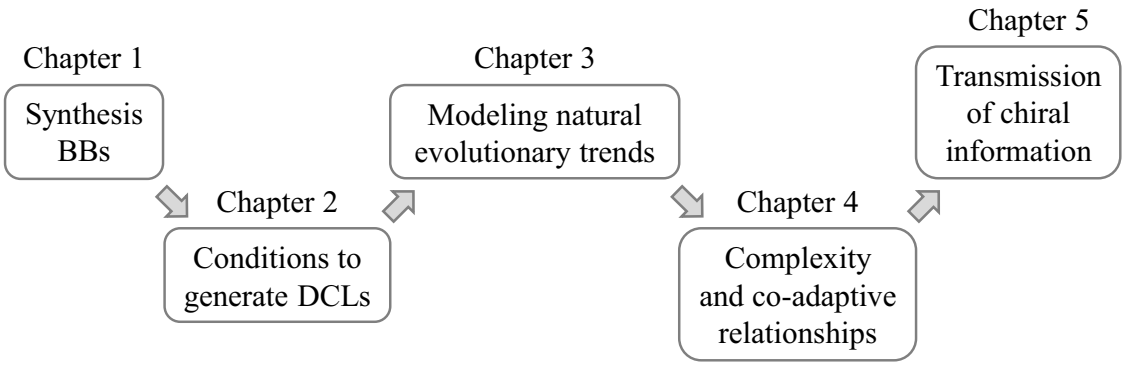
As evidenced in the following general objectives, the main idea and motivation behind this thesis is the use of dynamic combinatorial libraries as adaptive complex chemical systems for the minimalistic experimental modeling of different processes of biological interest. With this purpose, a bio-inspired design of the dynamic libraries, incorporating tailored structural information, together with a suitable external stimulus, are envisioned to allow for the emergence of natural evolutionary trends, co-adaptive relationships and even chiral self-sorting phenomena.

References

(1) Steed, J. W.; Atwood, J. L. *Supramolecular chemistry*; John Wiley & Sons, 2013.
(2) Beer, P. D.; Gale, P. A.; Smith, D. K. *Supramolecular chemistry*; Oxford University Press, 1999.
(3) Lehn, J.-M. *Angew. Chem. Int. Ed.* **1988**, *27*, 89.
(4) Lehn, J.-M. *Science* **1985**, *227*, 849.
(5) Uhlenheuer, D. A.; Petkau, K.; Brunsveld, L. *Chem. Soc. Rev.* **2010**, *39*, 2817.
(6) Montoya, J. M.; Pimm, S. L.; Solé, R. V. *Nature* **2006**, *442*, 259.
(7) Azaele, S.; Pigolotti, S.; Banavar, J. R.; Maritan, A. *Nature* **2006**, *444*, 926.
(8) Borgatti, S. P.; Mehra, A.; Brass, D. J.; Labianca, G. *Science* **2009**, *323*, 892.
(9) Schweitzer, F.; Fagiolo, G.; Sornette, D.; Vega-Redondo, F.; Vespignani, A.; White, D. R. *Science* **2009**, *325*, 422.
(10) Economides, N. *Int. J. Ind. Organ.* **1996**, *14*, 673.
(11) Lehn, J.-M. *Angew. Chem. Int. Ed.* **2013**, *52*, 2836.
(12) Kindermann, M.; Stahl, I.; Reimold, M.; Pankau, W. M.; von Kiedrowski, G. *Angew. Chem. Int. Ed.* **2005**, *44*, 6750.
(13) Stankiewicz, J.; Eckardt, L. H. *Angew. Chem. Int. Ed.* **2006**, *45*, 342.
(14) Whitesides, G. M.; Ismagilov, R. F. *Science* **1999**, *284*, 89.
(15) Li, J.; Nowak, P.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 9222.
(16) Hunt, R. A. R.; Otto, S. *Chem. Commun.* **2011**, *47*, 847.
(17) Ludlow, R. F.; Otto, S. *Chem. Soc. Rev.* **2008**, *37*, 101.
(18) Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. *Chem. Rev.* **2013**, *114*, 285.
(19) Wagner, N.; Ashkenasy, G. *Chem. Eur. J.* **2009**, *15*, 1765.
(20) de Silva, A. P.; Uchiyama, S. *Nat. Nanotechnol.* **2007**, *2*, 399.
(21) Collier, C. P.; Wong, E. W.; Belohradský, M.; Raymo, F. M.; Stoddart, J. F.; Kuekes, P. J.; Williams, R. S.; Heath, J. R. *Science* **1999**, *285*, 391.
(22) Epstein, I. R.; Pojman, J. A.; Steinbock, O. *Chaos* **2006**, *16*, 037101.
(23) Ferino, I.; Rombi, E. *Catal. Today* **1999**, *52*, 291.
(24) Cougnon, F. B. L.; Sanders, J. K. M. *Acc. Chem. Res.* **2011**, *45*, 2211.
(25) Reek, J. N.; Otto, S. *Dynamic combinatorial chemistry*; John Wiley & Sons, 2010.
(26) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652.
(27) Lehn, J.-M. *Chem. Soc. Rev.* **2007**, *36*, 151.
(28) Aastrup, T.; Barboiu, M. *Constitutional dynamic chemistry*; Springer Science & Business Media, 2012; Vol. 322.
(29) Hu, L.; Schaufelberger, F.; Timmer, B. J. J.; Flos, M. A.; Ramström, O. In *Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons, Inc.: 2000.
(30) Brady, P. A.; Bonar-Law, R. P.; Rowan, S. J.; Suckling, C. J.; Sanders, J. K. M. *Chem. Commun.* **1996**, 319.
(31) Lehn, J.-M. *Chem. Eur. J.* **1999**, *5*, 2455.
(32) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Chem. Eur. J.* **2008**, *14*, 2153.
(33) Otto, S.; Severin, K. In *Creative Chemical Sensor Systems*; Schrader, T., Ed.; Springer Berlin Heidelberg: 2007; Vol. 277, p 267.
(34) Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H. *Angew. Chem. Int. Ed.* **2014**, *53*, 3259.
(35) Herrmann, A. *Chem. Soc. Rev.* **2014**, *43*, 1899.
(36) López-Senín, P.; Gómez-Pinto, I.; Grandas, A.; Marchán, V. *Chem. Eur. J.* **2011**, *17*, 1946.

(37) Faggi, E.; Moure, A.; Bolte, M.; Vicent, C.; Luis, S. V.; Alfonso, I. *J. Org. Chem.* **2014**, *79*, 4590.
(38) Riddell, I. A.; Smulders, M. M. J.; Clegg, J. K.; Hristova, Y. R.; Breiner, B.; Thoburn, J. D.; Nitschke, J. R. *Nat. Chem.* **2012**, *4*, 751.
(39) You, L.; Berman, J. S.; Anslyn, E. V. *Nat. Chem.* **2011**, *3*, 943.
(40) Buryak, A.; Severin, K. *Angew. Chem. Int. Ed.* **2005**, *44*, 7935.
(41) Fanlo-Virgós, H.; Alba, A.-N. R.; Hamieh, S.; Colomb-Delsuc, M.; Otto, S. *Angew. Chem. Int. Ed.* **2014**, *53*, 11346.
(42) Dydio, P.; Breuil, P.-A. R.; Reek, J. N. H. *Isr. J. Chem.* **2013**, *53*, 61.
(43) Hafezi, N.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 12861.
(44) Perez-Fernandez, R.; Pittelkow, M.; Belenguer, A. M.; Lane, L. A.; Robinson, C. V.; Sanders, J. K. M. *Chem. Commun.* **2009**, 3708.
(45) Tauk, L.; Schröder, A. P.; Decher, G.; Giuseppone, N. *Nat. Chem.* **2009**, *1*, 649.
(46) Chang, T.; Rozkiewicz, D. I.; Ravoo, B. J.; Meijer, E. W.; Reinhoudt, D. N. *Nano Lett.* **2007**, *7*, 978.
(47) Lehn, J.-M. *Angew. Chem. Int. Ed.* **2015**, *54*, 3276.
(48) Moulin, E.; Cormos, G.; Giuseppone, N. *Chem. Soc. Rev.* **2012**, *41*, 1031.
(49) Ponnuswamy, N.; Cougnon, F. B. L.; Clough, J. M.; Pantoş, G. D.; Sanders, J. K. M. *Science* **2012**, *338*, 783.
(50) Cougnon, F. B. L.; Jenkins, N. A.; Pantoş, G. D.; Sanders, J. K. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 1443.
(51) Belowich, M. E.; Valente, C.; Smaldone, R. A.; Friedman, D. C.; Thiel, J.; Cronin, L.; Stoddart, J. F. *J. Am. Chem. Soc.* **2012**, *134*, 5243.
(52) Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. *Science* **2005**, *308*, 667.
(53) Otto, S. *Acc. Chem. Res.* **2012**, *45*, 2200.
(54) Kovaříček, P.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 9446.
(55) Campaña, A. G.; Carlone, A.; Chen, K.; Dryden, D. T. F.; Leigh, D. A.; Lewandowska, U.; Mullen, K. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 5480.
(56) von Delius, M.; Geertsema, E. M.; Leigh, D. A. *Nat. Chem.* **2010**, *2*, 96.
(57) Campbell, V. E.; de Hatten, X.; Delsuc, N.; Kauffmann, B.; Huc, I.; Nitschke, J. R. *Nat. Chem.* **2010**, *2*, 684.
(58) Jin, Y.; Wang, Q.; Taynton, P.; Zhang, W. *Acc. Chem. Res.* **2014**, *47*, 1575.
(59) Jin, Y.; Yu, C.; Denman, R. J.; Zhang, W. *Chem. Soc. Rev.* **2013**, *42*, 6634.
(60) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2002**, *41*, 898.
(61) Ruff, Y.; Garavini, V.; Giuseppone, N. *J. Am. Chem. Soc.* **2014**, *136*, 6333.

GENERAL OBJECTIVES



The general objectives of the present thesis were established covering the sequence of topics and challenges schematically listed above. The sequence starts with the setup of the building blocks and experimental conditions needed to generate useful dynamic combinatorial libraries, and follows with the use of these libraries as bio-inspired complex chemical systems.

The specific general objectives are:

1) To design, synthesize and characterize a set of new dithiol building blocks with peptide-like structural information.

2) To find suitable experimental conditions in which these building blocks can reversibly combine generating useful dynamic combinatorial libraries under thermodynamic control.

3) To explore the potential ability of a minimalistic dynamic combinatorial library to show adaptive trends in parallel with the evolution of biological systems.

4) To study the adaptive process of a large and diverse dynamic library with many co-adaptive relationships operating simultaneously in the same molecular network.

5) To investigate how chiral information is transmitted in dynamic combinatorial libraries leading to chiral self-sorting processes.

CHAPTER 1

Design, synthesis and characterization of the building blocks

1.1. Introduction

1.1.1. Precedents of C_2-symmetric minimalistic pseudopeptides

Life is based on very complex chemical systems composed by a limited number of structural units.[1] Among these basic elements, amino acids (AAs) are the building blocks of a large number of different biomolecules. The combination of only twenty AAs forming long polymeric chains allows building of an enormous number of peptides and proteins with different biological functions: enzyme catalysis, defense, transport, support, motion and regulation.[2] Fascinated and inspired by the astonishing structural and functional properties of these biomolecules, chemists have developed abiotic peptides mimicking properties of natural systems. In this regard, the building of small model compounds has been used for the understanding of the structural parameters and factors determining the properties of interest in natural peptides and proteins.[3]

A recurrent strategy for constructing bio-inspired minimalistic molecules is to combine natural building blocks that provide functional elements with abiotic fragments that serve as structural scaffolds. Pseudopeptides are molecules incorporating natural amino acids and non-natural structural parts, and have widely been used for the development of compounds with many important applications. Examples are reported in which pseudopeptidic backbones are used for preparing structured and stimuli-responsive materials,[4-7] peptide nucleic acids (PNAs),[8-9] catalysts,[10-12] and bio-active compounds.[13-15] Additionally, pseudopeptides have also been used for the development of molecular receptors for a variety of templates including anions,[16-19] inorganic cations[20-22] and peptides.[23-24]

Some of the simplest pseudopeptides reported in the literature involve two amino acids attached to a central diamino spacer (Figure 1.1).[25] Despite their simplicity, the possibility of using different spacers, different amino acids (R) and different groups attached to the *N*-terminus (Y), provides a high level of potential molecular diversity and recognition sites for supramolecular interactions. The NH groups of the pseudopeptidic backbone allow for H-bonding interactions, while the spacer, the side chains and the groups attached to the *N*-terminus can be involved in hydrophobic and π-π interactions. Very importantly, the C_2-simmetry facilitates the convergence[26] of

functional groups and makes easier the preparation of these pseudopeptidic molecules by reducing the number of synthetic steps.[a,27]

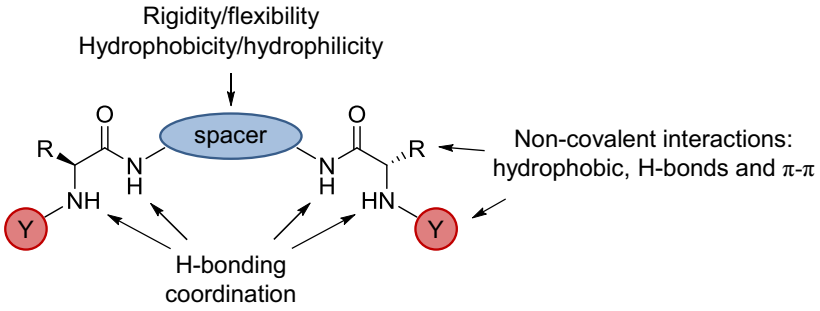


Figure 1.1. Schematic representation of a minimalistic C_2-symmetric pseudopaptide.

Minimalistic pseudopeptides with the general molecular structure represented in Figure 1.1 are designed and synthesized within the present thesis, and have previously been used for many different purposes. The versatility of this molecular architecture is illustrated in the following sections, where reported examples of open-chain and macrocyclic molecules are briefly commented.

1.1.1.1. Open-chain C_2-symmetric minimalistic pseudopeptides

A remarkable example of an open-chain C_2-symmetric pseudopeptide with the general structure depicted in Figure 1.1 is the gemini amphiphilic pseudopeptide (GAP) represented in Figure 1.2a.[28] The structure of this molecule was conceived as a "second generation" design after previous studies[5,29] in which a family of compounds with a similar molecular skeleton was found to be able to self-assemble into different nanostructures through the cooperative action of polar (H-bonding and dipole-dipole) and non-polar (van der Waals) intermolecular interactions. The amphiphilic compound consists of a flexible and hydrophilic central diethylenetriamine spacer[b] that joins two valine moieties with a long alkoxyphenyl hydrophobic tail attached to the *N*-terminus. In pure water, this compound self-assembles into vesicles at both acidic and neutral pH. Their formation occurs through the assembly of a folded conformation of the pseudopeptide in a gemini surfactant-like fashion[30] (Figure 1.2b). When dissolved in methanol or ethanol, the addition of water spontaneously causes the formation of stable hydrogels, which can be disassembled by heating or protonation. The presence of an

[a] Both identical arms of the molecule are generally simultaneously grown (*e.g.* section 1.3.2).
[b] The nitrogen atom in the middle of the central spacer is the responsible for its water solubility and is the main improvement compared with the "first generation" GAPs developed by the same authors.

organic co-solvent maintains the molecule in an extended conformation where the hydrophobic and H-bonding intermolecular interactions can operate in a preferred direction, leading to the formation of nanofibers upon self-assembly (Figure 1.2b).

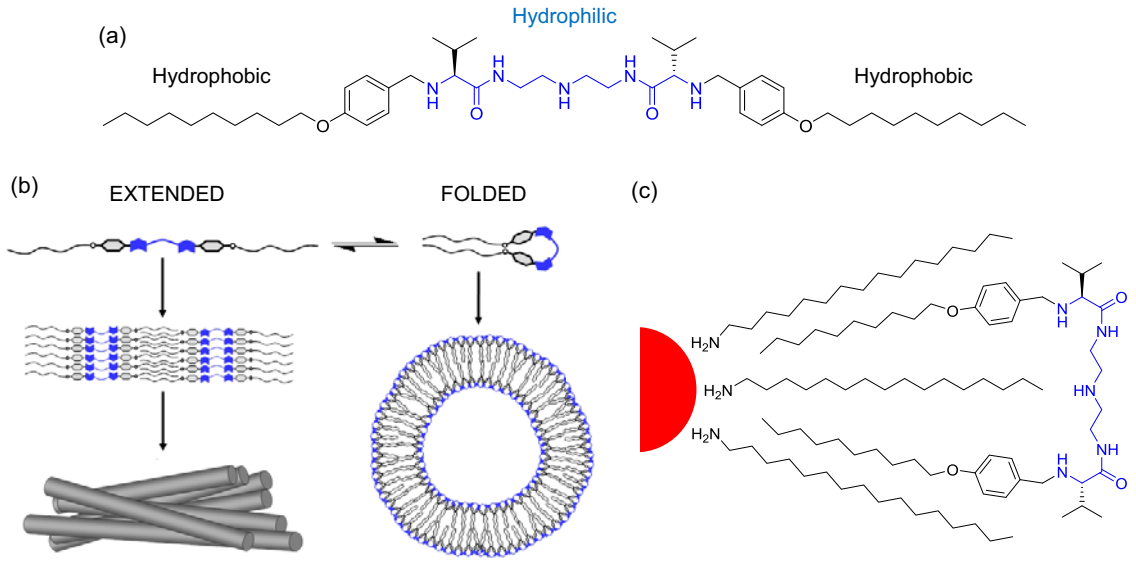


Figure 1.2. (a) Example of an open-chain C_2-symmetric minimalistic pseudopeptide and representation of (b) the proposed assembly of extended or folded conformations into fibers and vesicles, and (c) the interdigitation occurring between the hydrophobic ligands of the QDs and the alkyl chains of the pseudopeptidic compound (Figure modified from references[28,31]).

Additionally, within the context of a different investigation,[31] the same molecule proved to be able to stabilize hydrophobic quantum dots (QDs) in water. The amphiphile acts as an intercalator with hydrophobic ligands of CdSe/ZnS core-shell QDs (Figure 1.2c), transferring them from toluene to pure water.

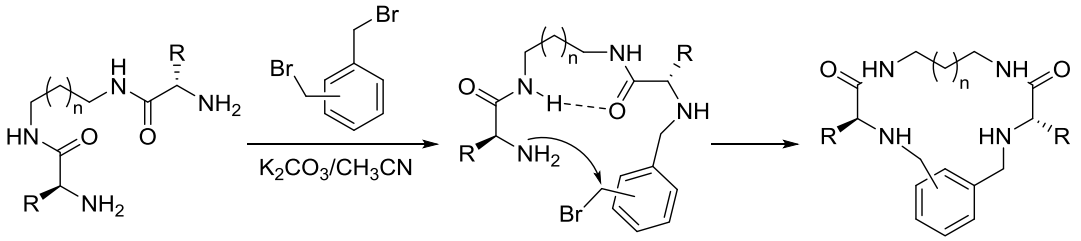
To summarize, the amphiphilic character of the molecule is the responsible for its tunable self-assembling behavior, as well as for its ability to transport hydrophobic QDs from non-polar solvents to water. This example perfectly illustrates that, by careful selection of the spacer, the amino acids and the groups attached to the *N*-termini; simple open-chain pseudopeptides with the general structure shown in Figure 1.1 can be designed with tailored properties, eventually leading to interesting and useful applications.

1.1.1.2. Macrocyclic C_2-symmetric minimalistic pseudopeptides

Chemists have been fascinated for many years by the inherent properties of macrocyclic molecules. Their decreased conformational freedom can provide a significant level of organization, combining high functional density and directionality.

Among all macrocyclic molecules, the ones containing amino acids are especially important compounds due to their interesting applications in molecular recognition,[32-34] biomedicine,[35-38] and materials science.[39-42] Therefore, the preparation of peptidic and pseudopeptidic macrocycles is a synthetic challenge of particular interest.[43-44] The synthesis of macrocyclic species in most cases is hampered by the macrocyclization step, which usually requires high dilution techniques, sophisticated protecting groups or tedious purification processes. One possible way to improve that process consists in favoring the approach of both ends of the linear precursor by means of preorganization mechanisms.[45] The following examples briefly comment on three preorganization mechanisms described for the synthesis of pseudopeptidic macrocycles with the general structure depicted in Figure 1.1.

The first mechanism consists in a *conformational preorganization*. In the example shown in Scheme 1.1,[27] the reaction of open-chain pseudopeptides derived from α,ω-alkyldiamines with meta- or para-bis(bromomethyl)benzene afforded the corresponding [1 + 1] macrocycles in 60-70% yields. The key step for the efficiency of the cyclization is the intramolecular reaction, which is promoted by a U-turn conformation favored by intramolecular H-bonding and solvophobic effects.



Scheme 1.1. Example of *conformational preorganization*. Reaction (S_N2) of open-chain pseudopeptides derived from α,ω-alkyldiamines with meta- or para-bis(bromomethyl)benzene to produce the corresponding [1 + 1] macrocycles (Figure modified from reference[27]).

The second mechanism is based on a *configurational preorganization* and requires the combination of chiral centers of appropriate configuration. In the example shown in Figure 1.3a,[46-47] a large-ring pseudopeptidic macrocycle is synthesized through a multicomponent [2 + 2] reaction, within a dynamic covalent mixture of imines. The process is a one-pot two-step cyclization based on a reductive amination reaction, and was found to be entirely governed by the structural information contained in the corresponding open-chain bis(amidoamine). A remarkable match/mismatch relationship between the configuration of the chiral centers of the cyclic spacer ((*R*,*R*)-cyclohexane-

1,2-diamine) and those of the pseudopeptidic frame (L- or D-amino acids) was observed (Figure 1.3b). The matched combination of configurations at the chiral centers afforded very good results (55-70% yields), whereas the mismatched combination hampered the cyclization.

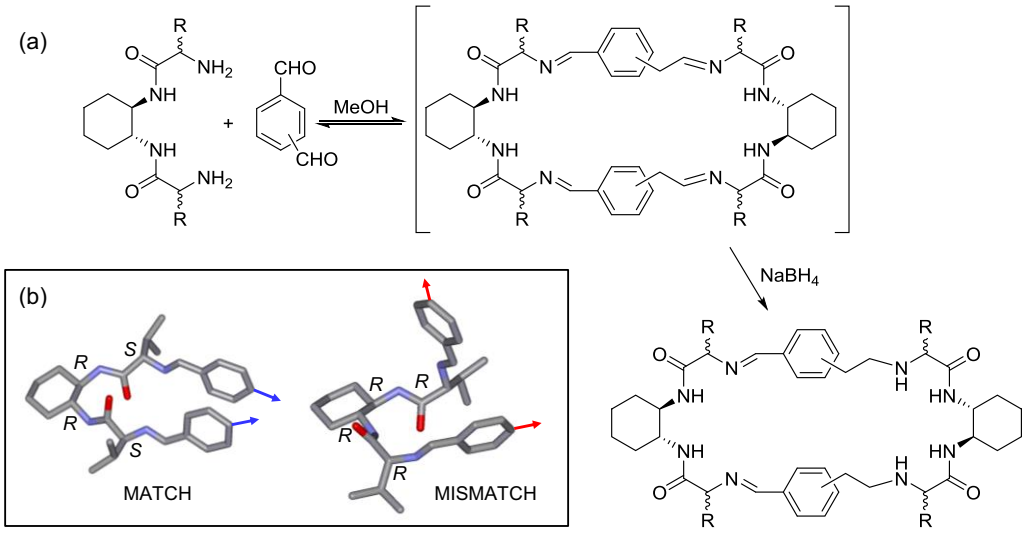


Figure 1.3. Example of *configurational preorganization*. (a) [2 + 2] reductive amination reaction between a pseudopeptidic bis(amidoamine) and a rigid planar aromatic dialdehyde. (b) Minimized geometries for the matched and mismatched configurations (Figure modified from references[46-47]).

The third mechanism is the *template induced preorganization* and represents an alternative and successful approach for those open-chain intermediates not properly preorganized. In the example depicted in Figure 1.4,[48-50] tetrabutyl ammonium terephthalate was selected as a template molecule in order to induce the preorganization of the previously commented mismatched configuration (Figure 1.3b). Thus, even though the system is structurally disfavored for the cyclization, the interactions with the anionic template force the suitable folding to yield the corresponding [2 + 2] macrocycle. In the absence of template, the linear molecule locates both ends at a long distance (Figure 1.4a). In the presence of the template, the H-bonds between the carboxylate groups of the terephthalate dianion and the amide protons force two alternated amide groups to rotate ~180°, locating both ends of the linear molecule at the appropriate distance and geometry for the C=N bond formation (Figure 1.4b).

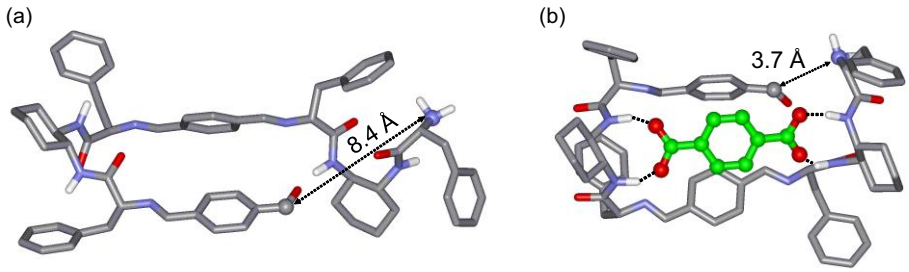


Figure 1.4. Example of a *template induced preorganization*. Minimized geometries for the aminoaldehyde intermediate of the mismatched combination represented in Figure 1.3, (a) in the absence of template and (b) in the presence of terephthalate (Figure modified from reference[48]).

In summary, preorganization of the linear precursor can improve the formation of macrocycles by different mechanisms. Moreover, regarding the purposes of the present thesis, these preorganization mechanisms can rule the outcome from a dynamic covalent mixture of pseudopeptidic oligomers.

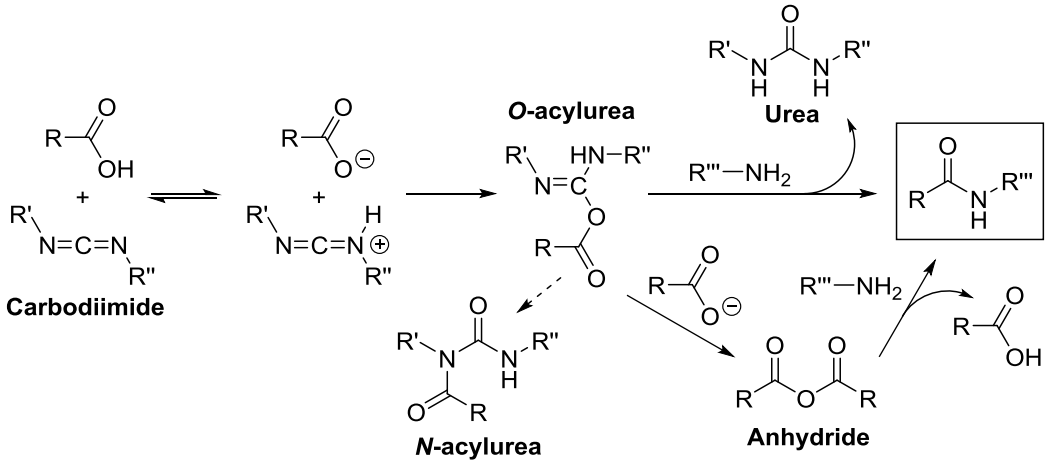
1.1.2. Amide bond formation

Amide bonds are typically synthesized from the reaction between carboxylic acids and amines. However, the reaction of these two functional groups does not occur spontaneously at room temperature but requires high temperatures (*e.g.* >200 °C)[51] for the necessary elimination of water. These conditions are typically detrimental to the integrity of the substrates and, for this reason, the activation of the carboxylic acid is usually necessary, converting the OH into a good leaving group prior to the addition of the amine. With this purpose, the so-called coupling agents[52-54] can be used to generate compounds such as acid chlorides, anhydrides, carbonic anhydrides or active esters. The following sections are focused on the use of carbodiimides, HOBt as an additive, and coupling reagents based on 1*H*-benzotriazole. These are the coupling agents used within the present thesis and the ones most widely employed for the synthesis of peptides and pseudopeptides.

1.1.2.1. Coupling using carbodiimides

Carbodiimides[55] were the first coupling agents to be synthesized and are still among the most widely employed reagents for the synthesis of amides. The mechanism for the carbodiimide-mediated coupling of carboxylic acids to amines is shown in Scheme 1.2. After a first acid/base reaction between the carbodiimide and the carboxylic acid, the carboxylate reacts with the protonated carbodiimide to give the corresponding *O*-acylurea. This intermediate can then yield a number of different products. The desired

amide can be obtained by either direct coupling with the amine or by formation of the carboxylic acid anhydride which subsequently reacts with the amine. The first of these two possibilities is the most common pathway and generates the urea of the corresponding carbodiimide as a by-product. This urea is normally easily eliminated from the crude mixture by filtration[c] or by aqueous washes.[d] Alternatively, in the absence of an efficient nucleophile, the *O*-acylurea intermediate can slowly rearrange forming the undesired *N*-acylurea side product.



Scheme 1.2. Mechanism for coupling carboxylic acids to amines using a carbodiimide as a coupling agent.

After generation of the *O*-acylurea, oxazolone formation can take place leading to undesired epimerization.[56] This side reaction is especially important when activating acid groups attached to the α position of an amide bond, *e.g.* in the consecutive coupling of amino acids.

1.1.2.2. The use of HOBt as an additive

In order to reduce the epimerization levels when using carbodiimides as coupling agents, 1-hydroxy-1*H*-benzotriazole (HOBt)[e,57] can be introduced as an additive, generally obtaining higher yields and much lower levels of epimerization.[f] The proposed mechanism[58] for the coupling of carboxylic acids to amines using

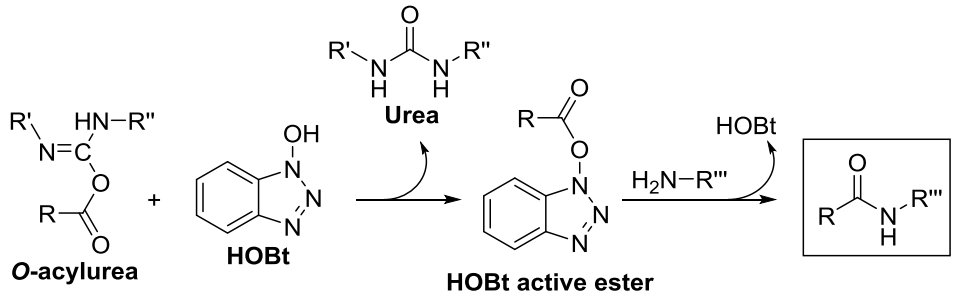
[c] As an example, the urea of *N,N'*-dicyclohexylcarbodiimide (DCCD)* can be eliminated from DMF by filtration. *Notice that DCCD is most commonly known as DCC. However, in this thesis DCC is used as the acronym for dynamic combinatorial chemistry.

[d] As an example, the urea of 1-ethyl-3-(dimethylaminopropyl)carbodiimide (EDC) can be eliminated from DMF/DCM by aqueous washes.

[e] In 1994 Carpino reported a related reagent, 1-hydroxy-7-azabenzotriazole (HOAt), with similar or even improved properties. Its proposed activation mechanism is the same as the one for HOBt.

[f] The main drawback associated to the use of HOBt/carbodiimide as coupling agents is the possible formation of the diazetidine derived from the carbodiimide. This process is catalyzed by HOBt.

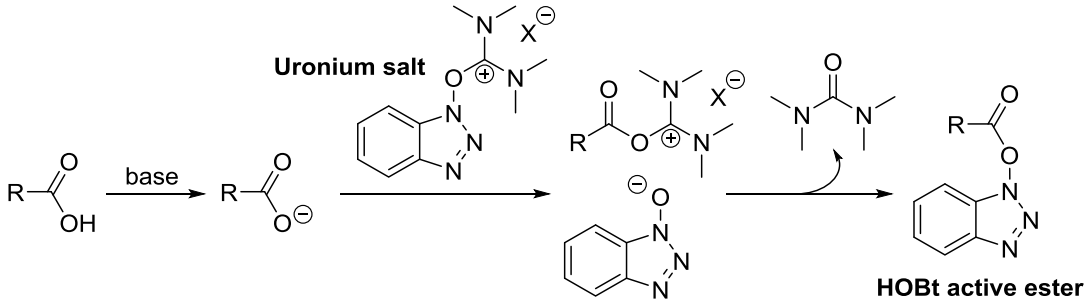
HOBt/carbodiimide as coupling agents is represented in Scheme 1.3. Initially, HOBt reacts with the *O*-acylurea to give the corresponding urea and the HOBt active ester. This has an enhanced reactivity since the OBt is a good leaving group and, moreover, it promotes the approach of the amine *via* H-bonding interactions. After the nucleophilic attack of the amine to the carbonyl of the active ester, the desired amide is formed and HOBt is recovered.



Scheme 1.3. Mechanism of activation by HOBt/carbodiimide.

1.1.2.3. Coupling reagents based on 1*H*-benzotriazole

Several activating agents are based on substituted 1*H*-benzotriazoles, including uronium/aminium[g] salts, phosphonium salts and immonium salts. In Scheme 1.4 the mechanism for the activation process using uranium/aminium salts is represented as an example. These reagents react with carboxylic acids to form the corresponding HOBt (or HOBt-like) active ester, which subsequently reacts with the amine to yield the desired amide.



Scheme 1.4. Mechanism of the activation process using uronium/aminium type reagents.

[g] Uronium and aminium isomers have been structurally identified and the true form depends on the solvent, the isolation method, the counter ion, *etc.* The main drawback associated to the use of uronium/aminium type reagents is the possible side reaction of the amine with the coupling agent to form a guanidinium side product.

1.2. Objectives

The main objective of the present Chapter is to design, synthesize and characterize a set of new molecular building blocks to be used for the generation of all dynamic combinatorial libraries studied in the following Chapters 2-5. The aim is to design simple and versatile molecules incorporating tailored structural information. The main structural requirements regarding their ability to generate useful dynamic libraries are the incorporation of molecular triggering moieties and a chromophore for easy monitoring. On the other hand, the specific requirements regarding their intended application are the incorporation of the following three features: i) peptide-like information, ii) differently charged functional groups, and iii) chiral information.

A design based on a common scaffold is envisioned to allow their easy preparation by means of a common synthetic pathway. Thus, an additional objective is to develop a versatile and general synthetic methodology for the preparation of the building blocks.

1.3. Results and discussion

1.3.1. Design of the building blocks

Thirteen pseudopeptidic dithiols (**1a-m** in Figure 1.5a) were designed in order to be used as bipodal building blocks (BBs) for the generation of disulfide-based dynamic combinatorial libraries. All of them were conceived with a common chemical backbone, with the only exception of **1m** that has a slightly simpler structure. The design of the BBs is based on a C_2-symmetric scaffold consisting of a central *m*-phenylenediamine that joins two identical arms, each formed by an amino acid residue. At the *N*-terminus of the two amino acids, dithiols **1a-l** incorporate a mercaptoacetyl moiety.

The function of the *m*-phenylenediamine is to link the two arms of the molecule in a rigid manner so that the BBs are preorganized for the formation of oligomeric macrocycles (section 1.1.1.2). The rigidity of this aromatic linker forces the two arms to stay apart, thereby preventing the formation of the entropically favored monomeric disulfide, *i.e.* the product of the intramolecular cyclization. Additionally, once diacylated, the resulting *m*-diamidophenyl chromophore allows sensitive HPLC-UV quantification at 254 nm.

The two equal natural α-amino acid residues present in each BB are the source of chemical diversity and provide compounds **1a-m** with different features depending on the nature of the corresponding side chain (Figure 1.5a). Thus, BB **1a** is based on glycine and does not have any side chain; BBs **1b-e** are based on asparagine, glutamine, serine and threonine respectively, and incorporate polar uncharged side chains with hydroxyl and amide functional groups; BBs **1f-g** are based on tyrosine and tryptophan respectively, and incorporate hydrophobic side chains with phenol and indole functional groups; BBs **1h-i** are based on aspartic and glutamic acids respectively, and incorporate the carboxylic acid functional group; BBs **1j-l** are based on ornithine, lysine and arginine respectively, and incorporate the amine and guanidine functional groups; and finally, BB **1m** is based on cysteine and incorporates the thiol functional group. In this latter case, the free *N*-terminus of the cysteine provides compound **1m** with two amine groups.

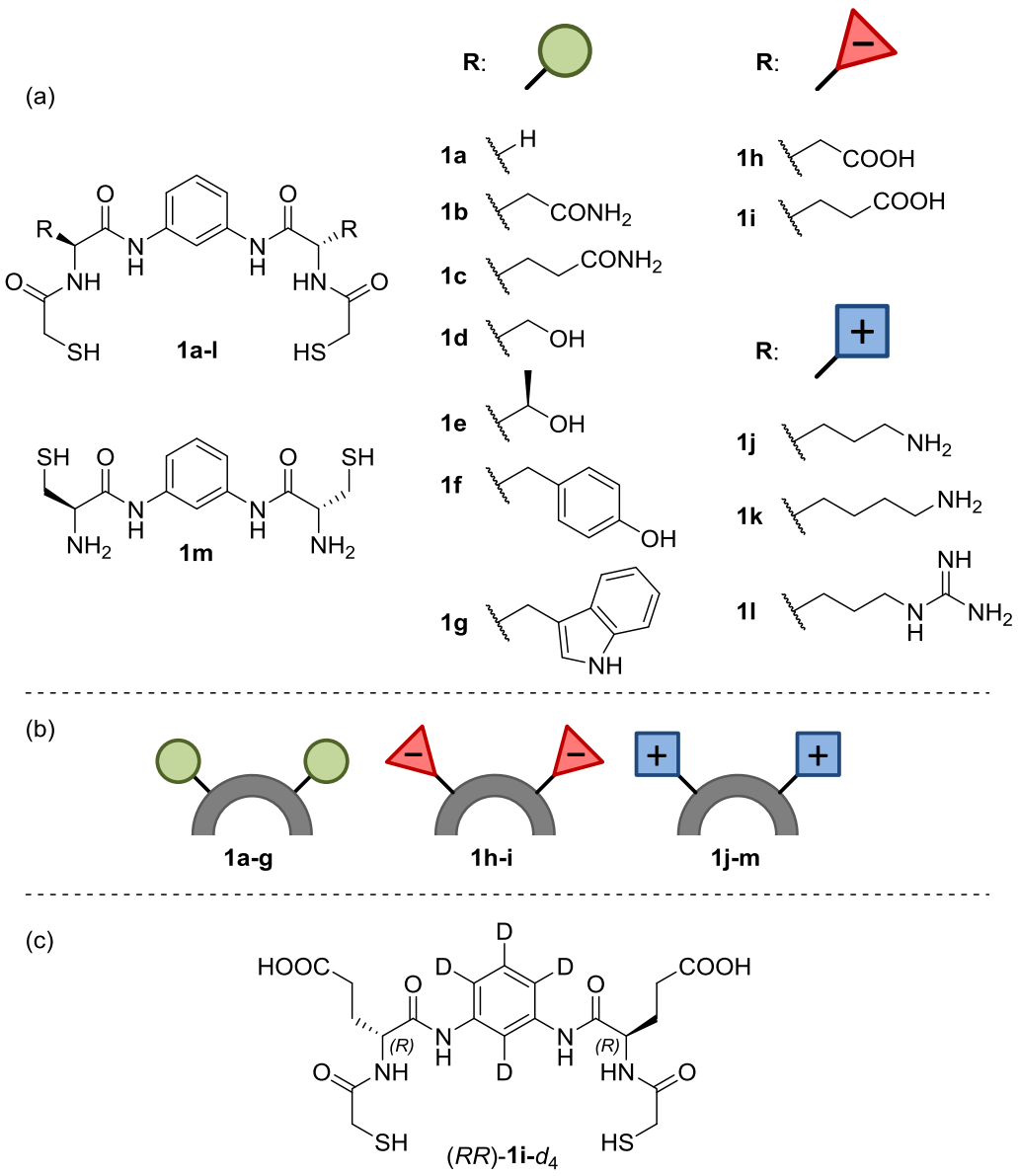


Figure 1.5. Chemical structure (a) and cartoon representation (b) of the thirteen BBs **1a-m**. (c) Chemical structure of (*RR*)-**1i**-d_4, the isotopically labeled enantiomeric form of **1i**.

The amino acid motif, far from being merely an easy way to introduce different side chains, is actually the most important structural part regarding the applications for which this set of BBs was designed. The amides along the pseudopeptidic backbone, together with the functional groups of the side chains, *a priori* make dithiols **1a-m** suitable BBs for the synthesis of molecular receptors.[h,16,23,59-60] The amino acid motif also provides peptide-like structural information, allowing the use of these BBs for the preparation of bio-inspired DCLs (Chapter 3). As represented by the cartoons in Figure 1.5b, at neutral pH the carboxylates in **1h-i** and the ammonium or guanidinium groups

[h] Preliminary tests were carried out regarding the use of BBs **1a-m** in the frame of DCC for the synthesis of molecular receptors for biologically relevant targets. However, this potential application is out of the scope of this thesis and will be the object of further investigations in our research group.

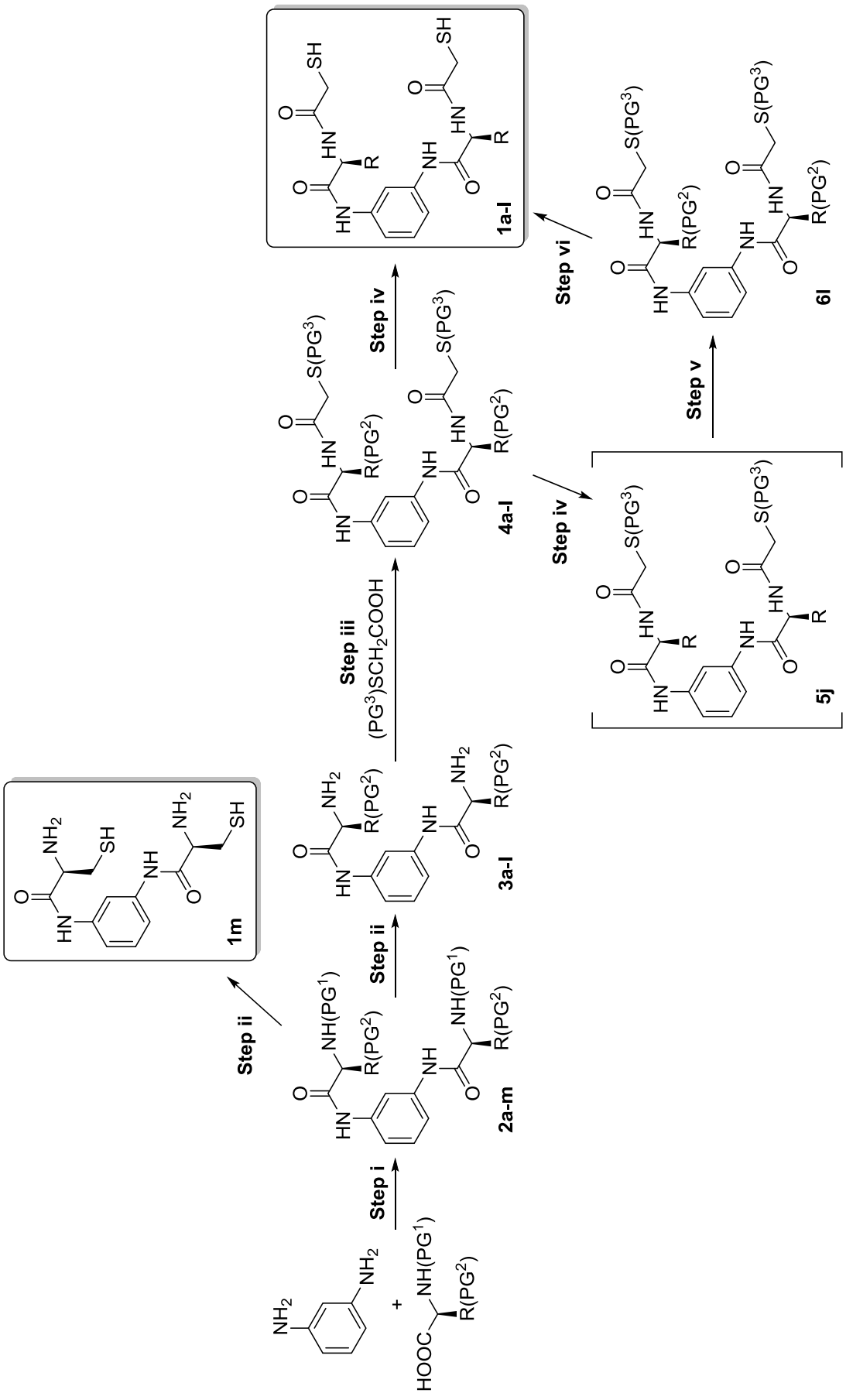
in **1j-m** have negative and positive charges respectively. These charged groups allow investigating the effect of the electrostatic interactions on the composition of DCLs (Chapter 4). And last but not least, the amino acid residues furnish dithiols **1b-m** with two stereogenic centers. The resulting chirality of the BBs has important implications in all applications mentioned above and allows investigating some specific aspects directly related to the stereoisomeric properties of the members of a DCL (Chapter 5). Regarding this latter application, in addition of the thirteen dithiols shown in Figure 1.5a, one last BB was designed as the enantiomeric form of the Glu-based BB **1i**. This additional dithiol ((*RR*)-**1i**-d_4 in Figure 1.5c) is equipped with two unnatural D-glutamic acid residues and incorporates an isotopically labeled *m*-phenylenediamine-2,4,5,6-d_4 linker.

The mercaptoacetyl moieties attached to the two *N*-termini in BBs **1a-l** incorporate the thiol function needed for the reversible connection of the BBs by the formation of disulfides (Chapter 2). In the particular case of BB **1m**, the cysteine amino acid already bears the thiol function and, therefore, the mercaptoacetyl moiety was not incorporated.

1.3.2. Synthesis of the building blocks

The retrosynthetic analysis of BBs **1a-m** is quite simple as it consists in considering the three previously described constituent parts (the *m*-phenylenediamine linker, the two amino acid residues and the two mercaptoacetyl moieties) as independent entities to be joined. This trivial retrosynthetic scheme entails the breakage of all amide bonds present in the BBs and delimits the scope of the synthetic "problem" to the consecutive formation of amides. The synthetic methodology for the formation of polyamides by the consecutive coupling of amines and carboxylic acids has extensively been investigated and developed in the field of peptide synthesis (section 1.1.2).

Taking advantage of the C_2-symmetry of dithiols **1a-m**, the proposed synthetic pathway starts from the central *m*-phenylenediamine linker and consists in simultaneously growing the two arms of the molecule. The growth proceeds in a *C*-terminal to *N*-terminal fashion and requires the use of three protecting groups (PGs). The incorporated amino acids have the *N*-terminus and the functional group of the side chain protected with PG^1 and PG^2 respectively. For the synthesis of compounds **1a-l**, the mercaptoacetic acid is incorporated with the thiol function protected with PG^3.



Scheme 1.5. Synthetic pathway for the synthesis of building blocks **1a-m**.

As shown in Scheme 1.5, in a first step *m*-phenylenediamine is diacylated with two units of the corresponding protected amino acid to give intermediates **2a-m**. For the particular case of the Cys-based BB, in a second and last step both PG^1 and PG^2 are removed from **2m** and dithiol **1m** is readily obtained. For the rest of the BBs, the synthetic pathway proceeds with the selective removal of PG^1, *i.e.* the deprotection of the *N*-termini, to give the free diamines **3a-l**. These intermediates are then diacylated with two units of the *S*-protected mercaptoacetic acid to form intermediates **4a-l**. In a fourth and last step, PG^2 and PG^3 are simultaneously removed to finally obtain dithiols **1a-k**. Unfortunately, this final deprotection was not possible for the Arg-based dithiol **1l** (section 1.3.2.4) and, consequently, an alternative synthetic route through intermediates **5j** and **6l** was formulated for the synthesis of this BB (section 1.3.2.5).

Table 1.1. Protecting groups for the synthesis of BBs **1a-m**.

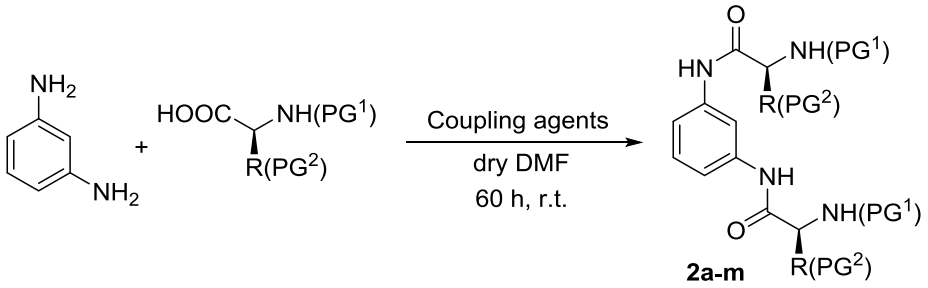
entry	**BB**	PG^1	PG^2	PG^3
1	**1a**	Boc	-	Trt
2	**1b**	Fmoc	Trt	
3	**1c**			
4	**1d**		^{t}Bu	
5	**1e**			
6	**1f**			
7	**1g**		Boc	
8	**1h**		^{t}Bu	
9	**1i**			
10	**1j**	Boc	Alloc	Trt
11	**1k**		Cbz	
12	**1l**	Boc	Cbz[a]/Boc[b]	
13	**1m**	Boc	Trt	-

[a] PG of the initially proposed synthetic route (intermediates **2l**, **3l** and **4l**). [b] PG of the alternative synthetic route (intermediate **6l**).

The identity of the protecting groups is disclosed in Table 1.1 and was carefully selected with criteria of orthogonality requirements and commercial availability. For the synthesis of compounds **1b-l** (entries 2-12 in Table 1.1), PG^1 and PG^2 were selected to be orthogonal, as the first one was selectively removed in the presence of the second one. For **1a** and **1m** (BBs based on Gly and Cys respectively, entries 1 and 13 in Table 1.1) the strategy was simpler: the synthesis of **1a** did not require the PG^2 protection and the synthesis of **1m** was carried out with non-orthogonal protecting groups. At the last step of the synthetic pathway (**step iv** in Scheme 1.5), PG^2 and PG^3 were simultaneously removed in order to obtain final dithiols **1a-k**. Hence, in most of the cases these two protecting groups were chosen to be non-orthogonal. An exception of that was the alternative synthetic route formulated for the Arg-based BB **1l**, in which PG^2 was selectively removed from the Orn-based intermediate **4j** in the presence of PG^3 (entry 10 in Table 1.1).

1.3.2.1. **Step i**: diacylation of *m*-phenylenediamine

For the diacylation of *m*-phenylenediamine with two units of the corresponding protected amino acid PG^1-L-AA(PG^2)-OH (**step i** in Scheme 1.5 and Scheme 1.6), the standard one-pot coupling methodology for peptide synthesis in solution phase was used.[53] The HOBt active ester of the corresponding amino acid was prepared *in situ* by the use of stoichiometric amounts of HOBt and a carbodiimide as coupling agents. Then, *m*-phenylenediamine was added over the mixture to react with two molecules of the active ester, forming the final diamides **2a-m**. Two different activating strategies were used depending on the nature of the *N*-terminus protection (PG^1). Thus, for those amino acids with the *N*-terminus protected with the acid-labile Boc group (synthesis of **2a** and **2j-m**, entries 1 and 10-13 in Table 1.1), the reaction was performed with either HOBt/EDC·HCl[61-62] or HOBt/HBTU[63] as coupling agents, and DIPEA was used as a catalytic base. Alternatively, for those amino acids with the *N*-terminus protected with the base-labile Fmoc group (synthesis of **2b-i**, entries 2-9 in Table 1.1), HOBt/DCCD[57-58] were implemented as coupling agents and the use of a catalytic base was avoided in order not to cleave the Fmoc protection of the *N*-termini.



Scheme 1.6. Synthesis of intermediates **2a-m**. from m

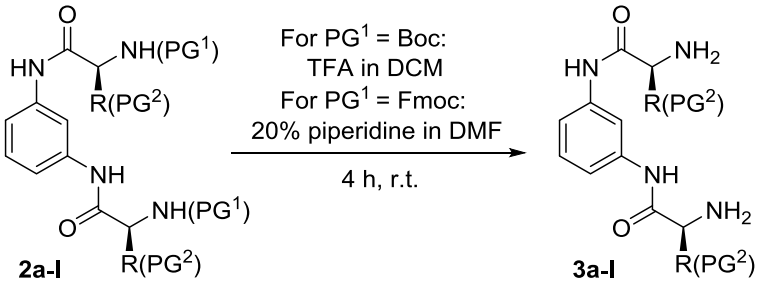
The reaction was carried out in dry DMF[i] and the mixtures were stirred at room temperature and under inert atmosphere of Ar (experimental details in section 1.5.4). The progress of the reaction was followed by TLC and UPLC-MS, observing the formation of the monoacylated intermediate within few hours, followed by the much slower formation of the diacylated product. Anilines are relatively poor nucleophiles due to electron pair delocalization. Moreover, the amide resulting from the first acylation further reduces the nucleophilicity of the monoacylated intermediate. For this reason, the second acylation was notably slower compared to the first one, and the

[i] The synthesis of intermediate **2m** was exceptionally performed in dry DCM. This was found not to be a good solvent for this reaction, with poor solubility of the starting reagents and a low reaction yield (29%).

reaction needed an overall of 60 hours to reach a high conversion.[j] After standard workup and purification procedures, pure products **2a-m** were obtained in 40-70% yield. This moderate yield was mainly attributed to the poor nucleophilicity of the monoacylated intermediate.

1.3.2.2. **Step ii**: deprotection of the *N*-termini (removal of PG^1)

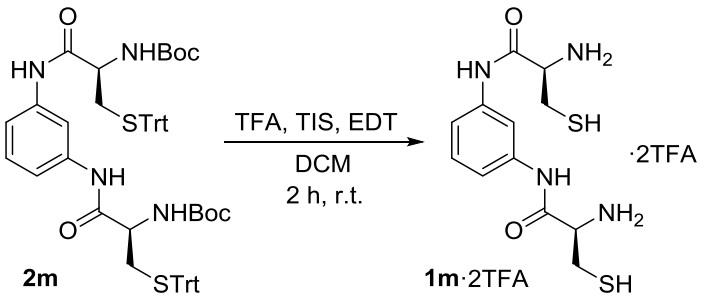
The deprotection of the two *N*-termini of intermediates **2a-l** (**step ii** in Scheme 1.5 and Scheme 1.7) was carried out by means of standard acidic or basic treatment depending on the identity of PG^1. Thus, for BBs **2a** and **2j-l** the Boc protecting group was removed by treatment with TFA in DCM,[64] obtaining the TFA salts of **3a** and **3j-l**. On the other hand, the Fmoc protecting group of BBs **2b-i** was removed by treatment with 20% (v/v) piperidine in dry DMF,[65] obtaining the free diamines **3b-i**. The reactions were carried out at room temperature and under stirring (experimental details in section 1.5.5) and in all cases the TLC and UPLC-MS analyses confirmed full conversion within few hours. Final products were precipitated in diethyl ether and intermediates **3a-l** were obtained in very good yields (>90% in most of the cases).



Scheme 1.7. Sinthesis of intermediates **3a-l** from **2a-l**.

For the particular case of the Cys-based BB, in addition to the protecting groups of the two *N*-termini (PG^1 = Boc), the protecting groups of the thiol functionalities (PG^2 = Trt) were also removed in this second synthetic step, and final dithiol **1m** was readily obtained (Scheme 1.8). The experimental conditions used for the simultaneous removal of PG^1 and PG^2 of **2m** were exactly the same as the ones discussed in section 1.3.2.4 for the simultaneous removal of PG^2 and PG^3 of **4a-i**.

[j] In spite of the long reaction time, in some cases a certain amount of the monoacylated intermediate was isolated in the final purification.

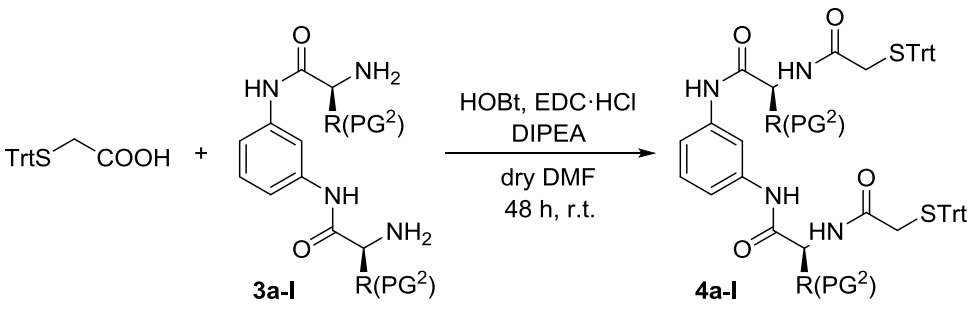


Scheme 1.8. Sinthesis of final product **1m** from **2m**.

1.3.2.3. **Step iii**: diacylation of the *N*-termini with $TrtSCH_2COOH$

First of all, the *S*-protected mercaptoacetic acid ($TrtSCH_2COOH$) was prepared as previously described,[66] by the nucleophilic substitution (S_N1) of triphenylmethanol with mercaptoacetic acid in acidic medium (experimental details in section 1.5.2). The trityl-protected mercaptoacetic acid was obtained in 83% yield.

For the coupling of intermediates **3a-l** with two molecules of the *S*-protected mercaptoacetic acid (**step iii** in Scheme 1.5 and Scheme 1.9), the one-pot coupling methodology already commented for **step i** (section 1.3.2.1) was implemented. In this case, HOBt/EDC·HCl were used as coupling agents and DIPEA was added as a catalytic base.



Scheme 1.9. Synthesis of intermediates **4a-l** from **3a-l**.

The reaction was performed in dry DMF and the mixtures were stirred at room temperature for 48 hours under inert atmosphere of Ar (experimental details in section 1.5.6). The stronger nucleophilicity of aliphatic amines **3a-l** compared with the aromatic *m*-phenylenediamine previously used in **step i**, allowed obtaining products **4a-l** in generally better yields (>70% in most of the cases)[k] and shorter reaction times.

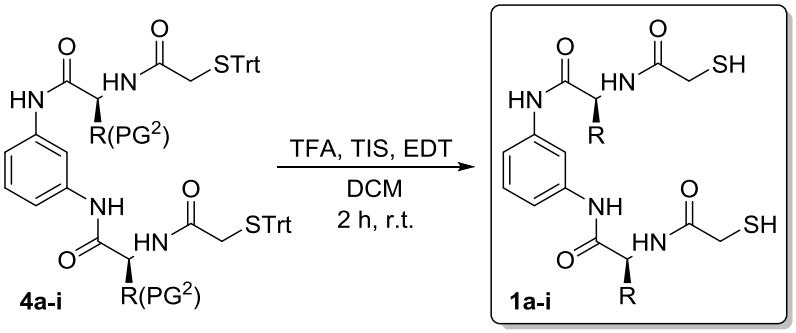
[k] An exception was the 15% yield obtained for the synthesis of **4j**. The reaction was not optimized and the low yield obtained might be the result of product loss during the workup and purification processes.

1.3.2.4. **Step iv**: final deprotection (removal of PG^2 and PG^3)

In the final deprotection step (**step iv** in Scheme 1.5) both PG^2 and PG^3 are simultaneously removed. With this purpose, different methodologies were used depending on the nature of PG^2. Thus, in the common case of an acid-labile PG^2, a standard acidic treatment was used. Alternatively, for the deprotection of intermediates **4j** and **4k**, specific protocols for the removal of the Alloc and Cbz protecting groups were implemented respectively. The unsuccessful attempts to synthesize BB **1l** by deprotection of intermediate **4l** are also mentioned below.

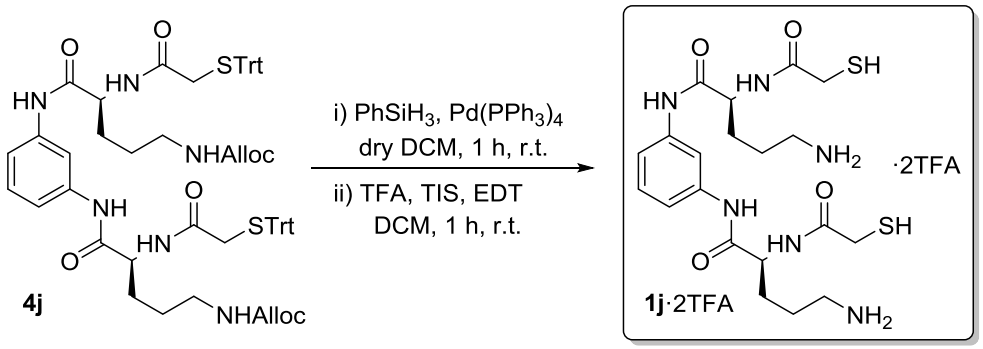
For the deprotection of the acid-labile ^{t}Bu, Boc and Trt protecting groups of intermediates **4a-i**, our first attempts consisted in using a standard deprotection cocktail containing TFA, EDT, TIS and H_2O (94/2.5/1.0/2.5, v/v/v/v). The function of TFA is to provide the acidic medium required for the deprotection, being the major solvent of the reaction mixture; whereas EDT, TIS and H_2O are used as scavengers in order to capture the carbocations generated during the deprotection process. Unfortunately, the use of this deprotection cocktail led to complex crude mixtures in which the desired product was present only in small amounts. After several tests, we eventually concluded that the presence of water was the responsible for the undesired side reactions, and we attributed this detrimental effect to the hydrolytic instability of certain amides in the presence of TFA/water mixtures.[67] Consequently, the use of water as a scavenger was avoided and the deprotection of compounds **4a-i** was carried out with only TFA, TIS and EDT,[1,68] together with a small amount of DCM to improve the solubility (Scheme 1.10). The reaction mixtures were stirred at room temperature for 2 hours, after which the products were precipitated in diethyl ether (experimental details in section 1.5.8). The crudes were subsequently purified by reversed-phase flash chromatography using mixtures of CH_3CN with 0.07% (v/v) TFA and H_2O with 0.1% (v/v) TFA as eluent. The presence of TFA in the eluent is crucial to avoid air oxidation of the thiols during the purification process. Pure dithiols **1a-i** were obtained in 40-50% yield. This moderate-low yield was mainly attributed to product loss during the purification.

[1] TIS and EDT are complementary in their behavior: EDT captures the produced carbocations very rapidly but in a reversible process, whereas TIS reacts with a constant rate an order of magnitude lower but in an irreversible process.



Scheme 1.10. Synthesis of final dithiols **1a-i** from **4a-i**.

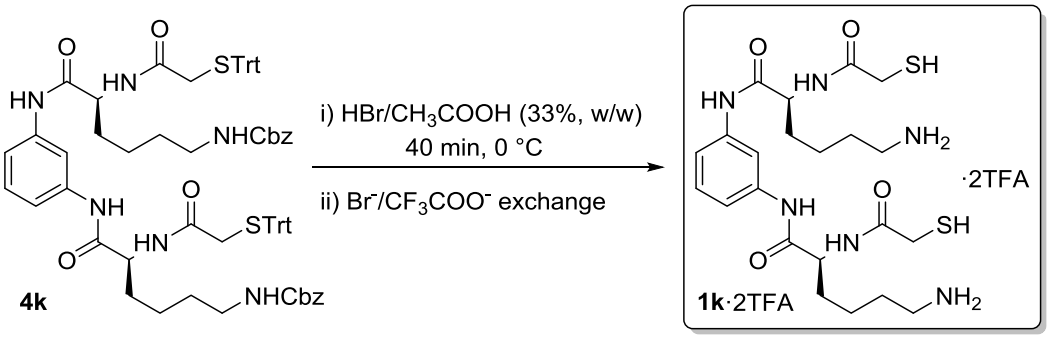
The final deprotection of **4j** was achieved by means of two consecutive deprotection procedures (Scheme 1.11). Firstly, the Alloc protecting groups were removed as previously described[69] with a catalytic amount of $Pd(PPh_3)_4$ in the presence $PhSiH_3$ as a scavenger for the allyl system (experimental details in section 1.5.8). Secondly, in order to remove the Trt protecting group, the obtained residue was directly subjected to the deprotection protocol described above for **4a-i**. After precipitation and purification processes, pure dithiol **1j** was obtained in 56% yield over the two steps.



Scheme 1.11. Synthesis of final dithiol **1j** from **4j**.

For the removal of the Cbz protecting groups of **4k**, our first attempts consisted in using standard hydrogenolysis conditions, *i.e.* a catalytic amount of Pd/C (10%, w/w) in MeOH at 1 atm of H_2. The reaction was followed by TLC and UPLC-MS, observing the fast removal of the Trt groups;[70] whereas the Cbz groups remained unaltered, even at long reaction times (>48 h). We hypothesized that the free thiol groups resulting from the Trt cleavage would passivate the catalyst and, thereby, stop the hydrogenolysis reaction. Consequently, we performed the same reaction with larger amounts of catalyst but no better results were obtained. In order to further force the reaction, the Pd/C

catalyst was replaced by palladium black. Unfortunately, this new catalyst appeared to promote partial cleavage of the CH_2–SH bond.[m,71]



Scheme 1.12. Synthesis of final dithiol **1k** from **4k**.

An alternative method was considered consisting in the use of dry HBr in CH_3COOH.[72] To a solution of compound **4k** in dry DCM, the scavenger TIS and a solution of HBr/CH_3COOH (33%, w/w) were added. After 1 hour stirring at room temperature, the subsequent UPLC-MS and ^{1}H NMR analyses of the crude product showed the successful removal of the two Cbz and the two Trt groups. However, under these conditions, a side reaction was found to be simultaneously taking place: the acetylation of the thiol groups by reaction with the acetic acid present in the medium. After screening different reaction conditions, we found out that, by shortening the reaction time and decreasing the temperature, the acetylation of the thiols can be importantly minimized while the conversion of the deprotection reaction remains very high. Thus, the reaction was eventually performed stirring the mixture at 0 °C for 40 min (Scheme 1.12, experimental details in section 1.5.8). After similar workup and purification processes as the ones previously commented for the synthesis of **1a-i**, the TFA salt[n] of **1k** was obtained in 84% yield.

The Arg-based intermediate **4l** has each of the guanidine moieties protected with two Cbz groups. Thus, the final deprotection of this compound entails the removal of an overall of four Cbz groups and two Trt groups. With this purpose, the methodology commented above for the synthesis of **1k** was tested. By treatment with HBr/CH_3COOH, the cleavage of the two Trt groups and the first two Cbz groups was found to be relatively fast, whereas the removal of the last two Cbz groups required

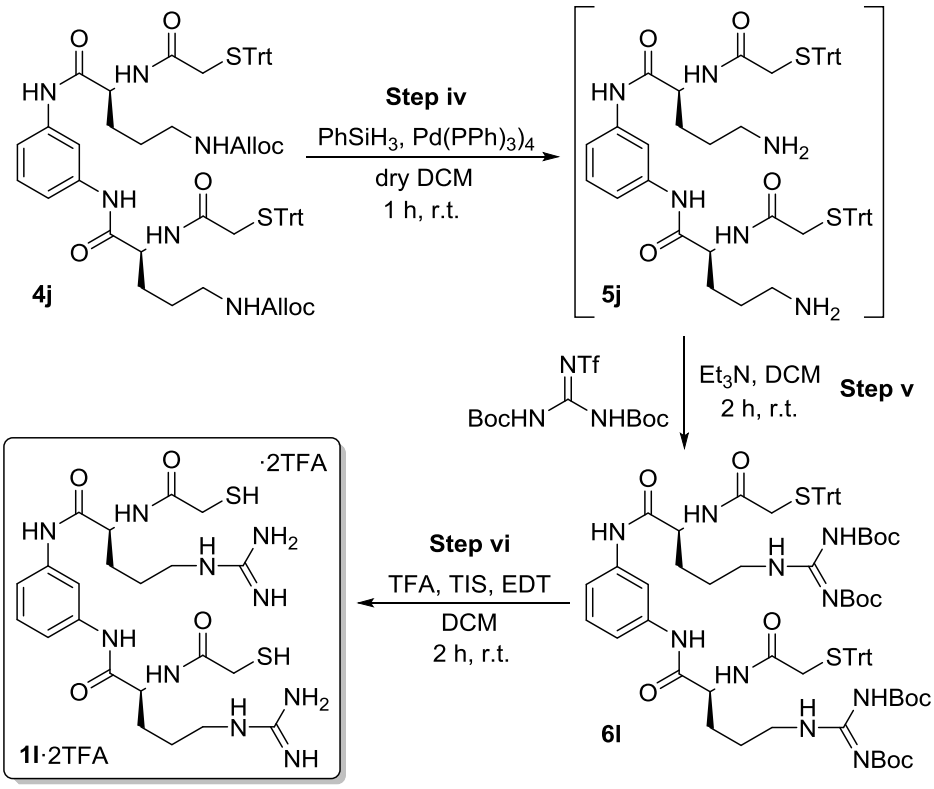
[m] Hydrodesulfurization (HDS) is an industrially important catalytic process extensively used to remove sulfur from natural gas and petroleum.
[n] After precipitation in diethyl ether, the hydrobromide salt of **1k** was obtained. Then, the bromide anions were exchanged by trifluoroacetate anions during the purification process and, thereby, the TFA salt of **1k** was eventually isolated.

several hours. Unavoidably, at such long reaction times the product was almost completely acetylated. Alternatively, in order to avoid the side reaction with the solvent, we tested two other methodologies: i) the use of dry HBr in EtOH, and ii) the standard hydrogenolysis protocol (catalytic amount of Pd/C in MeOH at 1 atm of H_2). Unfortunately, none of these alternatives allowed us to achieve the full deprotection of **4l**, and we eventually decided to reformulate the synthetic pathway for the synthesis of the Arg-based BB **1l**.

1.3.2.5. **Steps v**, **vi** and **vii**: alternative synthetic route for **1l**

The alternative route for the synthesis of the Arg-based BB **1l** starts with the selective deprotection of the two amines of the Orn-based intermediate **4j** (**step iv** in Scheme 1.13). By treatment with the neutral conditions required for the removal of the Alloc group,[o,69] the two amines of **4j** were selectively deprotected in the presence of the Trt groups, obtaining intermediate **5j** (experimental details in section 1.5.7). The next step is the key transformation of the new synthetic pathway and consists in the guanidination of the two amines of intermediate **5j** (**step v** in Scheme 1.13). In this reaction, the ornithine moieties of **5j** are converted into Boc-protected arginines and compound **6l** is obtained as a direct precursor of the Arg-based BB **1l**. Hence, intermediates **6l** and **4l** are analogous compounds with differently protected guanidines (Cbz and Boc protection for **4l** and **6l** respectively). The reaction was carried out with *N*,*N*′-di-Boc-*N*″-triflylguanidine, a reagent specifically designed for the preparation of Boc-protected guanidines from primary amines.[73] The crude diamine **5j** was mixed with a small excess of the guanidination reagent and Et_3N as a catalytic base (experimental details in section 1.5.7). After completion of the reaction, Et_3N and the by-product triflic amide were removed during a simple aqueous workup procedure. The product was purified by flash chromatography and pure compound **6l** was obtained in 63% yield over the two steps.

[o] Catalytic amount of $Pd(PPh_3)_4$ in the presence $PhSiH_3$, as previously commented for the synthesis of **1j**.

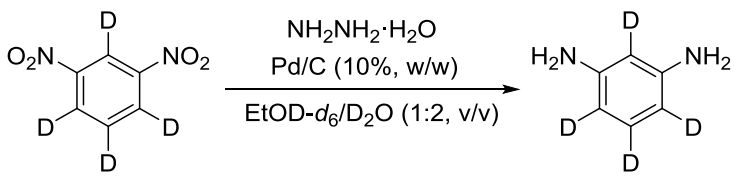


Scheme 1.13. Synthesis of final dithiol **1l** from **4j**, through intermediates **5j** and **6l**.

Finally, intermediate **6l** was fully deprotected in the same way as previously commented for the synthesis of **1a-i** (**step vi** in Scheme 1.13), and pure dithiol **1l** was obtained in 27% yield. This low yield was mainly attributed to product loss during the purification.

1.3.2.6. Synthesis of dithiol (*RR*)-**1i**-d_4

The synthesis of BB (*RR*)-**1i**-d_4 was carried out in the same way as already commented for its enantiomeric form **1i**. In this case, however, the starting reagents were obviously different: the protected L-glutamic acid was replaced by the corresponding commercially available enantiomer, and the *m*-phenylenediamine was replaced by the isotopically labeled *m*-phenylenediamine-2,4,5,6-d_4. This diamine was synthesized by reduction of the commercially available 1,3-dinitrobenzene-d_4 as shown in Scheme 1.14.



Scheme 1.14. Synthesis of *m*-phenylenediamine-2,4,5,6-d_4.

In this reaction, hydrazine reduces the two nitro groups of the starting material to give the corresponding diamine in a process catalyzed by the presence of Pd/C. This transformation was previously described for the reduction of a similar (but non-deuterated) 1,3-dinitrobenzene derivative.[74] Unfortunately, when reproducing the same reaction conditions for the reduction of 1,3-dinitrobenzene-d_4 (*i.e.* catalytic amount of Pd/C and reflux of ethanol for 2 hours), the deuterium labeling of the reagent was partially lost during the reaction, and only 1-2 of the 4 deuterium atoms were present at the resulting *m*-phenylenediamine (Figure 1.6a).

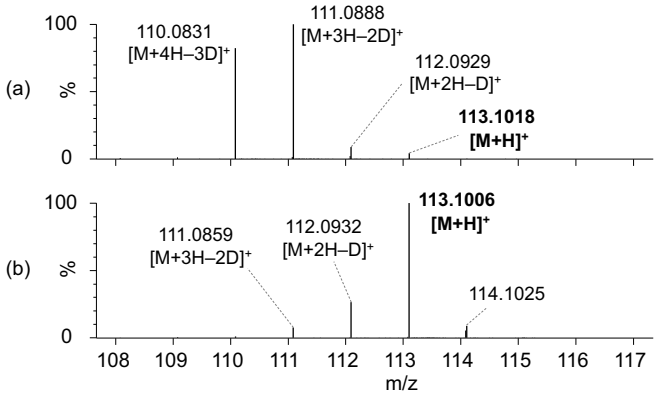


Figure 1.6. ESI(+)-TOF isotopic patterns of the product obtained using (a) non-deuterated and (b) deuterated solvents for the reaction represented in Scheme 1.14. Simulated isotopic pattern for $[M+H]^+$: 113.1011 (100.0%) and 114.1045 (6.5%).

Palladium(0) is known to catalyze the D/H exchange of aromatic deuterium atoms in the presence of protic solvents and under high temperatures.[75] Because of that, we decided to test $FeSO_4 \cdot 7H_2O$ as an alternative catalyst[76] but, again, the product showed an isotopic pattern similar to the one in Figure 1.6a. The 2,4 and 6 positions of the aromatic ring are activated by the two electron donating amino groups of the *m*-phenylenediamine. Thus, the exchange reaction in these three positions would be strongly favored regardless of the nature of the catalyst.[p] In agreement with that, only three of the four deuterium atoms of *m*-phenylenediamine are efficiently exchanging with the solvent, as evidenced by the absence of $[M+5H–4D]^+$ in the isotopic pattern shown in Figure 1.6a.

Finally, to definitely prevent the loss of the isotopic labeling, the reaction was performed in deuterated solvents as represented in Scheme 1.14, and the desired *m*-phenylenediamine-2,4,5,6-d_4 was obtained in 89% yield (experimental details in section

[p] At short reaction times, the MS analysis of the intermediate 3-nitroaniline-2,4,5,6-d_4 did not show D/H exchange because the presence of the electron withdrawing nitro group deactivates the substitution.

1.5.3). The use of non-deuterated hydrazine was the responsible for the small amount[q] of D/H exchange products observed at the corresponding isotopic pattern (Figure 1.6b). Even so, the achieved degree of deuterium labeling was considered to be enough for the identification purpose of Chapter 5.

[q] The MS analysis showed the presence of tetra-, tri- and dideuterated *m*-phenylenediamine in the proportion of 75%, 20% and 5%, respectively.

1.4. Conclusions

Fourteen new pseudopeptidic dithiols (**1a-m** and (*RR*)-**1i**-d_4) have been designed, synthesized and fully characterized in order to be used as bipodal building blocks for the generation of disulfide-based dynamic combinatorial libraries. The design of the building blocks is based on a C_2-symmetric scaffold consisting of a central *m*-phenylenediamine chromophore that rigidly joins two identical arms, each formed by an amino acid residue. The amino acid moieties are the most important structural part, as they provide the building blocks with peptide-like information, differently charged side chains and chiral information. At the *N*-terminus of the two amino acids, compounds **1a-l** and (*RR*)-**1i**-d_4 incorporate a mercaptoacetyl moiety.

The synthesis of these building blocks has been achieved *via* a common synthetic pathway consisting of four steps: two amine acylation reactions and two deprotection steps. Exceptionally, dithiols **1m** and **1l** have been prepared by shorter (two steps) and longer (six steps) synthetic pathways respectively.

1.5. Experimental section

1.5.1. General methods

Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka, Merck or Iris Biotech) and were used without further purification. Chromatographic purifications were performed on a Biotage® Isolera Prime™ equipment using Biotage® SNAP KP-Sil and Biotage® SNAP KP-C18-HS cartridges for normal- and reversed-phase purifications respectively. TLCs were performed using 6 x 3 cm SiO_2 pre-coated aluminium plates (ALUGRAM® SIL G/UV_{254}).

RP-HPLC analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using a reversed-phase X-Terra C_{18} (15 x 0.46 cm, 5 μm) column. (CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelengths were set at 220 and 254 nm. The temperature of the column was set at 25 °C. The HPLC samples were prepared by dilution with an acidic solution of 89% H_2O, 10% CH_3CN and 1% TFA.

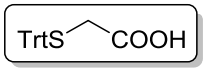
Nuclear Magnetic Resonance (NMR) spectroscopic experiments were carried out on a Varian INOVA 500 spectrometer (500 MHz for ^1H and 126 MHz for ^{13}C), a Varian Mercury 400 instrument (400 MHz for ^1H and 101 MHz for ^{13}C) and a Varian Unity 300 (300 MHz for ^1H and 75 MHz for ^{13}C). The chemical shifts (δ) are reported in ppm relative to trimethylsilane (TMS), and coupling constants (*J*) are reported in Hertz (Hz). Signal assignment was carried out using various 2D NMR spectra including ^1H-^1H gCOSY, ^1H-^{13}C gHSQC and ^1H-^{13}C gHMBC.[r] For assigning signals of ^1H NMR spectra de following abbreviations are used: s = singlet, d = doublet, t = triplet, q = quartet, ABq = AB quartet, quint = quintet, dd = doublet of doublets, dt = doublet of triplets, td = triplet of doublets, dq = doublet of quartets, qd = quartet of doublets, ddd = double doublet of doublets, ddt = double doublet of triplets, m = multiplet, and br = broad signal.

HRMS analyses were carried out at the IQAC Mass Spectrometry Facility, using a UPLC-ESI-TOF equipment: [Acquity UPLC® BEH C_{18} 1.7 mm, 2.1x100 mm, LCT Premier Xe, Waters]. (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH)

[r] NMR spectra of final products **1a-m** are provided in the electronic Annex accompanying this thesis.

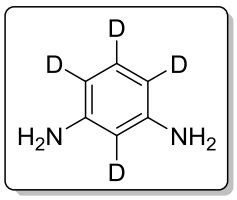
mixtures at 0.3 mL/min were used as mobile phase. The characterization of the pure products and intermediates was performed in flow injection analysis (FIA) mode.

1.5.2. Synthesis of tritylsulfanyl acetic acid



This compound was prepared as previously described.[66] To a solution of mercaptoacetic acid (4.60 g, 49.9 mmol) and triphenylmethanol (13.0 g, 49.9 mmol) in chloroform (50 mL), trifluoroacetic acid (TFA, 5.0 mL, 65 mmol) was added. After the mixture was stirred at room temperature for 2 hours, volatiles were removed *in vacuo*. The crude product was recrystallized from dichloromethane/hexane to give 13.9 g of tritylsulfanyl acetic acid (83% yield) as a white solid. Rf AcOEt/Hexane, 3:7, (v:v): 0.39. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.42 (d, *J* = 7.3 Hz, 6H, CH_{Ar}), 7.30 (t, *J* = 7.6 Hz, 6H, CH_{Ar}), 7.23 (t, *J* = 7.3 Hz, 3H, CH_{Ar}), 3.03 (s, 2H, CH_2). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 174.7 (1 x CO), 144.0 (3 x C_{Ar}), 129.6 (6 x CH_{Ar}), 128.3 (6 x CH_{Ar}), 127.2 (3 x CH_{Ar}), 67.4 (1 x C), 34.5 (1 x CH_2). HRMS (ESI–) calcd. for $C_{21}H_{18}O_2S$ $[2M–H]^-$ (m/z): 667.1982, found: 667.1996.

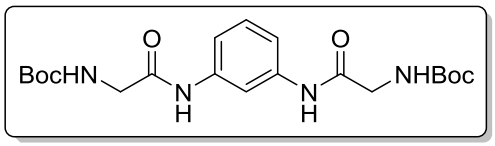
1.5.3. Synthesis of *m*-phenylenediamine-2,4,5,6-d_4



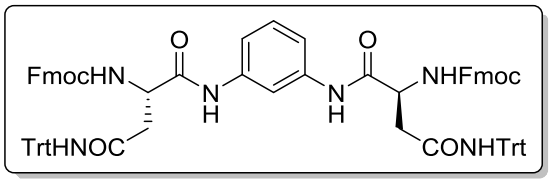
A mixture of 1,3-dinitrobenzene-d_4 (99% atom D, 723 mg, 4.20 mmol), 50 mg of Pd/C (10 wt. %), hydrazine monohydrate (8 mL, 165 mmol), 4.0 mL of EtOD-d_6, and 8.0 mL of D_2O was refluxed with stirring for 1 hour, after which complete conversion of the starting material was observed by TLC (Rf AcOEt/Hexane, 7:3 (v:v): 0.34). After Celite® filtration with abundant EtOH, the filtrate was concentrated under reduced pressure and purified by flash chromatography using hexane: AcOEt as eluent (from 60% to 70% AcOEt) to give 418 mg of *m*-phenylenediamine-2,4,5,6-d_4 (89% yield) as a pale white solid. HRMS (ESI+) calcd. for $C_6H_4D_4N_2$ $[M+H]^+$ (m/z): 113.1011, found: 113.1006. ^{13}C NMR (101 MHz, MeOD-d_4):[s] δ = 149.0 (2 x C_{Ar}), 130.2 (1 x CD_{Ar}), 107.3 (2 x CD_{Ar}), 104.1 (1 x CD_{Ar}).

[s] Due to partial D/H exchange (section 1.3.2.6), triplet multiplicity for CD_{Ar} carbons (indicating coupling of carbon to deuterium) was not observed.

1.5.4. Synthesis of intermediates [2a-m]

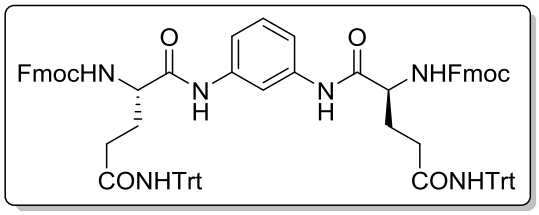


[**2a**]: to a solution of Boc-Gly-OH (1.06 g, 6.05 mmol) in dry DMF (15 mL), HOBt (908 mg, 6.72 mmol), EDC·HCl (1.29 g, 6.74 mmol) and DIPEA (2.3 mL, 13 mmol) were added. The resulting mixture was cooled down to 0ºC in an ice-water bath. Then, a solution of *m*-phenylenediamine (302 mg, 2.79 mmol) in dry DMF (10 mL) was added via cannula under inert atmosphere of Ar. The mixture was stirred at room temperature for 60 hours, after which complete conversion of the starting material was observed by TLC (Rf AcOEt/hexane, 7:3 (v:v): 0.34). The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 40% to 60% AcOEt) to give 517 mg of [**2a**] (44% yield) as a white solid. HRMS (ESI+) calcd. for $C_{20}H_{30}N_4O_6$ $[M+H]^+$ (m/z): 423.2238, found: 423.2234. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.49 (br s, 2H, N<u>H</u>COCH$_2$), 7.76 (s, 1H, CH_{Ar}), 7.29 (d, *J* = 8.0 Hz, 2H, CH_{Ar}), 7.24 (t, *J* = 7.9 Hz, 1H, CH_{Ar}), 5.53 (br s, 2H, N<u>H</u>Boc), 3.92 (d, *J* = 5.7 Hz, 4H, CH_2), 1.47 (s, 18H, CH_3). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 168.2 (2 x CO), 156.7 (2 x CO), 138.2 (2 x C_{Ar}), 129.7 (1 x CH_{Ar}), 116.0 (2 x CH_{Ar}), 111.5 (1 x CH_{Ar}), 80.9 (2 x C), 45.7 (2 x CH_2), 28.5 (6 x CH_3).

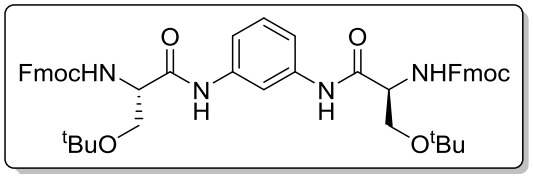


[**2b**]: to a solution of Fmoc-L-Asn(Trt)-OH (4.24 g, 7.11 mmol) in dry DMF (15 mL), HOBt (1.25 g, 9.27 mmol) and DCCD (2.23 g, 10.8 mmol) were added under inert atmosphere of Ar. The resulting mixture was cooled down to 0ºC in an ice-water bath. Then, a solution of *m*-phenylenediamine (334 mg, 3.09 mmol) in dry DMF (10 mL) was added via cannula under inert atmosphere of Ar. The mixture was stirred at room temperature for 60 hours, after which complete conversion of the starting material was observed by TLC (Rf AcOEt/hexane, 1:1 (v:v): 0.58). The mixture was filtered, and the filtrate was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 50% AcOEt) to give 2.05 g of [**2b**] (52% yield) as a white solid. HRMS (ESI+) calcd. for $C_{82}H_{68}N_6O_8$ $[M+H]^+$ (m/z): 1265.5171, found: 1265.5183. ^{1}H NMR (400 MHz, $CDCl_3$):

δ = 8.77 (br s, 2H, N<u>H</u>COC*H), 7.81–7.66 (m, 5H, CH_{Ar}), 7.61–7.51 (m, 4H, CH_{Ar}), 7.38 (t, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.32–7.04 (m, 37H, CH_{Ar}), 6.97 (s, 2H, CON<u>H</u>Trt), 6.54 (br s, 2H, N<u>H</u>Fmoc), 4.68 (br s, 2H, C*H), 4.51–4.32 (m, 4H, $COOCH_2$), 4.20 (t, *J* = 7.0 Hz, 2H, CH), 3.16 (d, *J* = 15.7 Hz, 2H, C<u>H</u>$_2$C*H), 2.66 (dd, *J* = 15.7, 6.9 Hz, 2H, C<u>H</u>$_2$C*H). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 170.8 (2 x CO), 169.0 (2 x CO), 156.4 (2 x CO), 144.2 (6 x C_{Ar}), 143.8 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.0 (2 x C_{Ar}), 129.3 (1 x CH_{Ar}), 128.7 (12 x CH_{Ar}), 128.2 (12 x CH_{Ar}), 127.9 (4 x CH_{Ar}), 127.3 (6 x CH_{Ar}), 127.3 (4 x CH_{Ar}), 125.3 (4 x CH_{Ar}), 120.1 (4 x CH_{Ar}), 116.3 (2 x CH_{Ar}), 111.7 (1 x CH_{Ar}), 71.2 (2 x C), 67.5 (2 x COO<u>C</u>H_2), 52.2 (2 x C*H), 47.2 (2 x CH), 38.9 (2 x <u>C</u>H_2C*H).

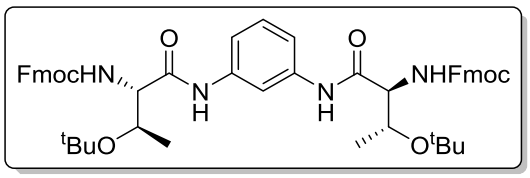


[**2c**]: this compound was obtained as described above for [**2b**], starting from Fmoc-L-Gln(Trt)-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 25% to 40% AcOEt, Rf AcOEt/hexane, 3:2 (v:v): 0.50) to give 1.12 g of [**2c**] (47% yield) as a white solid. HRMS (ESI+) calcd. for $C_{84}H_{72}N_6O_8$ $[M+H]^+$ (m/z): 1293.5484, found: 1293.5472. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.84 (s, 2H, N<u>H</u>COC*H), 7.89 (s, 1H, CH_{Ar}), 7.75 (d, *J* = 7.1 Hz, 4H, CH_{Ar}), 7.61–7.52 (m, 4H, CH_{Ar}), 7.38 (t, *J* = 7.1 Hz, 4H, CH_{Ar}), 7.31–7.18 (m, 34H, CH_{Ar}), 7.10 (t, *J* = 8.0 Hz, 1H, CH_{Ar}), 7.03 (s, 2H, N<u>H</u>Trt), 6.98 (d, *J* = 7.4 Hz, 2H, CH_{Ar}), 6.09 (d, *J* = 4.8 Hz, 2H, N<u>H</u>Fmoc), 4.43–4.30 (m, 4H, $COOCH_2$), 4.20 (t, *J* = 7.1 Hz, 2H, CH), 4.17–4.08 (m, 2H, C*H), 2.67–2.55 (m, 2H, C*HCH$_2$C<u>H</u>$_2$), 2.50–2.38 (m, 2H, C*HCH$_2$C<u>H</u>$_2$), 2.19–2.08 (m, 2H, C*HC<u>H</u>$_2$), 2.03–1.89 (m, 2H, C*HC<u>H</u>$_2$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 172.5 (2 x CO), 169.5 (2 x CO), 156.5 (2 x CO), 144.5 (6 x C_{Ar}), 143.9 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.2 (2 x C_{Ar}), 128.8 (13 x CH_{Ar}), 128.2 (12 x CH_{Ar}), 127.8 (4 x CH_{Ar}), 127.2 (10 x CH_{Ar}), 125.3 (4 x CH_{Ar}), 120.1 (4 x CH_{Ar}), 115.9 (2 x CH_{Ar}), 111.6 (1 x CH_{Ar}), 71.0 (2 x C), 67.2 (2 x COO<u>C</u>H_2), 54.4 (2 x C*H), 47.3 (2 x CH), 34.0 (2 x C*HCH$_2$<u>C</u>H$_2$), 30.4 (2 x C*H<u>C</u>H$_2$).

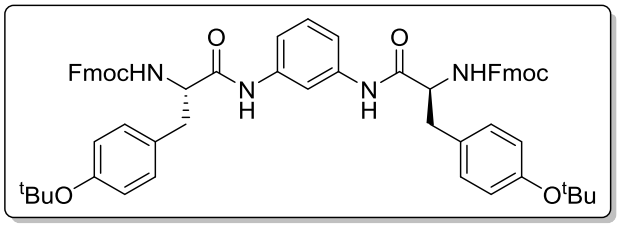


[**2d**]: this compound was obtained as described above for [**2b**], starting from Fmoc-L-Ser(^{t}Bu)-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 25% to 40% AcOEt, Rf AcOEt/hexane, 3:2 (v:v): 0.83) to give 1.05 g of [**2d**] (45% yield) as a white solid. HRMS (ESI+) calcd. for $C_{50}H_{54}N_4O_8$

$[M+H]^+$ (m/z): 839.4014, found: 839.4029. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.80 (br s, 2H, NHCOC*H), 7.96 (s, 1H, CH_{Ar}), 7.77 (d, *J* = 7.6 Hz, 4H, CH_{Ar}), 7.62 (d, *J* = 7.1 Hz, 4H, CH_{Ar}), 7.41 (t, *J* = 7.4 Hz, 4H, CH_{Ar}), 7.32 (t, *J* = 7.8 Hz, 4H, CH_{Ar}), 7.29–7.20 (m, 3H, CH_{Ar}), 5.87 (br s, 2H, C*HNHCO), 4.44 (d, *J* = 7.0 Hz, 4H, $COOCH_2$), 4.35 (br s, 2H, C*H), 4.25 (t, *J* = 6.9 Hz, 2H, CH), 3.92 (br s, 2H, C*HCH_2), 3.45 (t, *J* = 8.7 Hz, 2H, C*HCH_2), 1.28 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 168.5 (2 x CO), 156.2 (2 x CO), 143.9 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.9 (1 x CH_{Ar}), 127.9 (4 x CH_{Ar}), 127.2 (4 x CH_{Ar}), 125.2 (4 x CH_{Ar}), 120.2 (4 x CH_{Ar}), 115.5 (2 x CH_{Ar}), 111.0 (1 x CH_{Ar}), 75.1 (2 x C), 67.3 (2 x COOCH_2), 61.9 (2 x C*HCH_2), 54.8 (2 x C*H), 47.3 (2 x CH), 27.6 (6 x CH_3).

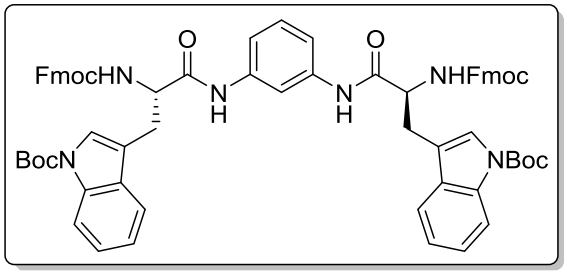


[**2e**]: this compound was obtained as described above for [**2b**], starting from Fmoc-L-Thr(^{t}Bu)-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 25% to 40% AcOEt, Rf AcOEt/hexane, 3:7 (v:v): 0.46) to give 1.61 g of [**2e**] (67% yield) as a white solid. HRMS (ESI–) calcd. for $C_{52}H_{58}N_4O_8$ $[M+HCOO]^-$ (m/z): 911.4237, found: 911.4254. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.24 (s, 2H, NHCOC*H), 7.92 (s, 1H, CH_{Ar}), 7.78 (d, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.63 (d, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.41 (t, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.37–7.19 (m, 7H, CH_{Ar}), 6.12 (d, *J* = 4.9 Hz, 2H, C*HNHCO), 4.49–4.21 (m, 10H, 4H x CH_2 + 2H x CH + 2H x C*HNH + 2H x C*HCH$_3$), 1.38 (s, 18H, C(CH_3)$_3$), 1.10 (d, *J* = 6.3 Hz, 6H, C*HCH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 167.6 (2 x CO), 156.2 (2 x CO), 143.8 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.8 (1 x CH_{Ar}), 127.9 (4 x CH_{Ar}), 127.2 (4 x CH_{Ar}), 125.3 (4 x CH_{Ar}), 120.2 (4 x CH_{Ar}), 115.3 (2 x CH_{Ar}), 110.9 (1 x CH_{Ar}), 76.3 (2 x C), 67.2 (2 x CH_2), 67.1 (2 x C*HCH$_3$), 59.1 (2 x C*HNH), 47.3 (2 x CH), 28.3 (6 x C(CH_3)$_3$), 16.9 (2 x C*HCH_3).

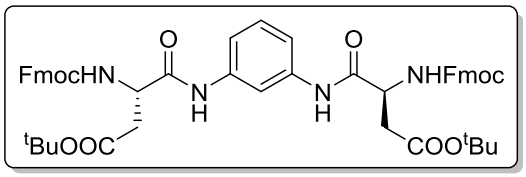


[**2f**]: this compound was obtained as described above for [**2b**], starting from Fmoc-L-Tyr(^{t}Bu)-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 25% to 40% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.46) to give 1.46 g of [**2f**] (53% yield) as a white solid. HRMS (ESI+) calcd. for $C_{62}H_{62}N_4O_8$ $[M+H]^+$ (m/z):

991.4640, found: 991.4622. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.91 (br s, 2H, NHCOC*H), 7.74 (d, *J* = 7.6 Hz, 4H, CH_{Ar}), 7.60–7.48 (m, 5H, CH_{Ar}), 7.37 (t, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.31–7.22 (m, 4H, CH_{Ar}), 7.14–6.99 (m, 7H, CH_{Ar}), 6.86 (d, *J* = 8.4 Hz, 4H, CH_{Ar}), 5.57 (br s, 2H, C*HNHCO), 4.50 (br s, 2H, C*H), 4.44–4.25 (m, 4H, $COOCH_2$), 4.19 (t, *J* = 6.9 Hz, 2H, CH), 3.14–2.93 (m, 4H, C*HCH_2), 1.26 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 169.5 (2 x CO), 156.5 (2 x CO), 154.6 (2 x C_{Ar}), 143.7 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 137.8 (2 x C_{Ar}), 131.1 (2 x C_{Ar}), 129.9 (4 x CH_{Ar}), 129.5 (1 x CH_{Ar}), 127.9 (4 x CH_{Ar}), 127.3 (4 x CH_{Ar}), 125.2 (4 x CH_{Ar}), 124.6 (4 x CH_{Ar}), 120.1 (4 x CH_{Ar}), 116.1 (2 x CH_{Ar}), 111.7 (1 x CH_{Ar}), 78.7 (2 x C), 67.4 (2 x $COOCH_2$), 57.3 (2 x C*H), 47.2 (2 x CH), 38.0 (2 x C*HCH_2), 28.9 (6 x CH_3).

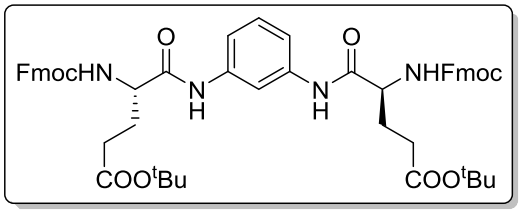


[**2g**]: this compound was obtained as described above for [**2b**], starting from Fmoc-L-Trp(Boc)-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 35% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.59) to give 1.33 g of [**2g**] (43% yield) as a white solid. HRMS (ESI–) calcd. for $C_{68}H_{64}N_6O_{10}$ $[M+HCOO]^-$ (m/z): 1169.4666, found:1169.5189. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.23 (br s, 2H, NHCOC*H), 8.08 (br s, 2H, CH_{Ar}), 7.71 (d, *J* = 7.6 Hz, 4H, CH_{Ar}), 7.62–7.40 (m, 9H, CH_{Ar}), 7.34 (t, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.29–6.98 (m, 11H, CH_{Ar}), 5.72 (br s, 2H, C*HNHCO), 4.67 (br s, 2H, C*H), 4.33 (br s, 4H, $COOCH_2$), 4.18–4.10 (m, 2H, $COOCH_2CH$), 3.30–3.06 (m, 4H, C*HCH_2), 1.56 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 169.7 (2 x CO), 156.6 (2 x CO), 149.6 (2 x CO), 143.7 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 137.8 (2 x C_{Ar}), 135.6 (2 x C_{Ar}), 130.2 (2 x C_{Ar}), 129.4 (1 x CH_{Ar}), 127.8 (4 x CH_{Ar}), 127.2 (4 x CH_{Ar}), 125.2 (4 x CH_{Ar}), 124.8 (2 x CH_{Ar}), 124.6 (2 x CH_{Ar}), 122.9 (2 x CH_{Ar}), 120.1 (4 x CH_{Ar}), 119.1 (2 x CH_{Ar}), 116.3 (2 x CH_{Ar}), 115.5 (2 x CH_{Ar}), 115.3 (2 x C_{Ar}), 112.0 (1 x CH_{Ar}), 83.9 (2 x C), 67.5 (2 x $COOCH_2$), 55.8 (2 x C*H), 47.1 (2 x CH), 28.2 (6 x CH_3), 28.1 (2 x C*HCH_2).

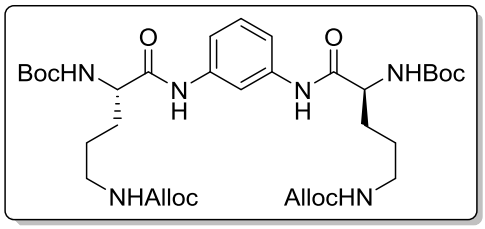


[**2h**]: this compound was obtained as described above for [**2b**], starting from Fmoc-L-Asp(^{t}Bu)-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 25% to 40% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.34)

to give 978 mg of [**2h**] (42% yield) as a white solid. HRMS (ESI+) calcd. for $C_{52}H_{54}N_4O_{10}$ $[M+Na]^+$ (m/z): 917.3732, found: 917.3764. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.57 (s, 2H, NHCOC*H), 7.80 (s, 1H, CH_{Ar}), 7.76 (d, *J* = 7.4 Hz, 4H, CH_{Ar}), 7.59 (d, *J* = 6.3 Hz, 4H, CH_{Ar}), 7.39 (t, *J* = 7.3 Hz, 4H, CH_{Ar}), 7.34 – 7.20 (m, 7H, CH_{Ar}), 6.11 (d, *J* = 7.4 Hz, 2H, C*HNHCO), 4.66 (br s, 2H, C*H), 4.45 (d, *J* = 6.4 Hz, 4H, $COOCH_2$), 4.23 (t, *J* = 6.9 Hz, 2H, CH), 2.96 (d, *J* = 16.0 Hz, 2H, C*HCH_2), 2.69 (dd, *J* = 17.0, 6.7 Hz, 2H, C*HCH_2), 1.45 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 171.5 (2 x CO), 168.7 (2 x CO), 156.4 (2 x CO), 143.8 (4 x C_{Ar}), 141.5 (4 x C_{Ar}), 138.2 (2 x C_{Ar}), 129.7 (1 x CH_{Ar}), 128.0 (4 x CH_{Ar}), 127.3 (4 x CH_{Ar}), 125.2 (4 x CH_{Ar}), 120.2 (4 x CH_{Ar}), 116.1 (2 x CH_{Ar}), 111.5 (1 x CH_{Ar}), 82.5 (2 x C), 67.5 (2 x $COOCH_2$), 51.9 (2 x C*H), 47.3 (2 x CH), 37.5 (2 x C*HCH_2), 28.2 (6 x CH_3).

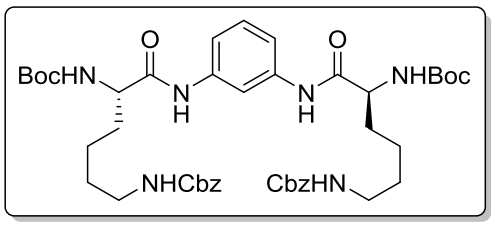


[**2i**]: this compound was obtained as described above for [**2b**], starting from Fmoc-L-Glu(^{t}Bu)-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 40% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.43) to give 1.79 g of [**2i**] (70% yield) as a white solid. HRMS (ESI+) calcd. for $C_{54}H_{58}N_4O_{10}$ $[M+H]^+$ (m/z): 923.4226, found: 923.4225. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.62 (s, 2H, NHCOC*H), 7.84 (s, 1H, CH_{Ar}), 7.75 (d, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.59 (t, *J* = 6.8 Hz, 4H, CH_{Ar}), 7.38 (t, *J* = 7.4 Hz, 4H, CH_{Ar}), 7.33–7.16 (m, 7H, CH_{Ar}), 5.94 (d, *J* = 7.1 Hz, 2H, C*HNHCO), 4.38–4.28 (m, 6H, 4H x $COOCH_2$ + 2H x C*H), 4.21 (t, *J* = 7.0 Hz, 2H, CH), 2.59 –2.46 (m, 2H, C*HCH$_2$CH_2), 2.43–2.30 (m, 2H, C*HCH$_2$CH_2), 2.22–2.09 (m, 2H, C*HCH_2), 2.04–1.93 (m, 2H, C*HCH_2), 1.46 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 173.4 (2 x CO), 169.7 (2 x CO), 156.7 (2 x CO), 143.8 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.2 (2 x C_{Ar}), 129.6 (1 x CH_{Ar}), 127.9 (4 x CH_{Ar}), 127.2 (4 x CH_{Ar}), 125.2 (4 x CH_{Ar}), 120.1 (4 x CH_{Ar}), 115.9 (2 x CH_{Ar}), 111.4 (1 x CH_{Ar}), 81.6 (2 x C), 67.4 (2 x $COOCH_2$), 55.2 (2 x C*H), 47.2 (2 x CH), 32.1 (2 x C*HCH$_2$CH_2), 28.3 (2 x C*HCH_2), 28.2 (6 x CH_3).

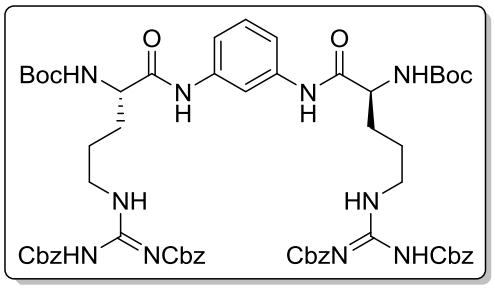


[**2j**]: this compound was obtained as described above for [**2a**], starting from Boc-L-Orn(Alloc)-OH. In this case HBTU was used instead of EDC·HCl. The residue was purified by flash

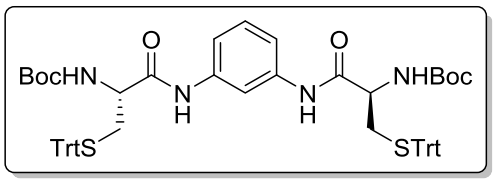
chromatography using AcOEt/hexane as eluent (from 45% to 55% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.23) to give 1.43 g of [**2j**] (85% yield) as a white solid. HRMS (ESI+) calcd. for $C_{34}H_{52}N_6O_{10}$ $[M+H]^+$ (m/z): 705.3818, found: 705.3813. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.92 (br s, 2H, NHCOC*H), 7.77 (br s, 1H, CH_{Ar}), 7.36–6.97 (m, 3H, CH_{Ar}), 5.89 (ddt, *J* = 17.2, 10.8, 5.6 Hz, 2H, $NHCOOCH_2CHCH_2$), 5.66 (br s, 2H, C*HNHCO), 5.28 (dq, *J* = 17.2, 1.6 Hz, 2H, $NHCOOCH_2CHCH_2$), 5.22–5.07 (m, 4H, 2H x $NHCOOCH_2CHCH_2$ + 2H x NHAlloc), 4.58 (d, *J* = 5.5 Hz, 4H, $NHCOOCH_2CHCH_2$), 4.43 (br s, 2H, C*H), 3.40 (br s, 2H, $CH_2NHAlloc$), 3.24–3.02 (m, 2H, $CH_2NHAlloc$), 1.96–1.77 (m, 2H, $C*HCH_2$), 1.76–1.54 (m, 6H, 2H x $C*HCH_2$ + 4H x $C*HCH_2CH_2$), 1.43 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 171.3 (2 x CO), 157.2 (2 x CO), 156.4 (2 x CO), 138.5 (2 x C_{Ar}), 133.0 (2 x $NHCOOCH_2CHCH_2$), 129.4 (1 x CH_{Ar}), 117.7 (2 x $NHCOOCH_2CHCH_2$), 115.7 (2 x CH_{Ar}), 111.5 (1 x CH_{Ar}), 80.3 (2 x C), 65.9 (2 x $NHCOOCH_2CHCH_2$), 53.9 (2 x C*H), 38.9 (2 x $CH_2NHAlloc$), 30.3 (2 x $C*HCH_2$), 28.5 (6 x CH_3), 26.6 (2 x $C*HCH_2CH_2$).



[**2k**]: this compound was obtained as described above for [**2a**], starting from Boc-L-Lys(Cbz)-OH. In this case EDC·HCl was replaced by HBTU and no HOBt was used. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 40% to 50% AcOEt, Rf AcOEt/hexane, 3:2 (v:v): 0.41) to give 1.77 g of [**2k**] (56% yield) as a white solid. HRMS (ESI+) calcd. for $C_{44}H_{60}N_6O_{10}$ $[M+H]^+$ (m/z): 833.4444, found: 833.4453. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.95 (br s, 2H, NHCOC*H), 7.73 (br s, 1H, CH_{Ar}), 7.55–7.16 (m, 12H, CH_{Ar}), 7.16–6.97 (m, 1H, CH_{Ar}), 5.72 (br s, 2H, C*HNHCO), 5.33–4.84 (m, 6H, 4H x $NHCOOCH_2$ + 2H x NHCbz), 4.24 (br s, 2H, C*H), 3.16 (br s, 4H, CH_2NHCbz), 2.00–1.59 (m, 4H, C*H CH_2), 1.58–1.18 (m, 26H, 4H x CH_2CH_2NHCbz + 4H x $C*HCH_2CH_2$ + 18H x CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 171.5 (2 x CO), 156.8 (2 x CO), 156.4 (2 x CO), 138.5 (2 x C_{Ar}), 136.7 (2 x C_{Ar}), 129.5 (1 x CH_{Ar}), 128.6 (4 x CH_{Ar}), 128.2 (6 x CH_{Ar}), 115.7 (2 x CH_{Ar}), 111.3 (1 x CH_{Ar}), 80.4 (2 x C), 66.8 (2 x $NHCOOCH_2$), 55.4 (2 x C*H), 40.7 (2 x CH_2NHCbz), 32.2 (2 x $C*HCH_2$), 29.5 (2 x CH_2CH_2NHCbz), 28.5 (6 x CH_3), 22.9 (2 x $C*HCH_2CH_2$).

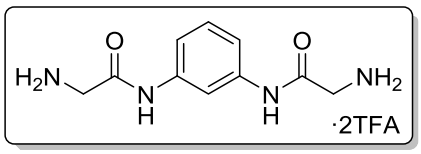


[**2l**]: this compound was obtained as described above for [**2a**], starting from Boc-L-Arg(Cbz)$_2$-OH. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 45% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.27) to give 360 mg of [**2l**] (44% yield) as a white solid. HRMS (ESI+) calcd. for $C_{60}H_{72}N_{10}O_{14}$ $[M+H]^+$ (m/z): 1157.5302, found: 1157.5328. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.62–9.11 (m, 4H, NH), 8.46 (br s, 2H, N<u>H</u>COC*H), 7.57 (s, 1H, CH_{Ar}), 7.44–7.19 (m, 20H, CH_{Ar}), 7.11–6.95 (m, 3H, CH_{Ar}), 5.71 (d, *J* = 8.0 Hz, 2H, C*HN<u>H</u>CO), 5.22 (s, 4H, COOC<u>H</u>$_2$), 5.10 (ABq, δ_A = 5.16, δ_B = 5.05, *J* = 12.5 Hz, 4H, $COOCH_2$), 4.35 (br s, 2H, C*H), 4.14–4.00 (m, 2H, C<u>H</u>$_2$NH), 3.99–3.83 (m, 2H, C<u>H</u>$_2$NH), 1.91–1.60 (m, 8H, 4H x C*HC<u>H</u>$_2$ + 4H x C*CH$_2$C<u>H</u>$_2$), 1.44 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 170.5 (2 x CO), 160.4 (2 x C-guanidine), 155.8 (6 x CO), 138.3 (2 x C_{Ar}), 134.4 (4 x C_{Ar}), 129.3 (1 x CH_{Ar}), 129.0 (8 x CH_{Ar}), 128.6 (8 x CH_{Ar}), 128.3 (4 x CH_{Ar}), 116.3 (2 x CH_{Ar}), 112.2 (1 x CH_{Ar}), 80.2 (2 x C), 69.7 (2 x COO<u>C</u>H$_2$), 67.9 (2 x COO<u>C</u>H$_2$), 54.5 (2 x C*H), 44.8 (2 x CH_2NH), 29.2 (2 x C*H<u>C</u>H$_2$), 28.5 (6 x CH_3), 24.9 (2 x C*HCH$_2$<u>C</u>H$_2$).

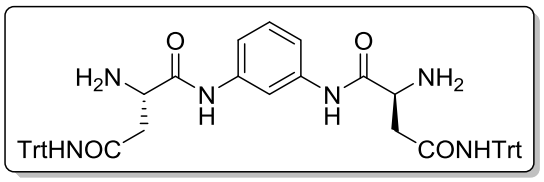


[**2m**]: this compound was obtained as described above for [**2a**], starting from Boc-L-Cys(Trt)-OH. In this case dry DCM was used instead of dry DMF as solvent. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 25% to 40% AcOEt, Rf AcOEt/hexane, 1:2 (v:v): 0.39) to give 1.49 g of [**2m**] (29% yield) as a white solid. HRMS (ESI+) calcd. for $C_{60}H_{62}N_4O_6S_2$ $[M+H]^+$ (m/z): 999.4184, found: 999.4145. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.02 (br s, 2H, N<u>H</u>COC*H), 7.64 (s, 1H, CH_{Ar}), 7.44 (d, *J* = 7.1 Hz, 12H, CH_{Ar}), 7.30 (t, *J* = 7.6 Hz, 12H, CH_{Ar}), 7.25–7.13 (m, 9H, CH_{Ar}), 4.81 (br s, 2H, C*HN<u>H</u>CO), 3.94 (br s, 2H, C*H), 2.85–2.70 (m, 2H, CH_2), 2.62 (dd, *J* = 13.2, 5.1 Hz, 2H, CH_2), 1.43 (s, 18H, CH_3). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 168.9 (2 x CO), 156.0 (2 x CO), 144.5 (6 x C_{Ar}), 138.1 (2 x C_{Ar}), 129.7 (12 x CH_{Ar}), 129.5 (1 x CH_{Ar}), 128.2 (12 x CH_{Ar}), 127.1 (6 x CH_{Ar}), 115.7 (2 x CH_{Ar}), 111.0 (1 x CH_{Ar}), 80.9 (2 x C), 67.5 (2 x C), 54.4 (2 x C*H), 33.6 (2 x CH_2), 28.4 (6 x CH_3).

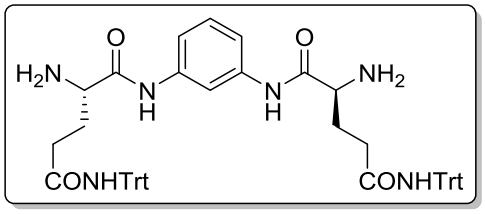
1.5.5. Synthesis of intermediates [3a-l]



[**3a**·2TFA]: to a solution of [**2a**] (139 mg, 0.29 mmol) in DCM (10 mL), TFA (1.5 mL) was added. The mixture was stirred at room temperature for 4 hours and then concentrated under reduced pressure. Diethyl ether was added over the residue and the product was filtered and washed with diethyl ether, finally obtaining 145 mg of [**3a**·2TFA] (quantitative yield) as a white solid. HRMS (ESI+) calcd. for $C_{10}H_{14}N_4O_2$ $[M+H]^+$ (m/z): 223.1190, found: 223.1190. ^{1}H NMR (500 MHz, MeOD-d_4): δ = 8.03 (t, J = 1.7 Hz, 1H, CH_{Ar}), 7.37–7.32 (m, 2H, CH_{Ar}), 7.32–7.28 (m, 1H, CH_{Ar}), 3.85 (s, 4H, CH_2). ^{13}C NMR (75 MHz, MeOD-d_4): δ = 165.5 (2 x CO), 139.8 (2 x C_{Ar}), 130.5 (1 x CH_{Ar}), 116.8 (2 x CH_{Ar}), 112.4 (1 x CH_{Ar}), 42.1 (2 x CH_2).

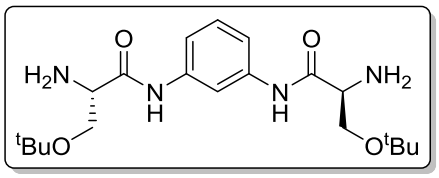


[**3b**]: compound [**2b**] (600 mg, 0.47 mmol) was dissolved in 4.0 mL of 20% piperidine in dry DMF. After several minutes stirring at room temperature the product precipitated as a white solid but the mixture was allowed to react for 4 hours until complete conversion of starting material. Diethyl ether was added over the reaction mixture and the product was filtered off and washed with diethyl ether, obtaining 293 mg of diamine [**3a**] (75% yield) as a white solid. HRMS (ESI+) calcd. for $C_{52}H_{48}N_6O_4$ $[M+H]^+$ (m/z): 821.3810, found: 821.3832. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 7.93 (s, 1H, CH_{Ar}), 7.38–7.31 (m, 2H, CH_{Ar}), 7.30–7.10 (m, 31H, CH_{Ar}), 3.77 (dd, J = 7.5, 5.5 Hz, 2H, C*H), 2.77 (dd, J = 15.3, 5.5 Hz, 2H, CH_2), 2.68 (dd, J = 15.3, 7.6 Hz, 2H, CH_2). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 174.6 (2 x CO), 172.5 (2 x CO), 145.9 (6 x C_{Ar}), 140.0 (2 x C_{Ar}), 130.1 (1 x CH_{Ar}), 130.0 (12 x CH_{Ar}), 128.7 (12 x CH_{Ar}), 127.8 (6 x CH_{Ar}), 117.0 (2 x CH_{Ar}), 112.8 (1 x CH_{Ar}), 71.7 (2 x C), 54.0 (2 x C*H), 42.3 (2 x CH_2).

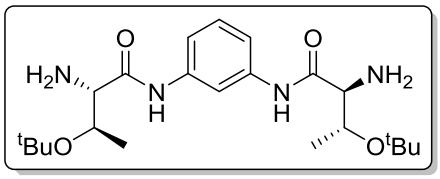


[**3c**]: 531 mg of [**3c**] (white solid, 74% yield) were obtained from [**2c**] as described above for [**3b**]. HRMS (ESI+) calcd. for $C_{54}H_{52}N_6O_4$ $[M+H]^+$ (m/z): 849.4123, found: 849.4135. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.46 (s, 2H, NHCOC*H), 7.82 (t, J = 1.8 Hz, 1H, CH_{Ar}), 7.35–7.19 (m, 33H, CH_{Ar}), 6.93 (s, 2H, NHTrt), 3.40 (t, J = 6.5 Hz, 2H, C*H), 2.53–2.45 (m, 4H, C*HCH$_2$CH$_2$), 2.13–1.94 (m,

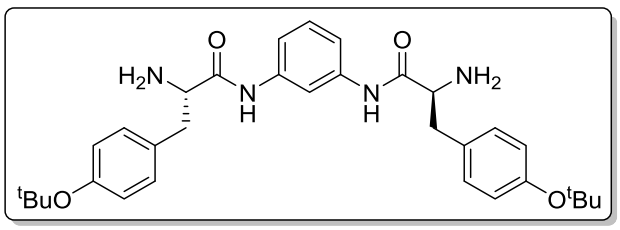
4H, C*HC<u>H</u>$_2$), 1.68 (br s, 4H, NH_2). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 173.3 (2 x CO), 171.8 (2 x CO), 144.7 (6 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.7 (1 x CH_{Ar}) , 128.8 (12 x CH_{Ar}), 128.1 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 115.2 (2 x CH_{Ar}), 110.5 (1 x CH_{Ar}), 70.7 (2 x C), 54.8 (2 x C*H), 34.1 (2 x C*HCH$_2$C<u>H</u>$_2$), 31.0 (2 x C*HC<u>H</u>$_2$).



[**3d**]: 522 mg of [**3d**] (white solid, quantitative yield) were obtained from [**2d**] as described above for [**3b**]. HRMS (ESI+) calcd. for $C_{20}H_{34}N_4O_4$ $[M+H]^+$ (m/z): 395.2653, found: 395.2672. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.54 (s, 2H, NH), 7.91 (t, *J* = 2.1 Hz, 1H, CH_{Ar}), 7.39–7.35 (m, 2H, CH_{Ar}), 7.29–7.23 (m, 1H, CH_{Ar}), 3.67 (dd, *J* = 7.2, 3.3 Hz, 2H, CH_2), 3.62–3.55 (m, 4H, 2H x C*H + 2H x CH_2), 2.00 (br s, 4H, NH_2), 1.21 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 171.6 (2 x CO), 138.6 (2 x C_{Ar}), 129.6 (1 x CH_{Ar}), 115.0 (2 x CH_{Ar}), 110.4 (1 x CH_{Ar}), 73.8 (2 x C), 63.8 (2 x CH_2), 56.0 (2 x C*H), 27.7 (6 x CH_3).

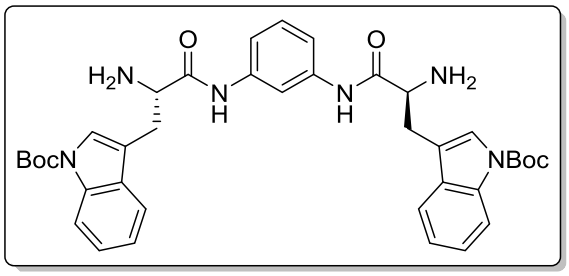


[**3e**]: 445 mg of [**3e**] (white solid, 64% yield) were obtained from [**2e**] as described above for [**3b**]. HRMS (ESI+) calcd. for $C_{22}H_{38}N_4O_4$ $[M+H]^+$ (m/z): 423.2966, found: 423.2956. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.63 (s, 2H, N<u>H</u>COC*H), 7.82 (t, *J* = 2.1 Hz, 1H, CH_{Ar}), 7.40–7.33 (m, 2H, CH_{Ar}), 7.32–7.23 (m, 1H, CH_{Ar}), 4.23 (qd, *J* = 6.3, 2.5 Hz, 2H, C*<u>H</u>CH_3), 3.25 (d, *J* = 2.5 Hz, 2H, C*<u>H</u>NH_2), 1.95 (br s, 4H, NH_2), 1.21 (d, *J* = 6.3 Hz, 6H, C*HC<u>H</u>$_3$), 1.17 (s, 18H, C(C<u>H</u>$_3$)$_3$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 172.3 (2 x CO), 138.7 (2 x C_{Ar}), 129.6 (1 x CH_{Ar}), 115.0 (2 x CH_{Ar}), 110.4 (1 x CH_{Ar}), 74.4 (2 x C), 67.8 (2 x <u>C</u>*HCH_3), 60.4 (2 x C*HNH_2), 28.6 (6 x C(<u>C</u>H_3)$_3$), 20.5 (2 x C*H<u>C</u>H_3).

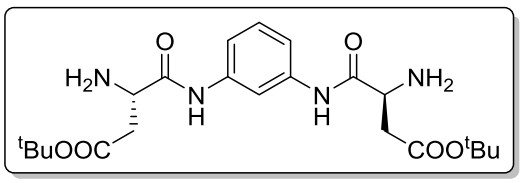


[**3f**]: 812 g of [**3f**] (white solid, quantitative yield) were obtained from [**2f**] as described above for [**3b**]. HRMS (ESI+) calcd. for $C_{32}H_{42}N_4O_4$ $[M+H]^+$ (m/z): 547.3279, found: 547.3280. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 9.48 (s, 2H, N<u>H</u>COC*H), 7.92 (t, *J* = 2.1 Hz, 1H, CH_{Ar}), 7.40 (dd, *J* = 8.1, 2.0 Hz, 2H, CH_{Ar}), 7.32–7.25 (m, 1H, CH_{Ar}), 7.14 (d, *J* = 8.4 Hz, 4H, CH_{Ar}), 6.95 (d, *J* = 8.4 Hz, 4H, CH_{Ar}), 3.70 (dd, *J* = 9.6, 3.8 Hz, 2H, C*H), 3.32 (dd, *J* = 14.0, 3.8 Hz, 2H, C*HC<u>H</u>$_2$), 2.72 (dd, *J* = 13.9, 9.7 Hz, 2H,

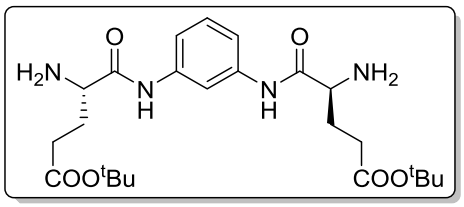
C*HC<u>H</u>$_2$), 1.99 (br s, 4H, NH_2), 1.33 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 172.8 (2 x CO), 154.5 (2 x C_{Ar}), 138.4 (2 x C_{Ar}), 132.5 (2 x C_{Ar}), 129.8 (4 x CH_{Ar}), 129.7 (1 x CH_{Ar}), 124.6 (4 x CH_{Ar}), 115.2 (2 x CH_{Ar}), 110.4 (1 x CH_{Ar}), 78.6 (2 x C), 57.0 (2 x C*H), 40.2 (2 x CH_2), 29.0 (6 x CH_3).



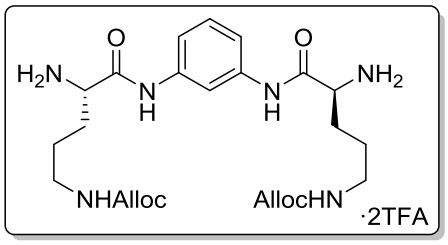
[**3g**]: 781 mg of [**3g**] (white solid, quantitative yield) were obtained from [**2g**] as described above for [**3b**]. HRMS (ESI+) calcd. for $C_{38}H_{44}N_6O_6$ $[M+H]^+$ (m/z): 681.3395, found: 681.3399. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.57 (s, 2H, N<u>H</u>COC*H), 8.14 (d, *J* = 8.3 Hz, 2H, CH_{Ar}), 7.94 (t, *J* = 2.1 Hz, 1H, CH_{Ar}), 7.65 (dd, *J* = 7.7, 1.0 Hz, 2H, CH_{Ar}), 7.50 (s, 2H, CH_{Ar}), 7.43 (dd, *J* = 7.8, 2.1 Hz, 2H, CH_{Ar}), 7.38–7.22 (m, 5H, CH_{Ar}), 3.84 (dd, *J* = 9.8, 3.6 Hz, 2H, C*H), 3.49 (ddd, *J* = 14.7, 3.7, 1.2 Hz, 2H, C*HC<u>H</u>$_2$), 2.88 (dd, *J* = 14.8, 9.6 Hz, 2H, C*HC<u>H</u>$_2$), 1.66 (s, 18H, CH_3), 1.58 (s, 4H, NH_2). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 172.7 (2 x CO), 149.7 (2 x CO), 138.4 (2 x C_{Ar}), 135.8 (2 x C_{Ar}), 130.3 (2 x C_{Ar}), 129.7 (1 x CH_{Ar}), 124.9 (2 x CH_{Ar}), 124.3 (2 x CH_{Ar}), 122.9 (2 x CH_{Ar}), 119.3 (2 x CH_{Ar}), 116.7 (2 x CH_{Ar}), 115.5 (2 x CH_{Ar}), 115.2 (2 x C_{Ar}), 110.5 (1 x CH_{Ar}), 83.9 (2 x C), 55.4 (2 x C*H), 30.6 (2 x CH_2), 28.3 (6 x CH_3).



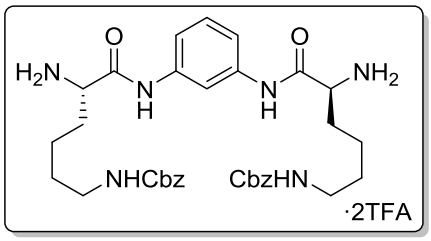
[**3h**]: 531 mg of [**3h**] (white solid, quantitative yield) were obtained from [**2h**] as described above for [**3b**]. HRMS (ESI+) calcd. for $C_{22}H_{34}N_4O_6$ $[M+H]^+$ (m/z): 451.2551, found: 451.2560. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 7.92 (t, *J* = 2.0 Hz, 1H, CH_{Ar}), 7.38–7.34 (m, 2H, CH_{Ar}), 7.29–7.24 (m, 1H, CH_{Ar}), 3.75 (dd, *J* = 6.7, 6.0 Hz, 2H, C*H), 2.74 (dd, *J* = 16.2, 6.0 Hz, 2H, CH_2), 2.62 (dd, *J* = 16.2, 6.7 Hz, 2H, CH_2), 1.43 (s, 18H, CH_3). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 174.5 (2 x CO), 172.1 (2 x CO), 140.0 (2 x C_{Ar}), 130.2(1 x CH_{Ar}), 116.9 (2 x CH_{Ar}), 112.8 (1 x CH_{Ar}), 82.3 (2 x C), 53.6 (2 x C*H), 41.5 (2 x CH_2), 28.3 (6 x CH_3).



[**3i**]: 1.05 g of [**3i**] (white solid, quantitative yield) were obtained from [**2i**] as described above for [**3b**]. HRMS (ESI+) calcd. for $C_{24}H_{38}N_4O_6$ $[M+H]^+$ (m/z): 479.2864, found: 479.2882. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 7.94 (t, J = 2.0 Hz, 1H, CH_{Ar}), 7.37–7.33 (m, 2H, CH_{Ar}), 7.30–7.23 (m, 1H, CH_{Ar}), 3.45 (dd, J = 7.2, 6.1 Hz, 2H, C*H), 2.44–2.30 (m, 4H, C*HCH$_2$C$\underline{H}_2$), 2.07–1.96 (m, 2H, C*HC$\underline{H}_2$), 1.92–1.80 (m, 2H, C*HC$\underline{H}_2$), 1.43 (s, 18H, CH_3). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 175.5 (2 x CO), 174.2 (2 x CO), 140.0 (2 x C_{Ar}), 130.2 (1 x CH_{Ar}), 117.1(2 x CH_{Ar}), 113.1 (1 x CH_{Ar}), 81.7 (2 x C), 56.1 (2 x C*H), 32.7 (2 x C*HCH$_2$C$\underline{H}_2$), 31.5 (2 x C*HC$\underline{H}_2$), 28.3 (6 x CH_3).

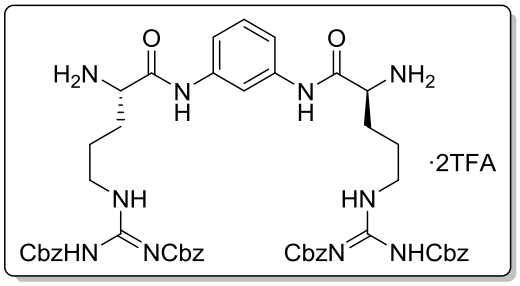


[**3j**·2TFA]: 1.27 g of [**3j**·2TFA] (white solid, 94% yield) were obtained from [**2j**] as described above for [**3a**·2TFA]. HRMS (ESI+) calcd. for $C_{24}H_{36}N_6O_6$ $[M+H]^+$ (m/z): 505.2769, found: 505.2786. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 8.08 (t, J = 1.8 Hz, 1H, CH_{Ar}), 7.40–7.35 (m, 2H, CH_{Ar}), 7.34–7.29 (m, 1H, CH_{Ar}), 5.90 (ddt, J = 17.3, 10.6, 5.4 Hz, 2H, NHCOOCH$_2$C$\underline{H}$CH$_2$), 5.27 (dd, J = 17.3, 1.7 Hz, 2H, NHCOOCH$_2$CHC$\underline{H}_2$), 5.15 (dd, J = 10.3, 1.0 Hz, 2H, NHCOOCH$_2$CHC$\underline{H}_2$), 4.52 (dt, J = 5.4, 1.5 Hz, 4H, NHCOOC$\underline{H}_2$CHCH$_2$), 4.02 (t, J = 6.5 Hz, 2H, C*H), 3.18 (td, J = 6.8, 1.7 Hz, 4H, C$\underline{H}_2$NHAlloc), 2.0 –1.85 (m, 4H, C*HC$\underline{H}_2$), 1.74–1.55 (m, 4H, C*HCH$_2$C$\underline{H}_2$). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 168.5 (2 x CO), 159.0 (2 x CO), 139.7 (2 x C_{Ar}), 134.4 (2 x NHCOOCH$_2$C$\underline{H}$CH$_2$), 130.5 (1 x CH_{Ar}), 117.5 (2 x NHCOOCH$_2$CH$\underline{C}$H$_2$), 117.3 (2 x CH_{Ar}), 112.9 (1 x CH_{Ar}), 66.4 (2 x NHCOO$\underline{C}$H$_2$CHCH$_2$), 54.8 (2 x C*H), 40.8 (2 x $\underline{C}$H$_2$NHAlloc), 30.1 (2 x C*HC$\underline{H}_2$), 26.6 (2 x C*HCH$_2$C$\underline{H}_2$).



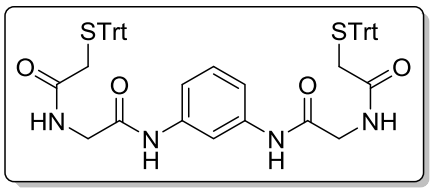
[**3k**·2TFA]: 474 mg of [**3k**·2TFA] (white solid, 93% yield) were obtained from [**2k**] as described above for [**3a**·2TFA]. HRMS (ESI+) calcd. for $C_{34}H_{44}N_6O_6$ $[M+H]^+$ (m/z): 633.3395, found: 633.3389. ^{1}H NMR (500 MHz, MeOD-d_4): δ = 8.10 (t, J = 2.0 Hz, 1H, CH_{Ar}), 7.41–7.24 (m, 13H, CH_{Ar}), 5.01 (ABq, δ_A = 5.04, δ_B = 4.99, J = 12.5 Hz, 4H, NHCOOC$\underline{H}_2$), 3.95 (t, J = 6.5 Hz, 2H, C*H), 3.13 (t, J = 6.8 Hz, 4H, C$\underline{H}_2$NHCbz), 2.04–1.84 (m, 4H, C*HC$\underline{H}_2$), 1.56 (quint, J = 7.0 Hz, 4H,

C<u>H</u>$_2$CH$_2$NHCbz), 1.51–1.38 (m, 4H, C*HCH$_2$C<u>H</u>$_2$). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 168.6 (2 x CO), 159.0 (2 x CO), 139.7 (2 x C$_{Ar}$), 138.3 (2 x C$_{Ar}$), 130.5 (1 x CH$_{Ar}$), 129.4 (4 x CH$_{Ar}$), 128.9 (2 x CH$_{Ar}$), 128.7 (4 x CH$_{Ar}$), 117.4 (2 x CH$_{Ar}$), 113.0 (1 x CH$_{Ar}$), 67.4 (2 x NHCOO<u>C</u>H$_2$), 55.1 (2 x C*H), 41.1 (2 x <u>C</u>H$_2$NHCbz), 32.3 (2 x C*H<u>C</u>H$_2$), 30.5 (2 x <u>C</u>H$_2$CH$_2$NHCbz), 23.0 (2 x C*HCH$_2$<u>C</u>H$_2$).



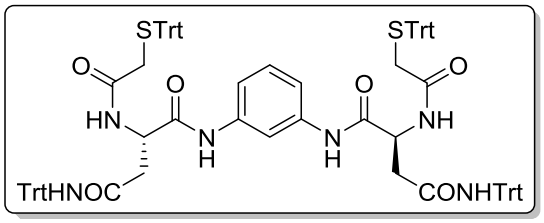
[**3l**·2TFA]: 332 mg of [**3l**·2TFA] (white solid, 89% yield) were obtained from [**2l**] as described above for [**3a**·2TFA]. HRMS (ESI+) calcd. for $C_{50}H_{56}N_{10}O_{10}$ $[M+H]^+$ (m/z): 957.4254, found: 957.4248. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 8.04 (s, 1H, CH$_{Ar}$), 7.46–7.18 (m, 23H, CH$_{Ar}$), 5.22 (s, 4H, COOCH$_2$), 5.09 (ABq, δ_A = 5.11, δ_B = 5.07, J = 12.4 Hz, 4H, COOCH$_2$), 4.04 (t, J = 6.6 Hz, 2H, C*H), 3.95 (t, J = 7.3 Hz, 4H, C<u>H</u>$_2$NH), 2.08–1.68 (m, 8H, 4H x C*HC<u>H</u>$_2$ + 4H x C*CH$_2$C<u>H</u>$_2$). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 168.4 (2 x CO), 161.9 (2 x C-guanidine), 156.7 (2 x CO), 139.7 (2 x CO), 138.2 (2 x C$_{Ar}$), 136.3 (4 x C$_{Ar}$), 130.6 (1 x CH$_{Ar}$), 129.8 (8 x CH$_{Ar}$), 129.5 (8 x CH$_{Ar}$), 129.2 (4 x CH$_{Ar}$), 117.4 (2 x CH$_{Ar}$), 113.0 (1 x CH$_{Ar}$), 70.3 (2 x COO<u>C</u>H$_2$), 68.1 (2 x COO<u>C</u>H$_2$), 54.5 (2 x C*H), 44.9 (2 x CH$_2$NH), 29.1 (2 x C*H<u>C</u>H$_2$), 24.9 (2 x C*HCH$_2$<u>C</u>H$_2$).

1.5.6. Synthesis of intermediates [4a-l]

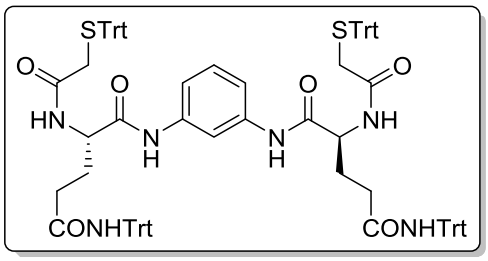


[**4a**]: tritylsulfanyl acetic acid (501 mg, 1.50 mmol) was dissolved in dry DMF (20 mL) and EDC·HCl (312 mg, 1.63 mmol), HOBt (228 mg, 1.69 mmol) and DIPEA (1.6 mL, 4.59 mmol) were added over the solution. The reaction mixture was cooled down to 0ºC in an ice-water bath and [**3a**·2TFA] (322 mg, 0.715 mmol) was added over the mixture. The mixture was stirred at room temperature under an inert atmosphere of Ar for 48 hours, and the formation of the product was followed by TLC (Rf AcOEt/hexane, 7:3 (v:v): 0.31). The mixture was diluted with DCM, washed with saturated aqueous NaHCO$_3$ and saturated aqueous NaCl, and dried under reduced pressure. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 40% to 60% AcOEt) to give 442 mg of [**4a**] (73% yield) as a white solid. HRMS (ESI–) calcd. for $C_{52}H_{46}N_4O_4S_2$ $[M–H]^–$ (m/z): 853.2888,

found: 853.2878. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.41 (s, 2H, N<u>H</u>COCH$_2$NH), 7.65 (s, 1H, CH_{Ar}), 7.42 (d, *J* = 7.5 Hz, 12H, CH_{Ar}), 7.29–7.21 (m, 15H, CH_{Ar}), 7.19 (t, *J* = 7.3 Hz, 6H, CH_{Ar}), 6.78 (t, *J* = 4.9 Hz, 2H, NHCOCH$_2$N<u>H</u>), 3.66 (d, *J* = 5.4 Hz, 4H, NHCOC<u>H</u>$_2$NH), 3.20 (s, 4H, C<u>H</u>$_2$STrt). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 169.7 (2 x CO), 167.0 (2 x CO), 144.0 (6 x C_{Ar}), 138.3 (2 x C_{Ar}), 129.6 (13 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 116.0 (2 x CH_{Ar}), 111.1 (1 x CH_{Ar}), 68.1 (2 x C), 45.0 (2 x NHCO<u>C</u>H$_2$NH), 35.8 (2 x <u>C</u>H$_2$STrt).

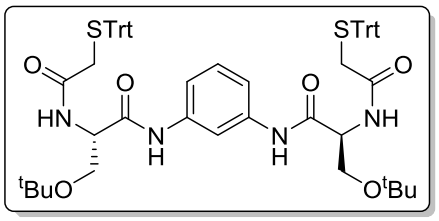


[**4b**]: this compound was obtained as described above for [**4a**], starting from [**3b**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 40% to 60% AcOEt, Rf AcOEt/hexane, 1:1 (v:v): 0.46) to give 343 mg of [**4b**] (73% yield) as a white solid. HRMS (ESI+) calcd. for $C_{94}H_{80}N_6O_6S_2$ $[M+H]^+$ (m/z): 1453.5654, found: 1453.5665. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.84 (s, 2H, N<u>H</u>COC*H), 7.60 (t, *J* = 2.0 Hz, 1H, CH_{Ar}), 7.54–7.00 (m, 65H, 2H x C*HN<u>H</u>CO + 63H x CH_{Ar}), 6.91 (s, 2H, N<u>H</u>Trt), 4.49 (td, *J* = 7.5, 3.0 Hz, 2H, C*H), 3.05 (ABq, δ_A = 3.08, δ_B = 3.02, *J* = 15.7 Hz, 4H, C<u>H</u>$_2$STrt), 2.96–2.83 (m, 2H, C<u>H</u>$_2$C*H), 2.39 (dd, *J* = 15.7, 7.8 Hz, 2H, C<u>H</u>$_2$C*H). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 170.7 (2 x CO), 169.0 (2 x CO), 168.3 (2 x CO), 144.2 (6 x C_{Ar}), 144.1 (6 x C_{Ar}), 138.11 (2 x C_{Ar}), 129.7 (12 x CH_{Ar}), 129.2 (1 x CH_{Ar}), 128.7 (12 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 128.2 (12 x CH_{Ar}), 127.3 (6 x CH_{Ar}), 127.1 (6 x CH_{Ar}), 116.3 (2 x CH_{Ar}), 111.8 (1 x CH_{Ar}), 71.1 (2 x C), 67.9 (2 x C), 50.7 (2 x C*H), 38.4 (2 x <u>C</u>H$_2$C*H), 36.2 (2 x <u>C</u>H$_2$STrt).

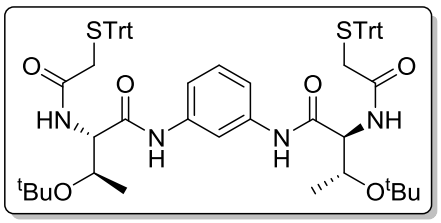


[**4c**]: this compound was obtained as described above for [**4a**], starting from [**3c**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 40% AcOEt, Rf AcOEt/hexane, 1:1 (v:v): 0.23) to give 593 mg of [**4c**] (66% yield) as a white solid. HRMS (ESI+) calcd. for $C_{96}H_{84}N_6O_6S_2$ $[M+H]^+$ (m/z): 1481.5967, found: 1481.5916. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.76 (s, 2H, N<u>H</u>COC*H), 7.73 (t, *J* = 1.9 Hz, 1H, CH_{Ar}), 7.44–7.38 (m, 12H, CH_{Ar}), 7.31–6.96 (m, 55H, 51H x CH_{Ar} + 2H x C*HN<u>H</u>CO + 2H x N<u>H</u>Trt), 4.03 (q, *J* = 6.8 Hz, 2H, C*H), 3.06 (ABq, δ_A = 3.07, δ_B = 3.05, *J* = 15.8Hz, 4H, C<u>H</u>$_2$STrt), 2.59–2.49 (m, 2H,

C*HCH$_2$C$\underline{H}_2$), 2.39–2.29 (m, 2H, C*HCH$_2$C$\underline{H}_2$), 2.03–1.92 (m, 2H, C*HC$\underline{H}_2$), 1.81–1.70 (m, 2H, C*HC$\underline{H}_2$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 172.5 (2 x CO), 168.8 (2 x CO), 168.6 (2 x CO), 144.5 (6 x C_{Ar}), 144.1 (6 x C_{Ar}), 138.2 (2 x C_{Ar}), 129.7 (12 x CH_{Ar}), 129.2 (1 x CH_{Ar}), 128.8 (12 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 128.1 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 127.1(6 x CH_{Ar}), 115.9 (2 x CH_{Ar}), 111.4 (1 x CH_{Ar}), 70.9 (2 x C), 68.0 (2 x C), 53.1 (2 x C*H), 36.3 (2 x $\underline{C}H_2$STrt), 34.2 (2 x C*HCH$_2$$\underline{C}H_2$), 30.3 (2 x C*H$\underline{C}H_2$).

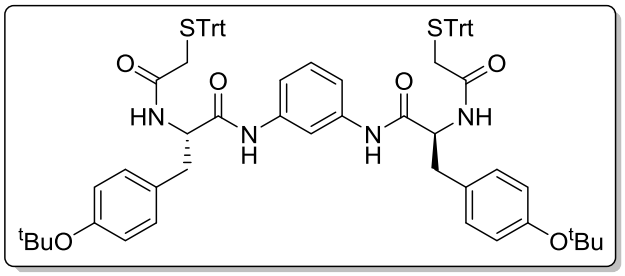


[**4d**]: this compound was obtained as described above for [**4a**], starting from [**3d**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 35% to 45% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.27) to give 612 mg of [**4d**] (51% yield) as a white solid. HRMS (ESI+) calcd. for $C_{62}H_{66}N_4O_6S_2$ $[M+H]^+$ (m/z): 1027.4497, found: 1027.4492. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.68 (s, 2H, N$\underline{H}$COC*H), 7.79 (t, *J* = 1.8 Hz, 1H, CH_{Ar}), 7.43 (d, *J* = 7.3 Hz, 12H, CH_{Ar}), 7.28 (t, *J* = 7.6 Hz, 12H, CH_{Ar}), 7.25–7.18 (m, 9H, CH_{Ar}), 7.10 (d, *J* = 5.8 Hz, 2H, C*HN$\underline{H}$CO), 4.24–4.18 (dt, *J* = 9.7, 4.6 Hz, 2H, C*H), 3.71 (dd, *J* = 8.6, 4.3 Hz, 2H, C*HC$\underline{H}_2$), 3.20–3.08 (m, 6H, 2H x C*HC$\underline{H}_2$ + 4H x C$\underline{H}_2$STrt), 1.22 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 168.7 (2 x CO), 168.2 (2 x CO), 144.1 (6 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.8 (1 x CH_{Ar}), 129.7 (12 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.1 (6 x CH_{Ar}), 115.5 (2 x CH_{Ar}), 110.9 (1 x CH_{Ar}), 75.0 (2 x C), 68.0 (2 x C), 61.0 (2 x C*H$\underline{C}H_2$), 53.5 (2 x C*H), 36.2 (2 x $\underline{C}H_2$STrt), 27.6 (6 x CH_3).

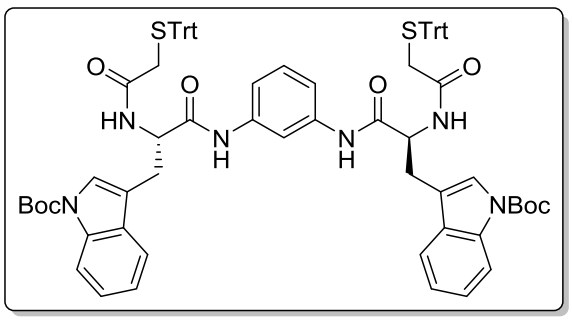


[**4e**]: this compound was obtained as described above for [**4a**], starting from [**3e**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 20% to 35% AcOEt, Rf AcOEt/hexane, 3:7 (v:v): 0.28) to give 608 mg of [**4e**] (57% yield) as a white solid. HRMS (ESI–) calcd. for $C_{64}H_{70}N_4O_6S_2$ $[M+HCOO]^-$ (m/z): 1099.4719, found: 1099.4722. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 9.16 (s, 2H, N$\underline{H}$COC*H), 7.70 (s, 1H, CH_{Ar}), 7.43 (d, *J* = 7.5 Hz, 12H, CH_{Ar}), 7.33–7.17 (m, 23H, 21H x CH_{Ar} + 2H x C*HN$\underline{H}$CO), 4.24–4.13 (m, 4H, 2H x C*$\underline{H}$NH + 2H x C*$\underline{H}$CH$_3$), 3.07 (ABq, δ_A = 3.10, δ_B = 3.03, *J* = 15.4 Hz, 4H, CH_2), 1.33 (s, 18H, C(C$\underline{H}_3$)$_3$), 0.94 (d, *J* = 6.4 Hz, 6H, C*HC$\underline{H}_3$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 168.5 (2 x CO), 167.4 (2 x CO), 144.1 (6 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.9 (1 x CH_{Ar}), 129.7 (12

x CH_{Ar}), 128.2 (12 x CH_{Ar}), 127.1 (6 x CH_{Ar}), 115.3 (2 x CH_{Ar}), 110.7 (1 x CH_{Ar}), 76.3 (2 x C), 67.9 (2 x C), 66.2 (2 x $\underline{C}$*HCH_3), 58.3 (2 x C*HNH), 36.6 (2 x CH_2), 28.3 (6 x $C(\underline{C}H_3)_3$), 17.0 (2 x C*H$\underline{C}H_3$).

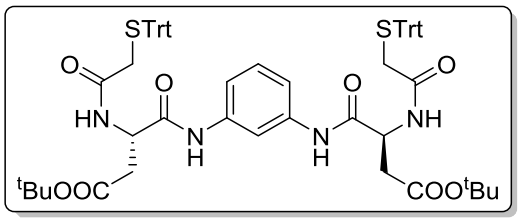


[**4f**]: this compound was obtained as described above for [**4a**], starting from [**3f**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 40% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.39) to give 1.07 g of [**4f**] (79% yield) as a white solid. HRMS (ESI–) calcd. for $C_{74}H_{74}N_4O_6S_2$ $[M+HCOO]^-$ (m/z): 1223.5032, found: 1223.4956. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.73 (s, 2H, N$\underline{H}$COC*H), 7.51 (t, *J* = 2.0 Hz, 1H, CH_{Ar}), 7.36 (d, *J* = 7.2 Hz, 12H, CH_{Ar}), 7.29–7.22 (m, 14H, CH_{Ar}), 7.18 (t, *J* = 7.3 Hz, 6H, CH_{Ar}), 7.14–7.09 (m, 1H, CH_{Ar}), 7.06 (d, *J* = 8.4 Hz, 4H, CH_{Ar}), 6.86 (d, *J* = 8.4 Hz, 4H, CH_{Ar}), 6.61 (d, *J* = 7.3 Hz, 2H, C*HN$\underline{H}$CO), 4.37 (q, *J* = 7.1 Hz, 2H, C*H), 3.12 (ABq, δ_A = 3.15, δ_B = 3.09, *J* = 16.3 Hz, 4H, C$\underline{H_2}$STrt), 2.91 (d, *J* = 7.1 Hz, 4H, C*HC$\underline{H_2}$), 1.28 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 169.0 (2 x CO), 168.4 (2 x CO), 154.6 (2 x C_{Ar}), 144.0 (6 x C_{Ar}), 138.0 (2 x C_{Ar}), 131.1 (2 x C_{Ar}), 130.0 (4 x CH_{Ar}), 129.6 (12 x CH_{Ar}), 129.4 (1 x CH_{Ar}), 128.4 (12 x CH_{Ar}), 127.3 (6 x CH_{Ar}), 124.6 (4 x CH_{Ar}), 115.9 (2 x CH_{Ar}), 111.3 (1 x CH_{Ar}), 78.6 (2 x C), 68.2 (2 x C), 55.9 (2 x C*H), 36.9 (2 x C*H$\underline{C}H_2$), 36.0 (2 x $\underline{C}H_2$STrt), 29.0 (6 x CH_3).

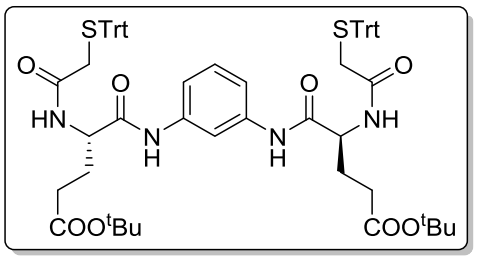


[**4g**]: this compound was obtained as described above for [**4a**], starting from [**3g**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 40% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.54) to give 977 mg of [**4g**] (88% yield) as a white solid. HRMS (ESI+) calcd. for $C_{80}H_{76}N_6O_8S_2$ $[M+Na]^+$ (m/z): 1335.5058, found: 1335.5060. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.10 (d, *J* = 8.4 Hz, 2H, CH_{Ar}), 7.94 (s, 2H, N$\underline{H}$COC*H), 7.55 (d, *J* = 7.8 Hz, 2H, CH_{Ar}), 7.46 (s, 1H, CH_{Ar}), 7.40 (s, 2H, CH_{Ar}), 7.34 (d, *J* = 7.4 Hz, 12H, CH_{Ar}), 7.29–7.02 (m, 25H, CH_{Ar}), 6.61 (d, *J* = 7.2 Hz, 2H, C*HN$\underline{H}$CO), 4.52 (q, *J* = 7.0 Hz, 2H, C*H), 3.21–2.98 (m, 8H, 4H x C$\underline{H_2}$STrt + 4H x C*HC$\underline{H_2}$), 1.59 (s, 18H, CH_3). ^{13}C NMR (101 MHz,

$CDCl_3$): δ = 169.2 (2 x CO), 168.5 (2 x CO), 149.6 (2 x CO), 144.0 (6 x C_{Ar}), 138.0 (2 x C_{Ar}), 135.6 (2 x C_{Ar}), 130.2 (2 x C_{Ar}), 129.6 (12 x CH_{Ar}), 129.4 (1 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 124.9 (2 x CH_{Ar}), 124.5 (2 x CH_{Ar}), 123.0 (2 x CH_{Ar}), 119.2 (2 x CH_{Ar}), 116.1 (2 x CH_{Ar}), 115.4 (2 x CH_{Ar}), 115.3 (2 x C_{Ar}), 111.7 (1 x CH_{Ar}), 83.8 (2 x C), 68.2 (2 x C), 54.5 (2 x C*H), 36.0 (2 x CO<u>C</u>H_2), 28.3 (6 x CH_3), 27.0 (2 x C*H<u>C</u>H_2).

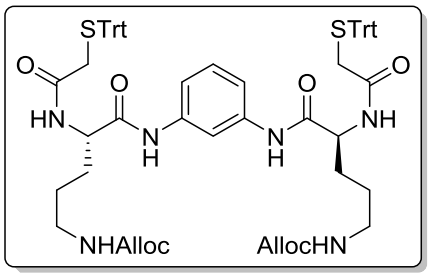


[**4h**]: this compound was obtained as described above for [**4a**], starting from [**3h**]. The residude was purified by flash chromatography using AcOEt/hexane as eluent (from 25% to 45% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.30) to give 644 mg of [**4h**] (69% yield) as a white solid. HRMS (ESI+) calcd. for $C_{64}H_{66}N_4O_8S_2$ $[M+Na]^+$ (m/z): 1105.4214, found: 1105.4236. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.54 (s, 2H, N<u>H</u>COC*H), 7.63 (s, 1H, CH_{Ar}), 7.40 (d, *J* = 7.9 Hz, 12H, CH_{Ar}), 7.31–7.16 (m, 23H, 21H x CH_{Ar} + 2H x C*HN<u>H</u>CO), 4.54 (td, *J* = 7.6, 4.0 Hz, 2H, C*H), 3.16 (ABq, δ_A = 3.18, δ_B = 3.14, *J* = 16.3 Hz, 4H, C<u>H</u>$_2$STrt), 2.69 (dd, *J* = 17.2, 3.9 Hz, 2H, C*HC<u>H</u>$_2$), 2.44 (dd, *J* = 17.1, 7.8 Hz, 2H, C*HC<u>H</u>$_2$), 1.43 (s, 18H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 171.3 (2 x CO), 169.0 (2 x CO), 168.0 (2 x CO), 144.0 (6 x C_{Ar}), 138.3 (2 x C_{Ar}), 129.6 (13 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.26 (6 x CH_{Ar}), 116.0 (2 x CH_{Ar}), 111.3 (1 x CH_{Ar}), 82.3 (2 x C), 68.2 (2 x C), 50.4 (2 x C*H), 36.6 (2 x CH_2), 36.0 (2 x CH_2), 28.2 (6 x CH_3).

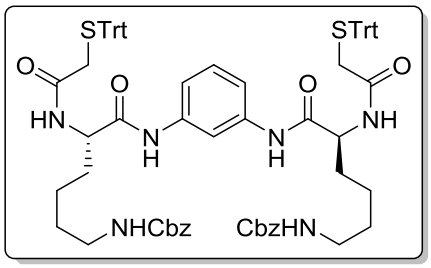


[**4i**]: this compound was obtained as described above for [**4a**], starting from [**3i**]. The residude was purified by flash chromatography using AcOEt/hexane as eluent (from 30% to 40% AcOEt, Rf AcOEt/hexane, 2:3 (v:v): 0.20) to give 285 mg of [**4i**] (74% yield) as a white solid. HRMS (ESI+) calcd. for $C_{66}H_{70}N_4O_8S_2$ $[M+H]^+$ (m/z): 1111.4708, found: 1111.4696. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.72 (s, 2H, N<u>H</u>COC*H), 7.78 (t, *J* = 1.8 Hz, 1H, CH_{Ar}), 7.44–7.37 (m, 12H, CH_{Ar}), 7.30–7.14 (m, 21H, CH_{Ar}), 6.80 (d, *J* = 7.2 Hz, 2H, C*HN<u>H</u>CO), 4.26 (q, *J* = 6.8 Hz, 2H, C*H), 3.14 (ABq, δ_A = 3.16, δ_B = 3.12, *J* = 16.1Hz, 4H, C<u>H</u>$_2$STrt), 2.45–2.34 (m, 2H, C*HCH$_2$C<u>H</u>$_2$), 2.28–2.17 (m, 2H, C*HCH$_2$C<u>H</u>$_2$), 2.08–1.96 (m, 2H, C*HC<u>H</u>$_2$), 1.83–1.71 (m, 2H, C*HC<u>H</u>$_2$), 1.44 (s, 18H, CH_3). ^{13}C NMR (101 MHz,

$CDCl_3$): δ = 173.1 (2 x CO), 169.0 (2 x CO), 168.8 (2 x CO), 144.0 (6 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.6 (12 x CH_{Ar}), 129.5 (1 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 115.7 (2 x CH_{Ar}), 111.0 (1 x CH_{Ar}), 81.3 (2 x C), 68.1 (2 x C), 53.5 (2 x C*H), 36.1 (2 x $\underline{C}H_2$STrt), 32.0 (2 x C*HCH$_2$$\underline{C}H_2$), 28.2 (6 x CH_3), 27.8 (2 x C*H$\underline{C}H_2$).

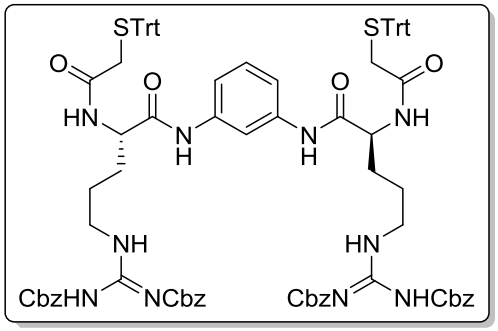


[**4j**]: this compound was obtained as described above for [**4a**], starting from [**3j**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 50% to 75% AcOEt, Rf AcOEt/hexane, 4:1 (v:v): 0.51) to give 281 mg of [**4j**] (15% yield) as a white solid. HRMS (ESI+) calcd. for $C_{66}H_{68}N_6O_8S_2$ $[M+Na]^+$ (m/z): 1159.4432, found: 1159.4454. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.75 (br s, 2H, N$\underline{H}$COC*H), 7.74 (br s, 1H, CH_{Ar}), 7.42–7.35 (m, 12H, CH_{Ar}), 7.33–7.23 (m, 15H, CH_{Ar}), 7.22–7.14 (m, 6H, CH_{Ar}), 6.79 (d, *J* = 6.1 Hz, 2H, C*HN$\underline{H}$CO), 5.86 (ddt, *J* = 16.3, 10.7, 5.5 Hz, 2H, NHCOOCH$_2$C$\underline{H}$CH$_2$), 5.25 (dd, *J* = 17.2, 1.6 Hz, 2H, NHCOOCH$_2$CHC$\underline{H}_2$), 5.14 (dd, *J* = 10.6, 1.4 Hz, 2H, NHCOOCH$_2$CHC$\underline{H}_2$), 5.02 (br s, 2H, N$\underline{H}$Alloc), 4.56 (d, *J* = 5.2 Hz, 4H, NHCOOC$\underline{H}_2$CHCH$_2$), 4.52 (s, 2H, C*H), 3.41 (br s, 2H, C$\underline{H}_2$NHAlloc), 3.21–2.96 (m, 6H, 2H x C$\underline{H}_2$NHAlloc + 4H x C$\underline{H}_2$STrt), 1.89–1.66 (m, 2H, C*HC$\underline{H}_2$), 1.58–1.35 (m, 6H, 2H x C*HC$\underline{H}_2$ + 4H x C*HCH$_2$C$\underline{H}_2$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 169.9 (2 x CO), 169.1 (2 x CO), 157.1 (2 x CO), 144.1 (6 x C_{Ar}), 138.5 (2 x C_{Ar}), 133.0 (2 x NHCOOCH$_2$$\underline{C}HCH_2$), 129.7 (12 x CH_{Ar}), 129.4 (1 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 117.7 (2 x NHCOOCH$_2$CH$\underline{C}$H$_2$), 115.6 (2 x CH_{Ar}), 111.0 (1 x CH_{Ar}), 68.0 (2 x C), 65.8 (2 x NHCOO$\underline{C}$H$_2$CHCH$_2$), 52.9 (2 x C*H), 39.7 (2 x $\underline{C}H_2$NHAlloc), 36.3 (2 x $\underline{C}H_2$STrt), 30.0 (2 x C*H$\underline{C}H_2$), 26.4 (2 x C*HCH$_2$$\underline{C}H_2$).



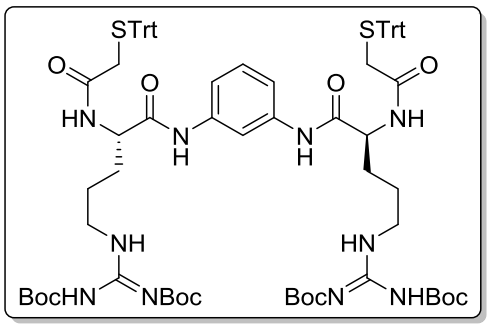
[**4k**]: this compound was obtained as described above for [**4a**], starting from [**3k**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 50% to 65% AcOEt, Rf AcOEt/hexane, 3:2 (v:v): 0.47) to give 1.70 g of [**4k**] (83% yield) as a white solid. HRMS (ESI+) calcd. for $C_{76}H_{76}N_6O_8S_2$ $[M+H]^+$ (m/z): 1265.5239, found: 1265.5187. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.64 (br s, 2H, N$\underline{H}$COC*H), 7.62 (br s, 1H, CH_{Ar}), 7.47–7.03 (m, 43H, CH_{Ar}), 6.60 (d, *J* = 7.4 Hz, 2H, C*HN$\underline{H}$CO), 5.03 (s, 4H, NHCOOC$\underline{H}_2$), 4.97 (br

s, 2H, N<u>H</u>Cbz), 4.23 (br s, 2H, C*H), 3.22–2.99 (m, 8H, 4H x C<u>H</u>$_2$NHCbz + 4H x C<u>H</u>$_2$STrt), 1.92–1.64 (m, 4H, C*HC<u>H</u>$_2$), 1.57–1.34 (m, 4H, C<u>H</u>$_2$CH$_2$NHCbz), 1.33–1.09 (m, 4H, C*HCH$_2$C<u>H</u>$_2$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 169.7 (2 x CO), 169.2 (2 x CO), 156.6 (2 x CO), 144.1 (6 x C_{Ar}), 138.4 (2 x C_{Ar}), 136.8 (2 x C_{Ar}), 129.6 (13 x CH_{Ar}), 128.6 (4 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 128.2 (6 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 115.9 (2 x CH_{Ar}), 111.4 (1 x CH_{Ar}), 68.1 (2 x C), 66.7 (2 x NHCOO<u>C</u>H$_2$), 54.3 (2 x C*H), 40.6 (2 x <u>C</u>H$_2$NHCbz), 36.2 (2 x <u>C</u>H$_2$STrt), 31.6 (2 x C*H<u>C</u>H$_2$), 29.5 (2 x <u>C</u>H$_2$CH$_2$NHCbz), 22.7 (2 x C*HCH$_2$<u>C</u>H$_2$).



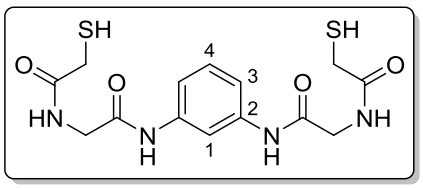
[**4l**]: this compound was obtained as described above for [**4a**], starting from [**3l**]. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 40% to 60% AcOEt, Rf AcOEt/hexane, 1:1 (v:v): 0.57) to give 266 mg of [**4l**] (70% yield) as a white solid. HRMS (ESI+) calcd. for $C_{92}H_{88}N_{10}O_{12}S_2$ $[M+H]^+$ (m/z): 1589.6097, found: 1589.6088. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.64–9.14 (m, 4H, NH), 8.41 (s, 2H, N<u>H</u>COC*H), 7.56 (s, 1H, CH_{Ar}), 7.45–7.09 (m, 53H, CH_{Ar}), 7.03 (d, *J* = 7.7 Hz, 2H, C*HN<u>H</u>CO), 5.22 (s, 4H, COOCH$_2$), 4.98 (ABq, δ_A = 5.07, δ_B = 4.89, *J* = 12.4 Hz, 4H, COOCH$_2$), 4.36 (q, *J* = 6.7 Hz, 2H, C*H), 4.10–3.97 (m, 2H, C<u>H</u>$_2$NH), 3.96–3.80 (m, 2H, C<u>H</u>$_2$NH), 2.96 (s, 4H, C<u>H</u>$_2$STrt), 1.81–1.46 (m, 8H, 4H x C*HC<u>H</u>$_2$ + 4H x C*CH$_2$C<u>H</u>$_2$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 169.3 (2 x CO), 169.0 (2 x CO), 160.8 (2 x C-guanidine), 155.8 (2 x CO), 147.0 (2 x CO), 144.1 (6 x C_{Ar}), 138.3 (2 x C_{Ar}), 134.5 (4 x C_{Ar}), 129.6 (12 x CH_{Ar}), 129.4 (1 x CH_{Ar}), 129.0 (8 x CH_{Ar}), 128.5 (8 x CH_{Ar}), 128.2 (16 x CH_{Ar}), 127.1 (6 x CH_{Ar}), 116.4 (2 x CH_{Ar}), 112.2 (1 x CH_{Ar}), 69.3 (2 x COO<u>C</u>H$_2$), 67.8 (2 x C), 67.5 (2 x COO<u>C</u>H$_2$), 53.8 (2 x C*H), 44.4 (2 x CH$_2$NH), 36.2 (2 x CO<u>C</u>H$_2$), 28.1 (2 x C*H<u>C</u>H$_2$), 25.0 (2 x C*HCH$_2$<u>C</u>H$_2$).

1.5.7. Synthesis of intermediates [5j and 6l]

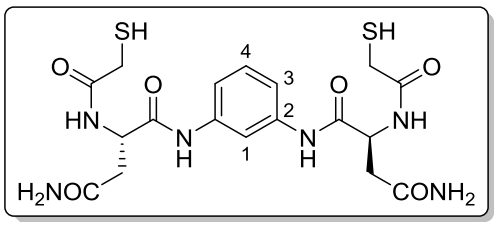


[**6l**]: to a solution of [**4j**] (132 mg, 0.116 mmol) in dry DCM (3.0 mL), $PhSiH_3$ (343 μL, 2.79 mmol) was added under inert atmosphere of Ar. Then a solution of $Pd(PPh_3)_4$ (18 mg, 15 μmol) in dry DCM (2.0 mL) was added. The mixture was stirred at room temperature for 1 hour, after which complete conversion of the starting material was observed by TLC. The crude mixture was filtered through a bed of Celite® and the filtrate was concentrated to dryness under reduced pressure, obtaining diamine [**5j**] as a brownish non-pure solid. HRMS (ESI+) calcd. for $C_{58}H_{60}N_6O_4S_2$ $[M+H]^+$ (m/z): 969.4190, found: 969.4178. The residue was redissolved in DCM (5.0 mL), and triethylamine (54 μL, 0.39 mmol) and *N*,*N*′-di-Boc-*N*″-triflylguanidine[73] (139 mg, 0.355 mmol) were added. The mixture was stirred for 2 hours under inert atmosphere of Ar, until the reaction was completed as evidenced by TLC (Rf AcOEt/Hexane, 2:3 (v:v): 0.30). The mixture was diluted with DCM, washed with 2M aqueous $NaHSO_4$, saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography using AcOEt/hexane as eluent (from 40% to 45% AcOEt) to give 106.7 mg of [**6l**] (63% yield over the last two steps) as a white solid. HRMS (ESI–) calcd. for $C_{80}H_{96}N_{10}O_{12}S_2$ $[M–H]^-$ (m/z): 1451.6578, found: 1451.6572. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.46 (s, 2H, NH), 8.51–8.33 (m, 4H, 2H x NHCOC*H + 2H x NH), 7.67–7.63 (m, 1H, CH_{Ar}), 7.41 (d, *J* = 7.2 Hz, 12H, CH_{Ar}), 7.31–7.16 (m, 21H, CH_{Ar}), 7.08 (d, *J* = 7.7 Hz, 2H, C*HNHCO), 4.34 (q, *J* = 6.8 Hz, 2H, C*H), 3.43 (q, *J* = 5.8 Hz, 4H, CH_2NH), 3.14 (s, 4H, CH_2STrt), 1.90–1.75 (m, 2H, C*HCH_2), 1.72–1.13 (m, 42H, 2H x C*HCH_2 + 4H x C*HCH$_2$CH_2 + 36H x CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 169.3 (2 x CO), 168.9 (2 x CO), 167.5 (2 x C-guanidine), 152.5 (2 x CO), 147.0 (2 x CO), 144.1 (6 x C_{Ar}), 138.6 (2 x C_{Ar}), 129.7 (12 x CH_{Ar}), 129.3 (1 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.2 (6 x CH_{Ar}), 116.2 (2 x CH_{Ar}), 111.8 (1 x CH_{Ar}), 90.1 (2 x C), 81.0 (2 x C), 68.0 (2 x C), 53.1 (2 x C*H), 41.9 (2 x CH_2NH), 36.3 (2 x CH_2STrt), 29.9 (2 x C*HCH_2), 28.1 (12 x CH_3), 25.5 (2 x C*HCH$_2$CH_2).

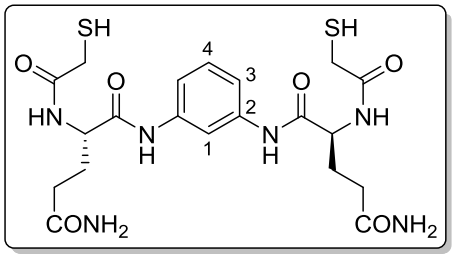
1.5.8. Synthesis of building blocks [1a-m]



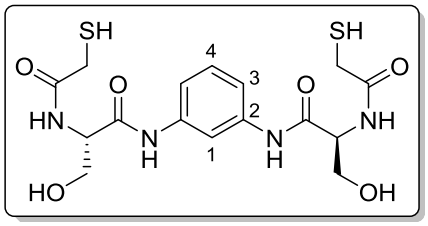
[**1a**]: to a solution of [**4a**] (189 mg, 0.221 mmol) in DCM (1.0 mL), TFA (8.5 mL), TIS (459 µL, 1.77 mmol) and EDT (222 µL, 2.65 mmol) were added rapidly and under stirring. The reaction mixture was stirred at room temperature for 2 hours, after which the solvents were partially evaporated using a N_2 flow. Diethyl ether was added over the reaction mixture and the product was filtered off and washed with diethyl ether, obtaining 59.4 mg of [**1a**] (73% yield) as a white solid. HRMS (ESI+) calcd. for $C_{14}H_{18}N_4O_4S_2$ $[M+H]^+$ (m/z): 371.0847, found: 371.0847. ^{1}H NMR (400 MHz, DMSO-d_6): δ = 9.99 (s, 2H, NHCOCH$_2$NH), 8.35 (t, J = 5.7 Hz, 2H, NHCOCH$_2$NH), 7.90 (s, 1H, H^1), 7.31–7.26 (m, 2H, H^3), 7.25–7.19 (m, 1H, H^4), 3.91 (d, J = 5.7 Hz, 4H, NHCOCH$_2$NH), 3.20 (d, J = 7.9 Hz, 4H, CH$_2$SH), 2.77 (t, J = 8.0 Hz, 2H, SH). ^{13}C NMR (101 MHz, DMSO-d_6): δ = 170.0 (2 x NHCOCH$_2$SH), 167.4 (2 x NHCOCH$_2$NH), 139.1 (2 x C^2), 128.9 (1 x C^4), 114.3 (2 x C^3), 110.2 (1 x C^1), 43.0 (2 x NHCOCH$_2$NH), 26.9 (2 x CH$_2$SH).



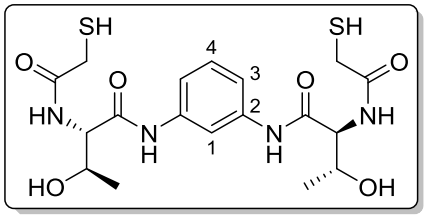
[**1b**]: this compound was obtained as described above for [**1a**], starting from [**4b**]. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 5% to 30% CH_3CN in H_2O) and 37.8 mg of [**1b**] (52% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{18}H_{24}N_6O_6S_2$ $[M+H]^+$ (m/z): 485.1277, found: 485.1279. ^{1}H NMR (400 MHz, DMSO-d_6): δ = 9.97 (s, 2H, NHCOC*H), 8.34 (d, J = 7.7 Hz, 2H, C*HNHCO), 7.93 (t, J = 2.0 Hz, 1H, H^1), 7.36 (s, 2H, NH$_2$), 7.29 (dd, J = 7.6, 2.0 Hz, 2H, H^3), 7.24–7.15 (m, 1H, H^4), 6.91 (s, 2H, NH$_2$), 4.67 (q, J = 7.1 Hz, 2H, C*H), 3.17 (d, J = 7.9 Hz, 4H, CH$_2$SH), 2.73 (t, J = 7.9 Hz, 2H, SH), 2.62–2.41 (m, 4H, CH$_2$C*H). ^{13}C NMR (101 MHz, DMSO-d_6): δ = 171.1 (2 x CONH$_2$), 169.5 (2 x COC*H), 169.4 (2 x COCH$_2$), 139.1 (2 x C^2), 128.6 (1 x C^4), 114.6 (2 x C^3), 110.9 (1 x C^1), 50.9 (2 x C*H), 37.1 (2 x CH$_2$C*H), 27.0 (2 x CH$_2$SH).



[**1c**]: this compound was obtained as described above for [**1a**], starting from [**4c**]. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 5% to 30% CH_3CN in H_2O) and 31.3 mg of [**1c**] (48% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{20}H_{28}N_6O_6S_2$ $[M+H]^+$ (m/z): 513.1590, found: 513.1592. ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.09 (s, 2H, N<u>H</u>COC*H), 8.30 (d, *J* = 7.7 Hz, 2H, C*HN<u>H</u>CO), 7.96 (t, *J* = 2.0 Hz, 1H, H^1), 7.35–7.27 (m, 4H, 2H x H^3 + 2H x NH_2), 7.26–7.20 (m, 1H, H^4), 6.77 (s, 2H, NH_2), 4.39 (td, *J* = 8.0, 5.7 Hz, 2H, C*H), 3.25–3.12 (m, 4H, C<u>H</u>$_2$SH), 2.75 (t, *J* = 8.0 Hz, 2H, SH), 2.22–2.04 (m, 4H, C*HCH$_2$C<u>H</u>$_2$), 1.99–1.88 (m, 2H, C*HC<u>H</u>$_2$), 1.88–1.76 (m, 2H, C*HC<u>H</u>$_2$). ^{13}C NMR (101 MHz, DMSO-d_6): δ = 173.3 (2 x $CONH_2$), 170.0 (2 x <u>C</u>OC*H), 169.6 (2 x <u>C</u>OCH$_2$), 139.0 (2 x C^2), 128.6 (1 x C^4), 114.3 (2 x C^3), 110.4 (1 x C^1), 53.1 (2 x C*H), 31.1 (2 x C*HCH$_2$<u>C</u>H$_2$), 27.8 (2 x C*H<u>C</u>H$_2$), 26.7 (2 x <u>C</u>H$_2$SH).

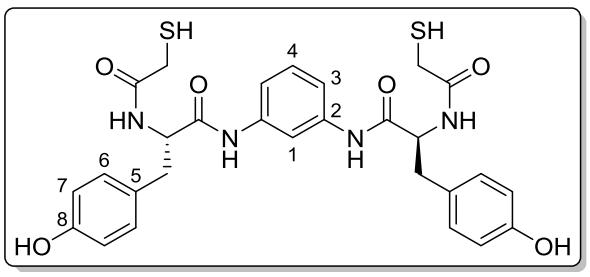


[**1d**]: this compound was obtained as described above for [**1a**], starting from [**4d**]. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 5% to 30% CH_3CN in H_2O) and 42.4 mg of [**1d**] (51% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{16}H_{22}N_4O_6S_2$ $[M+H]^+$ (m/z): 431.1059, found: 431.1054. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 7.92 (t, *J* = 2.0 Hz, 1H, H^1), 7.37–7.30 (m, 2H, H^3), 7.29–7.23 (m, 1H, H^4), 4.56 (t, *J* = 5.3 Hz, 2H, C*H), 3.93–3.82 (m, 4H, C<u>H</u>$_2$OH), 3.28 (s, 4H, C<u>H</u>$_2$SH). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 173.1 (2 x <u>C</u>OCH$_2$), 170.3 (2 x <u>C</u>OC*H), 139.7 (2 x C^2), 129.9 (1 x C^4), 117.3 (2 x C^3), 113.4 (1 x C^1), 62.8 (2 x <u>C</u>H$_2$OH), 57.2 (2 x C*H), 27.9 (2 x <u>C</u>H$_2$SH).

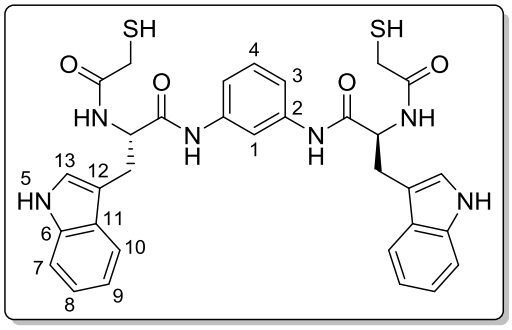


[**1e**]: this compound was obtained as described above for [**1a**], starting from [**4e**]; and 79.6 mg of [**1e**] (92% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{18}H_{26}N_4O_6S_2$ $[M+H]^+$ (m/z): 459.1372, found: 459.1371. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 9.80 (s, 2H, N<u>H</u>COC*H), 8.18 (d, *J* = 8.2 Hz, 2H, C*HN<u>H</u>CO),

7.95–7.89 (m, 1H, H^1), 7.36–7.30 (m, 2H, H^3), 7.29–7.23 (m, 1H, H^4), 4.48–4.42 (m, 2H, C*HNH), 4.24 (qd, *J* = 6.4, 3.9 Hz, 2H, C*HCH$_3$), 3.35–3.29 (m, 4H, CH$_2$), 1.24 (d, *J* = 6.3 Hz, 6H, CH$_3$). ^{13}C NMR (101 MHz, MeOD-*d*$_4$): δ = 173.3 (2 x COCH$_2$), 170.4 (2 x COC*H), 139.5 (2 x C^2), 129.8 (1 x C^4), 117.2 (2 x C^3), 113.3 (1 x C^1), 68.4 (2 x C*HCH$_3$), 60.6 (2 x C*HNH), 28.0 (2 x CH$_2$), 20.0 (2 x CH$_3$).

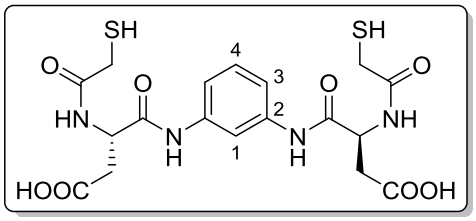


[**1f**]: this compound was obtained as described above for [**1a**], starting from [**4f**], and 63.7 mg of [**1f**] (91% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{28}H_{30}N_4O_6S_2$ $[M+H]^+$ (m/z): 583.1685, found: 583.1696. ^{1}H NMR (400 MHz, DMSO-*d*$_6$): δ = 10.11 (s, 2H, NHCOC*H), 9.17 (br s, 2H, OH), 8.34 (d, *J* = 8.1 Hz, 2H, C*HNHCO), 7.88 (t, *J* = 2.1 Hz, 1H, H^1), 7.32–7.25 (m, 2H, H^3), 7.25–7.18 (m, 1H, H^4), 7.06 (d, *J* = 8.5 Hz, 4H, H^6), 6.64 (d, *J* = 8.5 Hz, 4H, H^7), 4.58 (td, *J* = 8.4, 5.5 Hz, 2H, C*H), 3.12 (d, *J* = 7.9 Hz, 4H, CH$_2$SH), 2.92 (dd, *J* = 13.8, 5.4 Hz, 2H, C*HCH$_2$), 2.75 (dd, *J* = 13.8, 8.9 Hz, 2H, C*HCH$_2$), 2.63 (t, *J* = 7.9 Hz, 2H, SH). ^{13}C NMR (101 MHz, DMSO-*d*$_6$): δ = 170.0 (2 x COC*H), 169.4 (2 x COCH$_2$), 155.8 (2 x C^8), 139.0 (2 x C^2), 130.1 (4 x C^6), 128.8 (1 x C^4), 127.4 (2 x C^5), 114.9 (4 x C^7), 114.6 (2 x C^3), 110.6 (1 x C^1), 55.3 (2 x C*H), 37.1 (2 x C*HCH$_2$), 26.9 (2 x CH$_2$SH).

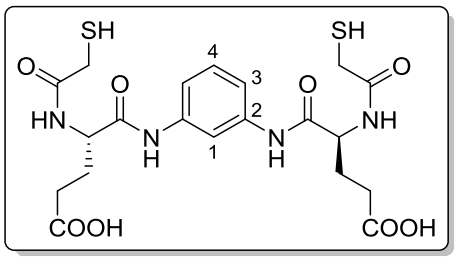


[**1g**]: this compound was obtained as described above for [**1a**], starting from [**4g**]. The product was purified by reversed-phase flash chromatography using a mixture of CH$_3$CN + 0.07% (v/v) TFA and H$_2$O + 0.1% (v/v) TFA as mobile phase (gradient: from 30% to 65% CH$_3$CN in H$_2$O) and 51.3 mg of [**1g**] (44% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{32}H_{32}N_6O_4S_2$ $[M+H]^+$ (m/z): 629.2004, found: 629.2003. ^{1}H NMR (400 MHz, DMSO-*d*$_6$): δ = 10.82 (d, *J* = 2.4 Hz, 2H, H^5), 10.17 (s, 2H, NHCOC*H), 8.37 (d, *J* = 7.9 Hz, 2H, C*HNHCO), 7.93 (s, 1H, H^1), 7.64 (d, *J* = 7.8 Hz, 2H, H^{10}), 7.35–7.27 (m, 4H, 2H x H^3 + 2H x H^7), 7.24–7.18 (m, 1H, H^4), 7.16 (d, *J* = 2.4 Hz, 2H, H^{13}), 7.05 (ddd, *J* = 8.1, 6.9, 1.2 Hz, 2H, H^8), 6.97 (ddd, *J* = 8.0, 7.0, 1.1 Hz, 2H, H^9), 4.72 (td, *J* = 8.0, 5.8 Hz,

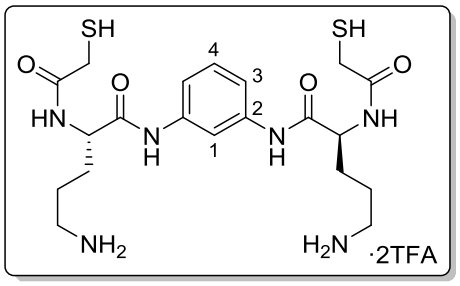
2H, C*H), 3.23–3.11 (m, 6H, 4H x C$\underline{H}_2$SH + 2H x C*HC$\underline{H}_2$), 3.02 (dd, *J* = 14.6, 8.2 Hz, 2H, C*HC$\underline{H}_2$), 2.65 (t, *J* = 8.0 Hz, 2H, SH). ^{13}C NMR (101 MHz, DMSO-d_6): δ = 170.3 (2 x $\underline{C}$OC*H), 169.4 (2 x $\underline{C}$OCH$_2$), 139.0 (2 x C^2), 136.0 (2 x C^6), 128.8 (1 x C^4), 127.3 (2 x C^{11}), 123.6 (2 x C^{13}), 120.9 (2 x C^8), 118.5 (2 x C^{10}), 118.2 (2 x C^9), 114.7 (2 x C^3), 111.3 (2 x C^7), 110.8 (1 x C^1), 109.6 (2 x C^{12}), 54.4 (2 x C*H), 28.0 (2 x C*H$\underline{C}$H$_2$), 27.0 (2 x CH$_2$SH).



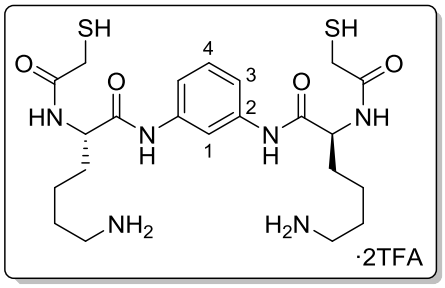
[**1h**]: this compound was obtained as described above for [**1a**], starting from [**4h**]. The product was purified by reversed-phase flash chromatography using a mixture of CH$_3$CN + 0.07% (v/v) TFA and H$_2$O + 0.1% (v/v) TFA as mobile phase (gradient: from 5% to 30% CH$_3$CN in H$_2$O) and 49.1 mg of [**1h**] (53% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{18}H_{22}N_4O_8S_2$ $[M+H]^+$ (m/z): 487.0957, found: 487.0956. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 7.85 (t, *J* = 2.1 Hz, 1H, H^1), 7.35–7.29 (m, 2H, H^3), 7.28–7.22 (m, 1H, H^4), 4.90–4.81 (m, 2H, C*H), 3.24 (s, 4H, C$\underline{H}_2$SH), 2.91 (dd, *J* = 16.6, 6.4 Hz, 2H, C*HC$\underline{H}_2$), 2.78 (dd, *J* = 16.6, 7.0 Hz, 2H, C*HC$\underline{H}_2$). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 173.6 (2 x COOH), 173.4 (2 x $\underline{C}$OCH$_2$), 170.9 (2 x $\underline{C}$OC*H), 139.8 (2 x C^2), 130.1 (1 x C^4), 117.6 (2 x C^3), 113.8 (1 x C^1), 52.3 (2 x C*H), 36.8 (2 x C*H$\underline{C}$H$_2$), 28.1 (2 x CH$_2$SH).



[**1i**]: this compound was obtained as described above for [**1a**], starting from [**4i**]. The product was purified by reversed-phase flash chromatography using a mixture of CH$_3$CN + 0.07% (v/v) TFA and H$_2$O + 0.1% (v/v) TFA as mobile phase (gradient: from 5% to 30% CH$_3$CN in H$_2$O) and 49.6 mg of [**1i**] (58% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{20}H_{26}N_4O_8S_2$ $[M+H]^+$ (m/z): 515.1270, found: 515.1271. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 7.90 (t, *J* = 2.0 Hz, 1H, H^1), 7.35–7.30 (m, 2H, H^3), 7.29–7.23 (m, 1H, H^4), 4.53 (dd, *J* = 8.7, 5.3 Hz, 2H, C*H), 3.24 (s, 4H, C$\underline{H}_2$SH), 2.45 (t, *J* = 7.6 Hz, 4H, C*HCH$_2$C$\underline{H}_2$), 2.25–2.12 (m, 2H, C*HC$\underline{H}_2$), 2.08–1.96 (m, 2H, C*HC$\underline{H}_2$). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 176.0 (2 x COOH), 173.1 (2 x $\underline{C}$OCH$_2$), 171.4 (2 x $\underline{C}$OC*H), 139.5 (2 x C^2), 129.8 (1 x C^4), 117.0 (2 x C^3), 113.1 (1 x C^1), 54.5 (2 x C*H), 30.7 (2 x C*HCH$_2$$\underline{C}H_2$), 28.3 (2 x C*H$\underline{C}H_2$), 27.7 (2 x CH$_2$SH).

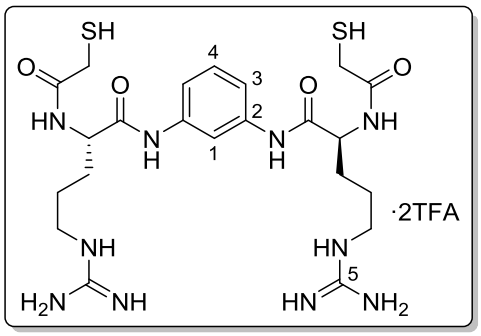


[**1j**]: to a solution of [**4j**] (133 mg, 0.117 mmol) in dry DCM (3.0 mL), $PhSiH_3$ (345 µL, 2.80 mmol) was added under inert atmosphere of Ar. Then a solution of $Pd(PPh_3)_4$ (18 mg, 15 µmol) in dry DCM (2.0 mL) was added. The mixture was stirred at room temperature for 1 hour, after which complete conversion of the starting material was observed by TLC. The crude mixture was filtered through a bed of Celite® and the filtrate was concentrated to dryness under reduced pressure. The resulting residue was redissolved in DCM (1.0 mL), and TFA (4.5 mL), TIS (242 µL, 0.933 mmol) and EDT (117 µL, 1.40 mmol) were added rapidly and under stirring. The reaction mixture was stirred at room temperature for 1 hour, after which the solvents were partially evaporated using a N_2 flow. Diethyl ether was added over the reaction mixture and the product was filtered and washed with diethyl ether. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 2% to 10% CH_3CN in H_2O) and 46.9 mg of [**1j**·2TFA] were obtained as a white solid (56% yield). HRMS (ESI+) calcd. for $C_{20}H_{32}N_6O_4S_2$ $[M+H]^+$ (m/z): 485.2005, found: 485.2007. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 8.02–7.97 (m, 1H, H^1), 7.34–7.18 (m, 3H, 2H x H^3 + 1H x H^4), 4.54 (dd, J = 8.0, 5.5 Hz, 2H, C*H), 3.25 (s, 4H, $C\underline{H}_2SH$), 3.07–2.90 (m, 4H, $C\underline{H}_2NH_3^+$), 2.05–1.89 (m, 2H, C*H $C\underline{H}_2$), 1.88–1.68 (m, 6H, 2H x $C^*HC\underline{H}_2$ + 4H, C*H $CH_2C\underline{H}_2$). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 173.5 (2 x $\underline{C}OCH_2$), 171.6 (2 x $\underline{C}OC^*H$), 139.9 (2 x C^2), 130.2 (1 x C^4), 117.5 (2 x C^3), 113.5 (1 x C^1), 54.8 (2 x C*H), 40.3 (2 x $CH_2NH_3^+$), 30.4 (2 x C*H $\underline{C}H_2$), 28.1 (2 x CH_2SH), 25.0 (2 x $C^*HCH_2\underline{C}H_2$).

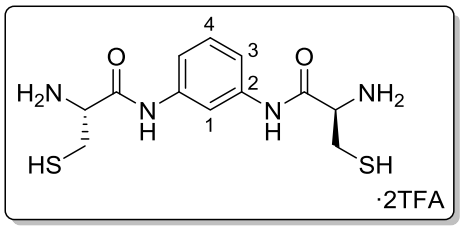


[**1k**]: a solution of [**4k**] (64.3 mg, 0.051 mmol) in dry DCM (3.3 mL) was cooled down to 0ºC in an ice-water bath. Then triisobutylsilane (TIS, 55 µL, 0.21 mmol) and 800 µL of a solution of HBr in CH_3COOH (33 wt. %) were added under stirring. After 40 minutes stirring at 0ºC, diethyl ether was added over the reaction mixture and the product was filtered off and washed with diethyl ether. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase

(gradient: from 2% to 12% CH_3CN in H_2O). During the purification the Br^- anions were exchanged by TFA^- and 31.8 mg of [**1k**·2TFA] (84% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{22}H_{36}N_6O_4S_2$ $[M+H]^+$ (m/z): 513.2318, found: 513.2319. ^1H NMR (400 MHz, MeOD-d_4): δ = 8.01–7.96 (m, 1H, H^1), 7.35–7.19 (m, 3H, 2H x H^3 + 1H x H^4), 4.49 (dd, J = 8.5, 5.5 Hz, 2H, C*H), 3.24 (s, 4H, $C\underline{H}_2SH$), 2.93 (t, J = 7.6 Hz, 4H, $C\underline{H}_2NH_3^+$), 2.01–1.87 (m, 2H, C*HC$\underline{H}_2$), 1.86–1.63 (m, 6H, 2H x C*HC$\underline{H}_2$ + 4H x $C\underline{H}_2CH_2NH_3^+$), 1.62–1.39 (m, 4H, $C\underline{H}_2CH_2C$*H). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 173.5 (2 x $\underline{C}OCH_2$), 172.1 (2 x $\underline{C}OC$*H), 139.9 (2 x C^2), 130.2 (1 x C^4), 117.5 (2 x C^3), 113.6 (1 x C^1), 55.3 (2 x C*H), 40.5 (2 x $CH_2NH_3^+$), 32.9 (2 x C*H$\underline{C}H_2$), 28.2 (2 x $\underline{C}H_2CH_2NH_3^+$), 28.1 (2 x CH_2SH), 23.8 (2 x C*HCH$_2\underline{C}H_2$).



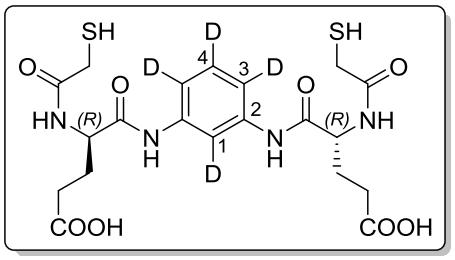
[**1l**]: this compound was obtained as described above for [**1a**], starting from [**6l**]. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 5% to 10% CH_3CN in H_2O) and 14.6 mg of [**1l**·2TFA] (27% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{22}H_{36}N_{10}O_4S_2$ $[M+H]^+$ (m/z): 569.2441, found: 569.2435. ^1H NMR (400 MHz, MeOD-d_4): δ = 7.99 (s, 1H, H^1), 7.33–7.20 (m, 3H, 2 x H^3 + 1 x H^4), 4.52 (dd, J = 8.3, 5.6 Hz, 2H, C*H), 3.29–3.11 (m, 8H, 4H x $C\underline{H}_2SH$ + 4H x $C\underline{H}_2NH$), 2.00–1.87 (m, 2H, C*HC$\underline{H}_2$), 1.86–1.57 (m, 6H, 2H x C*HC$\underline{H}_2$ + 4H x C*HCH$_2\underline{C}H_2$). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 172.1 (2 x $\underline{C}OCH_2$), 170.5 (2 x $\underline{C}OC$*H), 157.2 (2 x C^5), 138.4 (2 x C^2), 128.8 (1 x C^4), 116.1 (2 x C^3), 112.2 (1 x C^1), 53.6 (2 x C*H), 40.6 (2 x CH_2NH), 29.2 (2 x C*H$\underline{C}H_2$), 26.7 (2 x CH_2SH), 24.9 (2 x C*HCH$_2\underline{C}H_2$).



[**1m**]: this compound was obtained as described above for [**1a**], starting from [**2m**]. The product was purified by reversed-phase flash chromatography using a mixture of CH_3CN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 0% to 10% CH_3CN in H_2O) and 43.2 mg of [**1m**] (40% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{12}H_{18}N_4O_2S_2$ $[M+H]^+$ (m/z): 315.0949, found: 315.0950. ^1H NMR (400 MHz, MeOD-d_4): δ = 8.07 (t, J = 2.0

Hz, 1H, H^1), 7.41–7.36 (m, 2H, H^3), 7.36–7.30 (m, 1H, H^4), 4.14 (dd, *J* = 7.2, 5.1 Hz, 2H, C*H), 3.16 (dd, *J* = 14.7, 5.1 Hz, 2H, CH_2), 3.04 (dd, *J* = 14.7, 7.2 Hz, 2H, CH_2). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 166.7 (2 x CO), 139.6 (2 x C^2), 130.5 (1 x C^4), 117.5 (2 x C^3), 113.0 (1 x C^1), 56.9 (2 x C*H), 26.3 (2 x CH_2).

1.5.9. Synthesis of building block [(*RR*)-1i-d_4]



This compound was prepared exactly in the same way as described above for [**2i**, **3i**, **4i** and **1i**], starting from Fmoc-D-Glu(^{t}Bu)-OH and *m*-phenylenediamine-2,4,5,6-d_4. HRMS (ESI+) calcd. for $C_{20}H_{22}D_4N_4O_8S_2$ $[M+H]^+$ (m/z): 519.1516, found: 519.1501. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 4.53 (dd, *J* = 8.7, 5.3 Hz, 2H, C*H), 3.24 (s, 4H, C$\underline{H}_2$SH), 2.45 (t, *J* = 7.5 Hz, 4H, C*HCH$_2$C$\underline{H}_2$), 2.25–2.12 (m, 2H, C*HC$\underline{H}_2$), 2.09–1.95 (m, 2H, C*HC$\underline{H}_2$). ^{13}C NMR (101 MHz, MeOD-d_4):[t] δ = 176.4 (2 x COOH), 173.5 (2 x $\underline{C}$OCH$_2$), 171.8 (2 x $\underline{C}$OC*H), 139.8 (2 x C^2), 130.1 (1 x C^4), 117.3 (2 x C^3), 113.1 (1 x C^1), 54.9 (2 x C*H), 31.1 (2 x C*HCH$_2\underline{C}$H$_2$), 28.7 (2 x C*H$\underline{C}$H$_2$), 28.1 (2 x CH_2SH).

[t] Due to partial D/H exchange (section 1.3.2.6), triplet multiplicity for CD_{Ar} carbons (indicating coupling of carbon to deuterium) was not observed.

1.6. References

(1) Kuriyan, J.; Konforti, B.; Wemmer, D. *The molecules of life: Physical and chemical principles*; Garland Science, 2012.
(2) Whitford, D. *Proteins: structure and function*; John Wiley & Sons, 2013.
(3) Wu, Y.-D.; Gellman, S. *Acc. Chem. Res.* **2008**, *41*, 1231.
(4) Edwards, W.; Smith, D. K. *Chem. Commun.* **2012**, *48*, 2767.
(5) Rubio, J.; Alfonso, I.; Burguete, M. I.; Luis, S. V. *Soft Matter* **2011**, *7*, 10737.
(6) Alfonso, I.; Bru, M.; Burguete, M. I.; García-Verdugo, E.; Luis, S. V. *Chem. Eur. J.* **2010**, *16*, 1246.
(7) Suzuki, M.; Hanabusa, K. *Chem. Soc. Rev.* **2009**, *38*, 967.
(8) Nielsen, P. E.; Haaima, G. *Chem. Soc. Rev.* **1997**, *26*, 73.
(9) Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Norden, B. *Nature* **1994**, *368*, 561.
(10) Paradowska, J.; Pasternak, M.; Gut, B.; Gryzło, B.; Mlynarski, J. *J. Org. Chem.* **2012**, *77*, 173.
(11) Bauke Albada, H.; Rosati, F.; Coquière, D.; Roelfes, G.; Liskamp, R. M. J. *Eur. J. Org. Chem.* **2011**, *2011*, 1714.
(12) Rodríguez-Llansola, F.; Escuder, B.; Miravet, J. F. *J. Am. Chem. Soc.* **2009**, *131*, 11478.
(13) Micale, N.; Scarbaci, K.; Troiano, V.; Ettari, R.; Grasso, S.; Zappalà, M. *Med. Res. Rev.* **2014**, *34*, 1001.
(14) Zervoudi, E.; Saridakis, E.; Birtley, J. R.; Seregin, S. S.; Reeves, E.; Kokkala, P.; Aldhamen, Y. A.; Amalfitano, A.; Mavridis, I. M.; James, E. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 19890.
(15) Kokkonen, P.; Rahnasto-Rilla, M.; Kiviranta, P. H.; Huhtiniemi, T.; Laitinen, T.; Poso, A.; Jarho, E.; Lahtela-Kakkonen, M. *ACS Med. Chem. Lett.* **2012**, *3*, 969.
(16) Martí, I.; Bolte, M.; Burguete, M. I.; Vicent, C.; Alfonso, I.; Luis, S. V. *Chem. Eur. J.* **2014**, *20*, 7458.
(17) Moure, A.; Luis, S. V.; Alfonso, I. *Chem. Eur. J.* **2012**, *18*, 5496.
(18) Martí-Centelles, V.; Burguete, M. I.; Galindo, F.; Izquierdo, M. A.; Kumar, D. K.; White, A. J. P.; Luis, S. V.; Vilar, R. *J. Org. Chem.* **2012**, *77*, 490.
(19) Klemm, K.; Radić Stojković, M.; Horvat, G.; Tomišić, V.; Piantanida, I.; Schmuck, C. *Chem. Eur. J.* **2012**, *18*, 1352.
(20) Alfonso, I.; Burguete, I.; Luis, S. V.; Miravet, J. F.; Seliger, P.; Tomal, E. *Org. Biomol. Chem.* **2006**, *4*, 853.
(21) Luppi, G.; Garelli, A.; Prodi, L.; Broxterman, Q. B.; Kaptein, B.; Tomasini, C. *Org. Biomol. Chem.* **2005**, *3*, 1520.
(22) Furlan, R. L. E.; Ng, Y.-F.; Otto, S.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2001**, *123*, 8876.
(23) Faggi, E.; Moure, A.; Bolte, M.; Vicent, C.; Luis, S. V.; Alfonso, I. *J. Org. Chem.* **2014**, *79*, 4590.
(24) Alfonso, I.; Bolte, M.; Bru, M.; Burguete, M. I.; Luis, S. V.; Vicent, C. *Org. Biomol. Chem.* **2010**, *8*, 1329.
(25) Luis, S. V.; Alfonso, I. *Acc. Chem. Res.* **2014**, *47*, 112.
(26) Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Parris, K.; Williams, K.; Rebek, J. *J. Am. Chem. Soc.* **1989**, *111*, 1082.
(27) Becerril, J.; Bolte, M.; Burguete, M. I.; Galindo, F.; García-España, E.; Luis, S. V.; Miravet, J. F. *J. Am. Chem. Soc.* **2003**, *125*, 6677.

(28) Rubio, J.; Alfonso, I.; Burguete, M. I.; Luis, S. V. *Chem. Commun.* **2012**, *48*, 2210.
(29) Rubio, J.; Alfonso, I.; Bru, M.; Burguete, M. I.; Luis, S. V. *Tetrahedron Lett.* **2010**, *51*, 5861.
(30) Moran, M. C.; Pinazo, A.; Perez, L.; Clapes, P.; Angelet, M.; Garcia, M. T.; Vinardell, M. P.; Infante, M. R. *Green Chem.* **2004**, *6*, 233.
(31) Rubio, J.; Izquierdo, M. A.; Burguete, M. I.; Galindo, F.; Luis, S. V. *Nanoscale* **2011**, *3*, 3613.
(32) Kang, S. O.; Hossain, M. A.; Bowman-James, K. *Coord. Chem. Rev.* **2006**, *250*, 3038.
(33) Choi, K.; Hamilton, A. D. *Coord. Chem. Rev.* **2003**, *240*, 101.
(34) Conn, M. M.; Rebek, J. *Chem. Rev.* **1997**, *97*, 1647.
(35) Rodriguez-Vazquez, N.; Lionel Ozores, H.; Guerra, A.; Gonzalez-Freire, E.; Fuertes, A.; Panciera, M.; M Priegue, J.; Outeiral, J.; Montenegro, J.; García-Fandiño, R. *Curr. Top. Med. Chem.* **2014**, *14*, 2647.
(36) Walsh, C. T. *Science* **2004**, *303*, 1805.
(37) Loughlin, W. A.; Tyndall, J. D. A.; Glenn, M. P.; Fairlie, D. P. *Chem. Rev.* **2004**, *104*, 6085.
(38) Fernandez-Lopez, S.; Kim, H.-S.; Choi, E. C.; Delgado, M.; Granja, J. R.; Khasanov, A.; Kraehenbuehl, K.; Long, G.; Weinberger, D. A.; Wilcoxen, K. M.; Ghadiri, M. R. *Nature* **2001**, *412*, 452.
(39) Panciera, M.; Amorín, M.; Granja, J. R. *Chem. Eur. J.* **2014**, *20*, 10260.
(40) Brea, R. J.; Reiriz, C.; Granja, J. R. *Chem. Soc. Rev.* **2010**, *39*, 1448.
(41) Ashkenasy, N.; Horne, W. S.; Ghadiri, M. R. *Small* **2006**, *2*, 99.
(42) Bong, D. T.; Clark, T. D.; Granja, J. R.; Ghadiri, M. R. *Angew. Chem. Int. Ed.* **2001**, *40*, 988.
(43) White, C. J.; Yudin, A. K. *Nat. Chem.* **2011**, *3*, 509.
(44) Gibson, S. E.; Lecci, C. *Angew. Chem. Int. Ed.* **2006**, *45*, 1364.
(45) Diederich, F.; Stang, P. J.; Tykwinski, R. R. *Modern supramolecular chemistry: strategies for macrocycle synthesis*; John Wiley & Sons, 2008.
(46) Alfonso, I.; Bolte, M.; Bru, M.; Burguete, M. I.; Luis, S. V. *Chem. Eur. J.* **2008**, *14*, 8879.
(47) Bru, M.; Alfonso, I.; Burguete, M. I.; Luis, S. V. *Tetrahedron Lett.* **2005**, *46*, 7781.
(48) Bru, M.; Alfonso, I.; Bolte, M.; Burguete, M. I.; Luis, S. V. *Chem. Commun.* **2011**, *47*, 283.
(49) Alfonso, I.; Bolte, M.; Bru, M.; Burguete, M. I.; Luis, S. V.; Rubio, J. *J. Am. Chem. Soc.* **2008**, *130*, 6137.
(50) Bru, M.; Alfonso, I.; Burguete, M. I.; Luis, S. V. *Angew. Chem. Int. Ed.* **2006**, *45*, 6155.
(51) Jursic, B. S.; Zdravkovski, Z. *Synth. Commun.* **1993**, *23*, 2761.
(52) Valeur, E.; Bradley, M. *Chem. Soc. Rev.* **2009**, *38*, 606.
(53) Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61*, 10827.
(54) Han, S.-Y.; Kim, Y.-A. *Tetrahedron* **2004**, *60*, 2447.
(55) Williams, A.; Ibrahim, I. T. *Chem. Rev.* **1981**, *81*, 589.
(56) Anderson, G. W.; Callahan, F. M. *J. Am. Chem. Soc.* **1958**, *80*, 2902.
(57) König, W.; Geiger, R. *Chem. Ber.* **1970**, *103*, 788.
(58) Sheikh, M. C.; Takagi, S.; Yoshimura, T.; Morita, H. *Tetrahedron* **2010**, *66*, 7272.
(59) Rauschenberg, M.; Bandaru, S.; Waller, M. P.; Ravoo, B. J. *Chem. Eur. J.* **2014**, *20*, 2770.

(60) Kubik, S. *Chem. Soc. Rev.* **2009**, *38*, 585.
(61) Mahmoud, Khaled A.; Long, Y.-T.; Schatte, G.; Kraatz, H.-B. *Eur. J. Inorg. Chem.* **2005**, *2005*, 173.
(62) Baba, A. R.; Gowda, D. C. *Lett. Pept. Sci.* **2001**, *8*, 309.
(63) Fields, C.; Lloyd, D.; Macdonald, R.; Otteson, K.; Noble, R. *Peptide Res.* **1990**, *4*, 95.
(64) Blondelle, S. E.; Houghten, R. A. *Int. J. Pept. Protein Res.* **1993**, *41*, 522.
(65) Fields, G. B. In *Peptide Synthesis Protocols*; Springer: 1995, p 17.
(66) Kozikowski, A. P.; Chen, Y.; Gaysin, A.; Chen, B.; D'Annibale, M. A.; Suto, C. M.; Langley, B. C. *J. Med. Chem.* **2007**, *50*, 3054.
(67) Samaritoni, J. G.; Copes, A. T.; Crews, D. K.; Glos, C.; Thompson, A. L.; Wilson, C.; O'Donnell, M. J.; Scott, W. L. *J. Org. Chem.* **2014**, *79*, 3140.
(68) Pearson, D. A.; Blanchette, M.; Baker, M. L.; Guindon, C. A. *Tetrahedron Lett.* **1989**, *30*, 2739.
(69) Thieriet, N.; Alsina, J.; Giralt, E.; Guibé, F.; Albericio, F. *Tetrahedron Lett.* **1997**, *38*, 7275.
(70) Farkas, F.; Thurner, A.; Kovács, E.; Faigl, F.; Hegedűs, L. *Catal. Commun.* **2009**, *10*, 635.
(71) Friend, C. M.; Chen, D. A. *Polyhedron* **1997**, *16*, 3165.
(72) Ben-Ishai, D.; Berger, A. *J. Org. Chem.* **1952**, *17*, 1564.
(73) Feichtinger, K.; Zapf, C.; Sings, H. L.; Goodman, M. *J. Org. Chem.* **1998**, *63*, 3804.
(74) Wang, K.-L.; Liu, Y.-L.; Lee, J.-W.; Neoh, K.-G.; Kang, E.-T. *Macromolecules* **2010**, *43*, 7159.
(75) Matsubara, S.; Yokota, Y.; Oshima, K. *Chem. Lett.* **2004**, *33*, 294.
(76) Sharma, U.; Verma, P. K.; Kumar, N.; Kumar, V.; Bala, M.; Singh, B. *Chem. Eur. J.* **2011**, *17*, 5903.

CHAPTER 2

The effect of DMSO in the aqueous thiol-disulfide dynamic covalent chemistry of model pseudopeptides

The studies comprehended in this Chapter resulted in the following publication:

Atcher, J.; Alfonso, I. *RSC Adv.* **2013**, *3*, 25605.

2.1. Introduction

2.1.1. Disulfide bond

Disulfide bonds, also called disulfide bridges, are the covalent union between two substituted sulfur atoms (Figure 2.1). The union is relatively strong, with a bond dissociation energy of typically *ca.* 60 kcal·mol^{-1}.[1] The disulfide bond shows a conformational preference for dihedral C–S–S–C angles approaching 90°, although the energy barrier to rotation around the S–S bond is quite low (*ca.* 7.5 kcal·mol^{-1}).[2] The distance between the two sulfur atoms is about 2.05 Å.

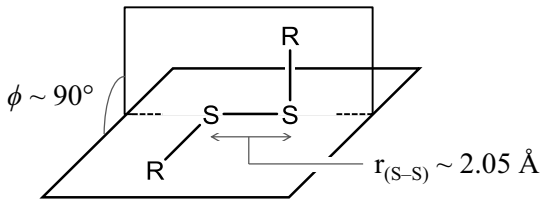


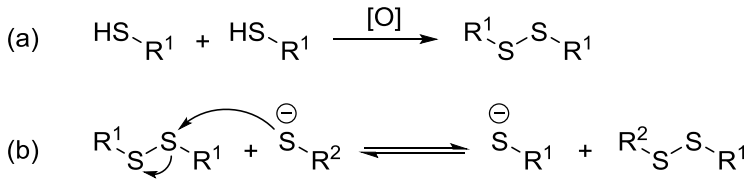
Figure 2.1. Representation of the dihedral angle (ϕ) and the bond length between the two sulfur atoms ($r_{(S–S)}$) of the disulfide bond.

Biologically important molecules containing the disulfide group are widely distributed in Nature. Disulfide bonds have an important role in many biological systems, most significantly in protein and peptide chemistry where they play a key part in the development of secondary structures.[3-4] Between thiol-containing cysteine residues, disulfide bridges frequently undergo "disulfide reshuffling" to attain a state where the prevailing linkages express the lowest energy folded structure.[3] Thus, disulfides in proteins are fundamental in the maintenance of biological activity and conformational stability. Besides, disulfide bonds are also known to be implicated in some other important biochemical processes, *e.g.* the triggering event in the cleavage of DNA by calichemicin and esperamicin,[5] etc.

2.1.2. Thiol oxidation and disulfide exchange processes

Disulfide bonds are usually generated from the irreversible oxidation of two thiol molecules (Scheme 2.1a). The reaction requires the presence of an oxidant reagent and can occur spontaneously without the presence of any external catalyst, or assisted *in vivo* by different enzymes.[3,6-7] Among the conventional methods for the formation of disulfide bonds in the laboratory, air oxidation in aqueous medium is the most commonly used.[8-9] The simple exposure of an aqueous solution of thiols to the air leads

to the slow oxidation of the thiols to generate the corresponding disulfides, while dissolved oxygen molecules are reduced to water. The reaction requires the presence of a catalytic amount of deprotonated thiolate, and thus neutral or slightly basic conditions are needed for the oxidation process to take place.[a] The disulfide bond formation is not only a very important naturally occurring process but has also significance in synthetic organic chemistry, especially in the field of peptides synthesis.[10-12]



Scheme 2.1. Thiol oxidation (a) and disulfide exchange (b) processes.

The second reaction depicted in Scheme 2.2 is the thiol-disulfide exchange process, also called disulfide metathesis or simply disulfide exchange. This reaction has a significant biological importance, as it plays an important role in the folding of proteins, among other biochemical processes.[13-14] The essence of the disulfide exchange mechanism[15] is the S_N2 nucleophilic attack of a thiolate anion on the sulfur atom of a disulfide moiety to cleave the original S–S bond and create another.[b] Although it involves the cleavage and formation of a strong covalent bond, the process is reversible and allows the disulfides to self-correct for reaching the most stable final situation. Similarly to the disulfide formation process (Scheme 2.1a), since the disulfide exchange also requires the presence of a catalytic amount of deprotonated thiolate to proceed, the reaction is highly sensitive to the pH of the medium and the pK_a of the two involved thiols (R^1–SH and R^2–SH in Scheme 2.1b).[1,16] In most cases, neutral or slightly basic conditions are needed to generate a sufficient thiolate concentration for the exchange process to occur.

As a consequence of the thiolate catalysis needed in both the disulfide formation and exchange processes, the two reactions depicted in Scheme 2.1 can be intentionally "frozen" by lowering the pH of the medium.[8]

[a] Typically, the pK_a of thiols is *ca.* 8-10, although it can be considerably lower for aromatic thiols.
[b] The determination of the reaction mechanism has been subject of investigation for a long time and it is well stablished now that the thiolate anion is the reacting species. This conclusion is based on two principal arguments: i) the empirical knowledge that a protonated thiol is a weak nucleophile relative to a thiolate anion, and ii) the experimental observation that the reaction is base-catalyzed when the thiol pK_a is greater than the pH of the solution.

2.1.3. Disulfide-based DCLs

After very scarce precursor studies in which equilibrium mixtures of disulfides had been used to study the lipid mixing in bilayers[17] and the template effects of a tripeptide on the composition of a very simple mixture of disulfides,[18] in 2000 Otto, Furlan and Sanders investigated and reported[19] for the first time the use of the disulfide linkage for the generation of reversible and diverse DCLs. At that time, the disulfide chemistry supposed a remarkable addition to the limited number of reactions that were available in the framework of DCC. The authors highlighted the advantages of the neutral or mildly basic conditions required for the disulfide chemistry,[c] and the fact that no external catalyst is required for the exchange process.

Disulfide-based DCLs are commonly generated simply by dissolution of thiol building blocks in water at pH 7-9 under air.[8] The stablishing of the dynamic process occurs through the interplay between the two previously commented processes: the thiol oxidation (Scheme 2.1a) and the disulfide exchange (Scheme 2.1b). The process is reversible as long as thiolate anions are present in solution, but the oxidation process is irreversible, and the exchange stops after the building blocks are fully oxidized. Thus, the whole process is finished when all thiols have been oxidized to disulfides, typically within several days,[19-20] allowing for easy analysis and purification of the generated library. As the final situation is not subjected to further equilibration processes, it is crucial for the generation of useful DCLs to work in such conditions in which the exchange reaction is faster than the oxidation of the thiols. Only if this condition is met, the concentration of each of the members of the resulting DCL will be dictated by its relative stability, and thus the library will have reached the thermodynamic equilibrium.

In contrast with the slow oxidation process of the extensively used air-oxidation in aqueous media (pH 7-9), that leads to equilibrated DCLs; in some cases the oxidation is intentionally accelerated in order to obtain kinetically trapped situations. With this purpose, stronger oxidants like I_2 are used to achieve a much faster oxidation of the thiols.[21] The almost immediate disulfides formation prevents the disulfide exchange to significantly take part in the process, thus reaching a non-equilibrated final situation, *i.e.* a kinetic trap.

[c] The mild conditions needed for the disulfide exchange were clearly in contrast with the previously developed imine chemistry that requires acidic media for reversibility.

Due to the controllability of the disulfide exchange reaction, the synthetic ease with which diverse thiol-bearing building blocks can be fashioned, and the soft reaction conditions compatible with biomolecules and aqueous media, disulfide exchange has become one of the most widely used reactions in DCC.[8] Additionally, the disulfide bonds have the remarkable advantage of being symmetrical linkages, as they are generated from the reaction of two units of the same organic functional group, *i.e.* two thiols. This allows the formation of highly diverse dynamic libraries from the mixture of a few thiol BBs, in contrast with the non-symmetrical reversible bonds (*e.g.* imines, esters, hydrazones, etc.), that require the specific union of two different functional groups to be formed. In the disulfide-based DCLs, *a priori* all possible combinations of BBs are allowed, with the intrinsic limitation imposed by the functionality degree of the BBs to be joined.[d,20,22]

Many chemists have explored the use of this reversible linkage to construct and study complex molecular architectures whose conventional synthesis would be either very tedious or impossible to achieve. In this regard, impressive examples have been reported in which the disulfide chemistry allows the preparation of macrocycles (Chapters 2-5),[21,23] catalysts,[24-26] capsular molecules,[22,27] catenanes,[28-31] molecular knots[32-33] and even self-replicating molecules.[34-37] Also remarkable is the fact that in disulfide systems, DCLs commonly display great sensitivity to templation effects, where addition of a new molecule alters their equilibrium position through supramolecular interactions with library constituents.

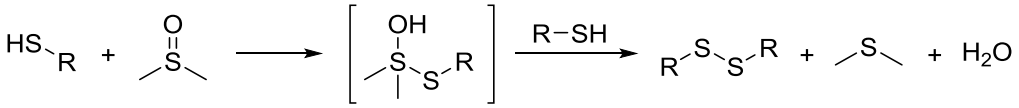
2.1.4. DMSO as an organic co-solvent

Dimethyl sulfoxide (DMSO) is a polar aprotic solvent that has extensively been used as an organic co-solvent in aqueous mixtures in order to increase the solubility of many organic molecules. It is one of the most powerful readily available organic solvents, being able to dissolve a great variety of substances, including both polar and nonpolar compounds. Water/DMSO mixtures have important industrial applications and, due to the relatively low toxicity of DMSO, aqueous mixtures with DMSO are even used in biological applications such as in the field of drug discovery.[38] Probably, the main

[d] The allowed combinations of BBs in a fully oxidized disulfide-based DCL are those in which the number of thiols to be joined is even, as two thiol groups are needed to form each of the disulfide bonds. This consideration is of special importance when preparing dynamic libraries by mixing BBs with different degrees of functionality (mono-,di-, trithiols, etc.).

drawback or limitation associated with the use of DMSO is its high boiling point (189 °C),[39] that makes very difficult to recover the compounds dissolved in this organic solvent.

In 1963 DMSO was reported[40] as a selective oxidizing reagent for the synthesis of disulfides from the corresponding thiols. Very interestingly, the formation of disulfides by DMSO was found to proceed in a wide range of pH, thus overcoming the limitation of the conventional oxidation by air that is applicable only at a narrow basic pH range.[9] Additionally, in contrast with the relatively long reaction times required for the disulfide formation by air oxidation, the use of DMSO allows faster oxidation rates, yet being a soft oxidizing process. In the proposed reaction pathway,[41] a thiol molecule combines with a molecule of DMSO in the rate determining step to form a sulfoxide-thiol adduct (Scheme 2.2). The adduct then reacts more rapidly with another molecule of thiol to form the observed products: the corresponding disulfide, dimethyl sulfide and water.



Scheme 2.2. Proposed reaction pathway for the oxidation of thiols to disulfides by DMSO.

The reaction mechanism can be slightly different depending on the pH, as different protonation states of the thiol group are involved. Very interestingly, in contrast with the previously commented oxidation by air, when DMSO is used as an oxidant the presence of deprotonated thiolate anions is not a required condition, reason why the reaction can also proceed at acidic pH.

On the other hand, DMSO has also been described to have an effect on the disulfide exchange reaction (Scheme 2.1b). As transition structures of the disulfide exchange mechanism have the charge density more widely distributed than reactants, hydrophobic environments are known to catalyze the process.[15] The rates for thiol-disulfide exchange in DMSO, and in general in polar aprotic solvents, are faster by a factor of approximately 10^3 than rates in polar protic solvents like water and methanol.[1,42-43]

At the time that the study corresponding to this Chapter was carried out, to the best of our knowledge there was only one example in the literature in which DMSO was used in small proportion (5.5% (v/v)) as a co-solvent in water for the preparation of

DCLs of disulfides.[44] However, in this study there is no explicit mention of the effect of this co-solvent on the kinetics or the thermodynamics of the system.

2.2. Objectives and hypothesis

The main objective of the present Chapter is to find suitable experimental conditions for the preparation of useful disulfide-based DCLs by the reversible union of the dithiol building blocks prepared in Chapter 1. With this purpose, we hypothesized that the incorporation of dimethyl sulfoxide (DMSO) as an organic co-solvent in aqueous mixtures would exert some beneficial effects on the solubility of the library members, as well as on the kinetics and thermodynamics of the disulfide formation and disulfide exchange processes.

This main objective can be divided into the following three specific aims:

i) To evaluate the use of DMSO as an organic co-solvent to increase the solubility of the library members in aqueous mixtures.

ii) To study the influence of three variables (the DMSO content, the pH of the mixtures and the presence of differently charged moieties around the thiol groups) on the rate of the disulfides formation.

iii) To investigate the effect of the same three variables on the ability of the generated disulfide-based DCLs to reach the thermodynamic equilibrium.

2.3. Results and discussion

2.3.1. Motivation for the use of DMSO as a co-solvent

For the preparation of disulfide-based dynamic libraries in water, the solubility of the organic building blocks in such a polar solvent is a recurring issue. In most of the cases the DCLs are designed to work under thermodynamic control and so it is tremendously important to ensure the total solubility of all involved species, including the initial BBs, the intermediates of the oxidation process and the final generated disulfides. In the case of precipitation, the composition of the whole library would be shifted towards the formation of the precipitate, obtaining a kinetic trap.[45]

In order to enhance the solubility of the BBs and avoid precipitation problems, two main strategies can be implemented. The first one is the most obvious and widely used strategy and consists in the increase of the polarity of the BBs by furnishing them with charged groups. These solubilizing groups are well solvated by water molecules and provide the BBs with good aqueous solubility. With this purpose, quite commonly the carboxylic acid group, carboxylate at the slightly basic conditions required for the thiol-disulfide reversible chemistry, is incorporated.[23,46-49] Alternatively, the second strategy consists in tuning the properties of the solvent instead of the properties of the solute. The incorporation of an organic co-solvent can decrease the polarity of the medium, thus preventing the hydrophobically driven aggregation and precipitation processes.[50-52] With this purpose many examples are reported in the literature in which an organic solvent, mainly acetonitrile,[25,53-57] is incorporated in the aqueous media for the better performance of the disulfide-based DCLs.

For the particular case of the BBs **1a-m** previously described in Chapter 1, as they are very different in terms of polarity, they are expected to have quite different solubilities in water. Thus, those BBs bearing charged groups at neutral pH, *i.e.* the negatively charged **1h-i** and the positively charged **1j-m**, are expected to be soluble in water. This was corroborated in section 2.3.3, where BBs **1h** and **1j** proved to be fully soluble at 2 mM concentration in water at pH 7.5 and 2.5 respectively. On the contrary, those BBs with no charges are more prone to be water insoluble. Preliminary solubility tests showed that, at mM concentrations, none of the neutral BBs **1a-g** is completely soluble in water.

To overcome this crucial solubility issue and with the aim to be able to generate chemically diverse DCLs by the mixture of differently charged BBs (Chapters 3, 4 and 5), the use of an organic co-solvent was envisioned. Different solubility experiments were performed using DMSO, methanol and acetonitrile in different proportions, and we finally opted for the use of DMSO, as it has some important advantages over the others. The reason for the election of this co-solvent was obviously not only the gain in solubility of the libraries, the initial motivation for the use of an organic co-solvent, but also the rest of the previously commented specific features attributed to this organic solvent (section 2.1.4). To summarize, the interesting features of DMSO that encouraged us to carry out the present study are these four: i) it is the most suitable co-solvent to increase the solubility of the libraries, ii) it is known to promote the rapid oxidation of the thiols, shortening the oxidation time required to prepare the dynamic libraries, iii) DMSO also promotes the fast disulfide exchange, a tremendously beneficial aspect when intending to prepare dynamic libraries under thermodynamic control, and iv) among the possible organic co-solvents, DMSO is the strongest solvating one and also the most efficient for breaking H-bonds, therefore being more comparable with the effect of water.[e]

2.3.2. The interplay between DMSO and pH

Before investigating the use of water/DMSO mixtures for the generation of disulfide-based DCLs, the interplay between the presence of DMSO and the acidity of the corresponding aqueous mixtures was considered. The addition of DMSO to an aqueous buffered solution is known to be able to alter the pH of the mixture, a key parameter in the chemistry of the disulfide bond. This pH variation strongly depends on the nature of the buffer.[58]

For the pH screenings carried out in the following sections 2.3.3 and 2.3.4, the aqueous phosphate-citrate buffer system, also known as McIlvine buffer,[59] was used to cover the pH from 2.5 to 7.5. Consequently, the effect of DMSO was tested on this particular buffer, for the whole pH range. Six McIlvine buffer solutions (pH 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5) were prepared as previously reported,[60] containing different amounts of sodium chloride as an inert salt in order to fix their ionic strength at 0.5 M.

[e] This characteristic of DMSO will be of special importance in Chapter 3, where bio-like conditions are desirable in order to be able to stablish certain parallelisms between the adaptation of a small disulfide-based DCL and the natural evolution of the halophilic proteins.

From these, two aliquots of each buffer were prepared, one with 10% and the other with 25% (v/v) DMSO. The mixtures were left to cool down to room temperature and the pH was measured with a regular glass electrode.

Table 2.1. Measured pH values of the buffer solutions with different %DMSO.

Entry	Buffer	Buffer + 10% DMSO	Buffer + 25% DMSO
1	2.50	2.75	3.23
2	3.49	3.73	4.19
3	4.53	4.77	5.22
4	5.53	5.80	6.31
5	6.52	6.85	7.46
6	7.66	7.98	8.61

As evidenced in Table 2.1., after the addition of DMSO the measured pH systematically increases. This increase is dependent on the pH of the initial aqueous buffer and the proportion of DMSO. Hence, the presence of 10% DMSO causes an increase of 0.2-0.4 pH units, while the presence of 25% DMSO causes an increase of 0.7-1.0 pH units. The DMSO-induced basification appears to be slightly smaller for the acidic buffers.

According to literature,[58] the "real pH" of an aqueous solution containing a certain amount of DMSO should be lower than the value directly measured with a glass electrode system. This difference between the measured pH and the "real pH" is the result of a change in the standard potential of the electrode and not the activity of H^+. However, for the DMSO proportions of 10% and 25% used herein, this difference is known to be very small. For simplicity, from now on the given pH values in the discussion are those of the corresponding aqueous buffer solutions with no DMSO.

2.3.3. Evaluation of the oxidation rate

The first factor to be studied regarding the use of DMSO as a co-solvent for the preparation of disulfide-based DCLs was the kinetics of the oxidation process (Scheme 2.1a). As previously discussed, the oxidation is required to be slow enough to allow the mixtures to fully equilibrate, but at the same time a moderately fast oxidation would be beneficial as it would allow to shorten the long oxidation time normally needed for the preparation of the libraries.

A preliminary oxidation reaction was performed for the Ser-based BB (**1d**) in order to first get an idea of the identity of the generated disulfides as well as the time required to complete their formation. The oxidation of 2 mM **1d** was performed at the commonly used slightly basic conditions (aqueous phosphate-citrate buffer at pH 7.5), incorporating 25% (v/v) DMSO. The reaction was followed by HPLC-UV and UPLC-MS, until complete conversion of the starting dithiol was observed.

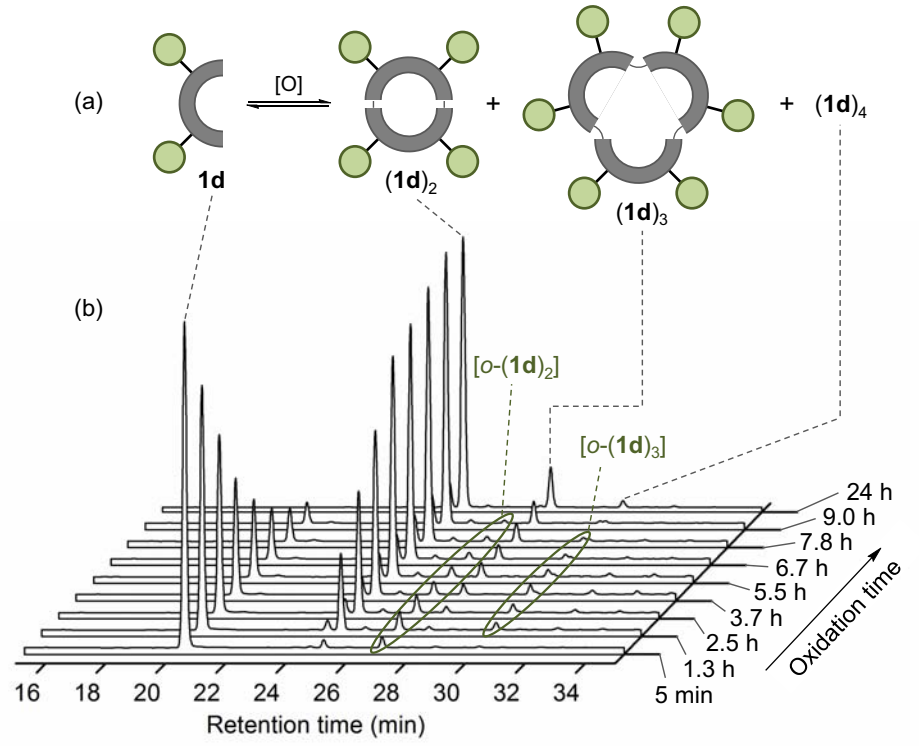


Figure 2.2. Schematic representation (a) and HPLC-UV traces (254 nm) at different reaction times (b) of the oxidation process for 2 mM **1d** in aqueous phosphate-citrate buffer (pH 7.5) with 25% (v/v) DMSO.

In Figure 2.2b, the HPLC traces corresponding to the first 9 hours of the oxidation process are shown, as well as the chromatogram of the analysis performed 1 day after the start of the reaction. As the oxidation process proceeds, BB **1d** (peak at 20.3 min in Figure 2.2b) is quite rapidly[f] consumed and three macrocyclic disulfides are generated as depicted in Figure 2.2a: the dimer $(\mathbf{1d})_2$ as the predominant species in the library (peak at 25.0 min in Figure 2.2b; ~85% of the overall HPLC area), the trimer $(\mathbf{1d})_3$ as a minor compound (peak at 28.0 min in Figure 2.2b; ~13% of the overall HPLC area), and a very small amount of the tetramer $(\mathbf{1d})_4$ (peak at 30.4 min in Figure 2.2b; ~2% of the overall HPLC area). The formation of these three products passes through

[f] In 9 hours ~95% of the initial dithiol **1d** is already consumed.

intermediate stages in which the open forms of the dimer and the trimer, *i.e.* the partially oxidized open dimer *o*-(**1d**)$_2$ and open trimer *o*-(**1d**)$_3$, are present in detectable amounts.

Interestingly, over the whole oxidation process the sum of the HPLC integrals of the reagent, the intermediates and the produced disulfides is constant (<2% variation between the HPLC analyses performed at the oxidation time of 5 min and 24 h). Therefore, at 254 nm the molar absorptivity of BB **1d** is the same whether as free dithiol or as part of any of the generated macrocyclic disulfides.

In order to expand the scope of the kinetic evaluation to different BBs and different experimental conditions, the well stablished Ellman's test[61] was preferred. In contrast with the HPLC monitoring, this analytical method doesn't provide us with information about the composition of the mixture, but only about the overall concentration of thiols present in the sample. However, as a clear advantage, when performed on a UV/Vis multi-well plate reader, the Ellman's test allows the simultaneous analysis of various oxidation processes, therefore being an ideal method for the broad screening of experimental conditions.

With this aim, three representative BBs were chosen bearing neutral (**1d**), acidic (**1h**) and basic (**1j**) polar residues in order to map the effect of different microenvironments around the thiol group.[62-63] The Ellman's test was used to evaluate the rate of disulfide formation, *i.e.* thiols oxidation, at pH values from 2.5 to 7.5 and different proportions of DMSO (10%, 25% and 100% (v/v)). Additionally, only for BBs **1h** and **1j** two lower DMSO proportions (0% and 2% (v/v)) were also tested at very specific pH values (pH 7.5 for **1h** and 2.5 for **1j**). At these experimental conditions, the corresponding charged carboxylate and ammonium groups of **1h** and **1j** respectively, provide good solubility in polar solvents and allow the use of these two BBs even in the absence of the organic co-solvent. For each of the assayed conditions, the initial rate (v_0) and the half-life time ($t_{1/2}$) values for the oxidation reaction were obtained from the corresponding kinetic data (Table 2.2).

Table 2.2. Initial rates (v_0, mM·h^{-1}) and half-life times ($t_{1/2}$, h) for the oxidation of **1d**,**h**,**j** at different pH values in water/DMSO mixtures.

entry	[**BB**]	pH	%DMSO[a]	v_0	$t_{1/2}$
1	**1d**	—	100	0.065	38.3
2	**1d**	7.5	10	0.81	2.8
3	**1d**	2.5-6.5[b]	10	1.12	1.95
4	**1d**	7.5	25	0.98	2.6
5	**1d**	2.5-6.5[b]	25	1.67	1.2
6	**1h**	7.5	0	0.094	25.0
7	**1h**	—	100	0.097	23.4
8	**1h**	7.5	2	0.36	6.3
9	**1h**	7.5	10	0.87	2.7
10	**1h**	2.5-6.5[b]	10	1.03	2.1
11	**1h**	7.5	25	1.31	1.6
12	**1h**	2.5-6.5[b]	25	1.90	0.96
13	**1j**	2.5	0	<0.001	—
14	**1j**	—	100	0.18	11.9
15	**1j**	2.5	2	0.31	7.8
16	**1j**	7.5	10	0.70	2.7
17	**1j**	2.5-6.5[b]	10	0.74	2.5
18	**1j**	7.5	25	1.04	2.2
19	**1j**	2.5-6.5[b]	25	1.65	1.2

[a] Volume % of DMSO in water. [b] The reactions were separately carried out at pH 2.5, 3.5, 4.5, 5.5 and 6.5 with very similar results.

Interestingly, the reactions performed in water/DMSO mixtures proceeded faster than in either pure DMSO or pure water (*e.g.* entries 6-12 in Table 2.2). Moreover, for the 0%-25% DMSO range, the higher content of DMSO also induces a slightly faster oxidation (entries 6, 8, 9 and 11 for **1h** and entries 13, 15, 17 and 19 for **1j**). For a suitable comparison, Figure 2.3 shows two representations of the kinetic data obtained for **1h**: the remaining free thiol (%) at different oxidation times (Figure 2.3a); and the initial rate and half-life time values as a function of the DMSO content (Figure 2.3b). The oxidation rate gradually increases with the amount of DMSO, being already remarkable even at low proportion of DMSO (2%) where the oxidation showed to be *ca.* four-fold faster than in pure water (entries 6 and 8 in Table 2.2). The addition of just 10% DMSO allows the practically full oxidation of the thiols within 24 h (<5% of remaining thiol), which means up to 10 times faster than in pure water.

Depending on the content of DMSO, the pH of the reaction mixture has different effects on the oxidation rate. As commonly known, in 100% water media (entries 6 and 13 in Table 2.2) the disulfides formation takes place slowly at pH 7.5, whereas at pH 2.5 the thiols are stable to the oxidation. In the presence of 2%-25% DMSO, the pH value between 2.5 and 6.5 practically does not affect the reaction rate within the experimental

errors. Remarkably, the presence of just 2% DMSO already enables the oxidation process even at markedly acidic pH (Figure 2.3 and entry 15 in Table 2.2). At pH 7.5 the overall thiol oxidation is slightly slower compared to pH 2.5-6.5 (*e.g.* entries 2 *vs.* 3 and 4 *vs.* 5 in Table 2.2). We hypothesized that at neutral pH, a considerable amount of deprotonated thiolate anion should increase the rate of the disulfide exchange process (Scheme 2.1b), which efficiently competes with the thiol oxidation.

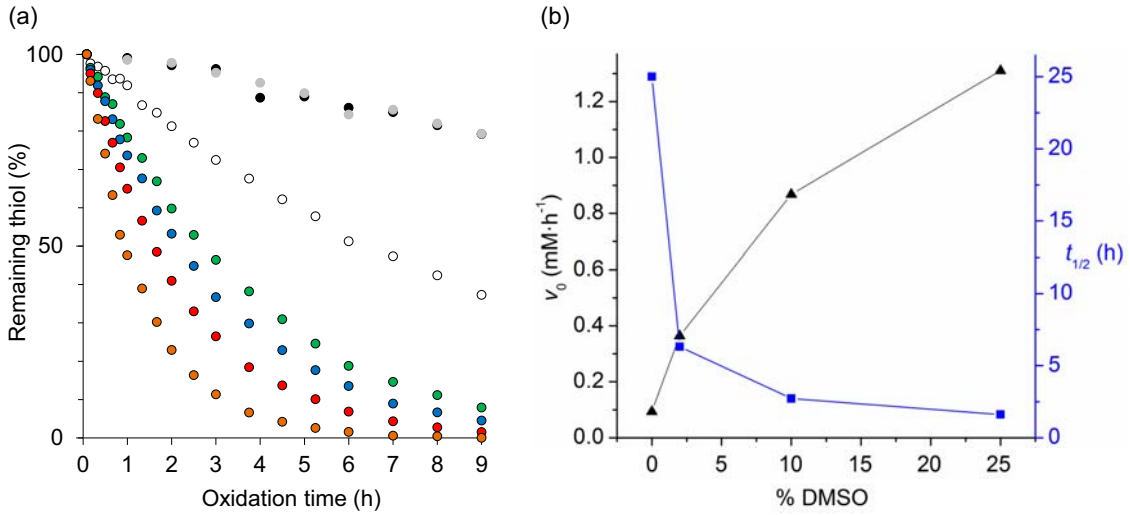


Figure 2.3. (a) Plot of the remaining thiol (%) over the oxidation time (h) for 2 mM **1h** at different oxidation conditions: 0% DMSO at pH 7.5 (●); pure DMSO (●); 2% (v/v) DMSO at pH 7.5 (○); 10% (v/v) DMSO at pH 7.5 (●); 10% (v/v) DMSO at pH 2.5 (●); 25% (v/v) DMSO at pH 7.5 (●); and 25% (v/v) DMSO at pH 2.5 (●). (b) Representation of the initial rate (v_0, mM·h^{-1}, left axis, black triangles) and the half-life time ($t_{1/2}$, h, right axis, blue squares) as a function of the DMSO content, for the oxidation of 2 mM **1h** at pH 7.5.

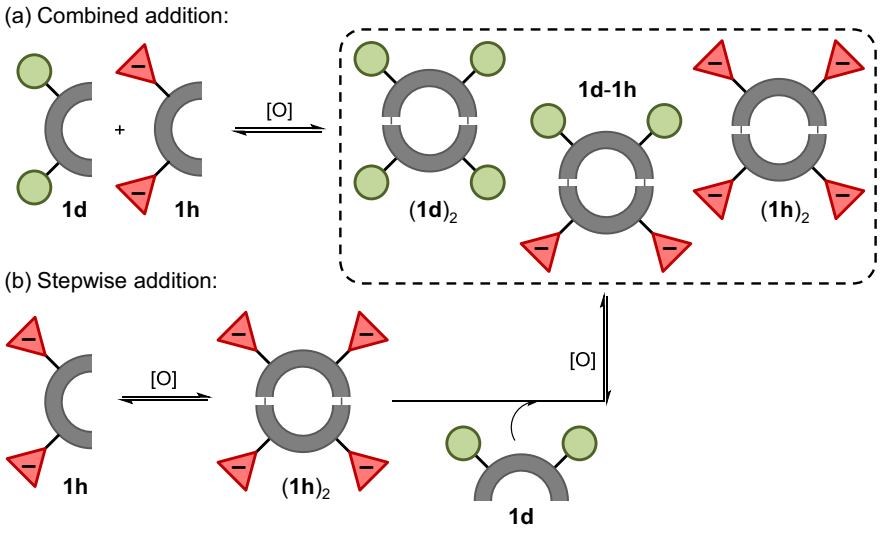
Regarding the nature of the BB, the presence of either negatively (**1h**) or positively charged groups (**1j**) has little effect on the rate of oxidation (*e.g.* entries 4, 11 and 18 or 5, 12 and 19). Thus, the different microenvironments around the thiol group do not seem to be able to cause any significant change on the kinetics of the disulfides formation.

To summarize, the presence of small amounts of DMSO as an organic co-solvent notably increases the thiols oxidation rate. This increase depends on the DMSO content and is already remarkable even at very small proportions of the organic co-solvent.

2.3.4. Evaluation of the reversibility

Since the presence of DMSO leads to the faster oxidation of the thiols, the kinetically controlled products could be obtained.[64] This situation is undesirable for the generation of most of the DCLs, usually designed to work under thermodynamic

control. Consequently, a simple experiment was designed in order to evaluate the reversibility of the libraries.



Scheme 2.3. Cartoon representation of the reversibility test. For simplicity, only the corresponding dimers are depicted.

As represented in Scheme 2.3a, two BBs (**1d** and **1h** in the example) were oxidized together allowing the formation of the two corresponding homodimers (**1d**)$_2$ and (**1h**)$_2$, the heterodimer **1d-1h**, and small amounts of most of the possible trimers, as observed by UPLC-MS. In parallel, BB **1b** was added over a pre-oxidized library generated from **1h** alone and consisting mainly of the corresponding homodimer (**1h**)$_2$ (Scheme 2.3b). If the exchange process is fast enough compared to the oxidation reaction, the composition of the library in the combined and stepwise additions must be the same, being the identical final situation an evidence of the reversibility of the libraries. On the contrary, if the exchange is too slow compared to the oxidation, the mixtures don't reach the thermodynamic equilibrium and the composition of the library in the combined and stepwise additions must be different.

Taking into account the prevalence of dimers in the mixtures, the parameter chosen to quantitatively characterize the composition of the libraries was the ratio between homo- and heterodimers. With this purpose, the thermodynamic constant corresponding to the exchange reaction between homo- and heterodimeric species was used. For a mixture generated from two BBs (**A** and **B**), the exchange constant $K_{[\mathbf{A},\mathbf{B}]}$ between the two homodimers (**AA** and **BB**) and their corresponding heterodimer (**AB**) is defined as shown in Equation 2.1. This constant presents the main advantage of being

dimensionless: its value does not depend on the actual concentration of any of the two BBs and this provides experimental robustness.

$$2\mathbf{AB} \rightleftharpoons \mathbf{AA} + \mathbf{BB} \qquad K_{[\mathbf{A},\mathbf{B}]} = \frac{[\mathbf{AA}]\cdot[\mathbf{BB}]}{[\mathbf{AB}]^2} \qquad \textbf{(Equation 2.1)}$$

$$\mathrm{RD}\ (\%) = \left|\frac{\left(K^{(\mathrm{b})}_{[\mathbf{A},\mathbf{B}]} - K^{(\mathrm{a})}_{[\mathbf{A},\mathbf{B}]}\right)}{K^{(\mathrm{a})}_{[\mathbf{A},\mathbf{B}]}}\right| \cdot 100 \qquad \textbf{(Equation 2.2)}$$

The reversibility test was performed at the same pH values tested in the kinetic evaluation (from 2.5 to 7.5) in both 10% and 25% (v/v) DMSO. Moreover, in addition of the pair of BBs **1d**+**1h** represented in Scheme 2.3, the equivalent reversibility test was also performed for **1d**+**1j** in order to study the effect of the nature of the side chains on the reversibility of the mixtures. For each of the tested experimental conditions the exchange constant was calculated both for the the combined addition $\left(K^{(\mathrm{a})}_{[\mathbf{A},\mathbf{B}]}\right)$ and the stepwise addition $\left(K^{(\mathrm{b})}_{[\mathbf{A},\mathbf{B}]}\right)$, and the relative difference (RD) between these two constants was calculated as shown in Equation 2.2.

Very interestingly, as shown in Table 2.3, for the two tested pairs of BBs and the two different proportions of DMSO, at pH ≥ 4.5 the composition of the mixtures prepared by combined and stepwise addition is the same within the experimental error, showing a <2% relative difference (RD) between the exchange constants calculated for the two addition protocols (entries 3-6 and 9-12 in Table 2.3). Thus, very remarkably, the libraries proved to reach the thermodynamic equilibrium at pH values down to 4.5. However, for the reactions performed at lower pH values (entries 1-2 and 7-8 in Table 2.3) a higher proportion of the homodimers was obtained in the stepwise addition compared to the combined addition ($K^{(\mathrm{b})}_{[\mathbf{A},\mathbf{B}]} > K^{(\mathrm{a})}_{[\mathbf{A},\mathbf{B}]}$). The calculated RD between the exchange constants obtained for the two addition protocols is relatively small at pH 3.5 (entries 2 and 8 in Table 2.3), while at pH 2.5 the difference is huge (entries 1 and 2 in Table 2.3). Therefore, at such acidic conditions, the mixtures are not able to reach the thermodynamic equilibrium.

Table 2.3. Values of the exchange constant between homo- and heterodimers of the libraries prepared by combined addition ($K^{(a)}_{[A,B]}$) and stepwise addition ($K^{(b)}_{[A,B]}$), as well as the values of their relative difference (RD), for the mixtures of **1d**+**1h** and **1d**+**1j**.

			1d + **1h**			**1d** + **1j**		
entry	pH	%DMSO	$K^{(a)}_{[1d,1h]}$	$K^{(b)}_{[1d,1h]}$	RD (%)	$K^{(a)}_{[1d,1j]}$	$K^{(b)}_{[1d,1j]}$	RD (%)
1	2.5	10	0.234	1.09	**>100**	0.240	2.87	**>100**
2	3.5	10	0.206	0.213	**3.4**	0.234	0.256	**9.4**
3	4.5	10	0.145	0.143	**<2**	0.248	0.248	**<2**
4	5.5	10	0.110	0.108	**<2**	0.247	0.246	**<2**
5	6.5	10	0.110	0.110	**<2**	0.250	0.248	**<2**
6	7.5	10	0.109	0.110	**<2**	0.253	0.256	**<2**
7	2.5	25	0.275	2.58	**>100**	0.234	2.81	**>100**
8	3.5	25	0.211	0.288	**36**	0.235	0.245	**4.3**
9	4.5	25	0.141	0.143	**<2**	0.236	0.237	**<2**
10	5.5	25	0.101	0.102	**<2**	0.257	0.256	**<2**
11	6.5	25	0.100	0.100	**<2**	0.265	0.264	**<2**
12	7.5	25	0.103	0.103	**<2**	0.274	0.269	**<2**

In summary, the analysis of the homo/heterodimers proportions by means of the exchange constant $K_{[A,B]}$ showed that the mixtures reach the thermodynamic equilibrium at pH ≥ 4.5, and this ability of the mixtures appears not to depend on the proportion of DMSO, at least within the range 10-25% DMSO, and not either the nature of the BBs. The generation of disulfide-based DCLs at acidic pH is very interesting, as it opens the way for the use of this chemistry for the study of biological systems that function at pH < 7, for instance, lysosomal enzymes.[65]

2.3.5. Study of the exchange process

The kinetic studies showed that, in the presence of DMSO, the pH value between 2.5 and 6.5 appears not to significantly alter the oxidation rate of the thiols. However, the acidity of the mixtures showed to be a key parameter regarding the ability of the libraries to reach the thermodynamic equilibrium. These two observations imply that the disulfides exchange process is affected by the pH of the medium, as the ability of the mixtures to reach the thermodynamic equilibrium only depends on whether the disulfide formation process (constant at pH 2.5-6.5, Scheme 2.1a) is faster or slower than the disulfide exchange process (Scheme 2.1b). In order to corroborate that, the exchange process was studied in a deeper detail by the evaluation of intermediate oxidation stages in the stepwise addition protocol performed for the reversibility tests.

The exchange process was studied for a selected pair of BBs (**1d** and **1h**) at 25% (v/v) DMSO and different pH values (Figure 2.4a-b for pH 4.5 and 2.4d-f for pH 2.5).

First of all, BB **1b** (in green) was pre-oxidized until complete formation of the corresponding disulfides 2.4a,d. To this pre-oxidized library, 1 equivalent of the second dithiol (**1h**, in red) was added. The resulting mixture was analyzed 40 minutes (Figure 2.4b,e) and > 24 hours (Figure 2.4c,f) after the addition of the second BB.

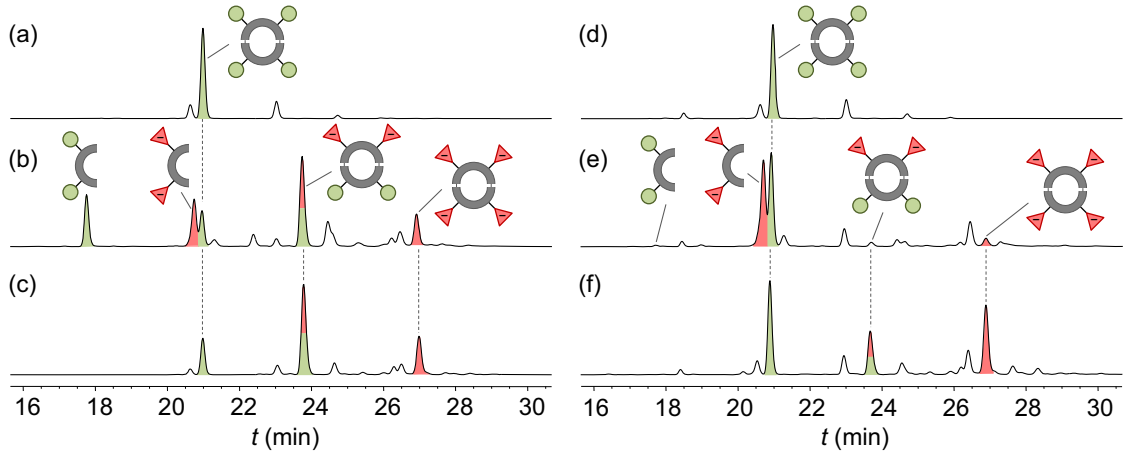


Figure 2.4. HPLC-UV traces (254 nm) of: (a,d) 2 mM **1d** fully oxidized in aqueous buffer with 25% (v/v) DMSO (>24 h); (b,e) the same library 40 min after the addition of 1 equivalent of **1h**; and (c,f) the same library >24 h after the addition of 1 equivalent of **1h**. Chromatograms (a-c) correspond to the libraries performed at pH 4.5, while chromatograms (d-f) correspond to the libraries performed at pH 2.5.

At pH 4.5, the addition of the second BB (**1h**) rapidly produces the growing of the peaks assigned to the free dithiol **1d** and the corresponding macrocyclic heterodimer **1d-1h**, demonstrating the fast disulfide exchange. Actually, the higher concentration of the heterodimer **1d-1h** compared to the homodimer $(\mathbf{1h})_2$ after only 40 minutes (Figure 2.4b) unambiguously proves that, at pH 4.5, the exchange process is much faster than the oxidation, reason why the libraries are able to reach the thermodynamic equilibrium. On the contrary, when the equivalent experiment was performed at pH 2.5, the behavior was notably different. In this case the mixture analyzed 40 minutes after the addition of the second BB (Figure 2.4e) contains almost undetectable amounts of the direct products of exchange (**1d** and **1d-1h**), demonstrating the slow disulfide exchange. At this acidic pH, since **1h** is not able to efficiently exchange with $(\mathbf{1d})_2$, the secondly added BB is oxidized with itself forming mainly the homodimer $(\mathbf{1h})_2$ and leading to a fully oxidized situation (Figure 2.4f) in which the two homodimers have a much larger presence in the mixture than the corresponding heterodimer. In this case, the low exchange rate is the responsible for not reaching the thermodynamic equilibrium.

2.3.6. Larger DCL in the limit experimental conditions

With the aim to illustrate the applicability of the found experimental conditions with DMSO as a co-solvent, a more diverse DCL of macrocyclic disulfides was set up. An equimolar mixture of **1d**+**1h**+**1j** was prepared in the limit experimental conditions for a thermodynamically controlled system: aqueous buffer at pH 4.5 with 25% (v/v) DMSO. The HPLC-UV and UPLC-MS analyses of the mixture obtained by oxidation of the three BBs together showed the presence of all the possible macrocyclic homo- and heterodimers, and most of the corresponding macrocyclic trimers (Figure 2.5a). For some of the peaks containing the Asp-based BB **1h**, an unexpected small amount of the dehydration product was also found (peaks marked with * in Figure 2.5a). For peptides and proteins it is well established[66-68] that the β-carbonyl group of Asp and Asn amino acids can acylate the amido NH group of the next residue producing a five-membered succinimide ring.[g] At neutral and basic pH the concentration of the succinimide derivative produced by this side reaction is always very low, whereas at acidic pH it can be an abundant product.[69] Moreover, in this particular case the acidity of the aromatic NH of **1h** can further favor the formation of the dehydration products.[70-71]

In a separate experiment, the DCL with only two of the BBs (**1d**+**1j**) was performed and, when the equilibrium was established (Figure 2.5b), the third BB (**1h**) was added and the mixture was allowed to re-equilibrate (Figure 2.5c). The composition of the HPLC traces of the libraries obtained either by combined or stepwise addition of the BBs (Figure 2.5a and 2.5c respectively) showed an identical composition regardless their preparation pathway, thus confirming that the dynamic chemical system reached the thermodynamic equilibrium composition.

[g] This reaction is one of the most common chemical modifications resulting in covalent damage to peptides and proteins.

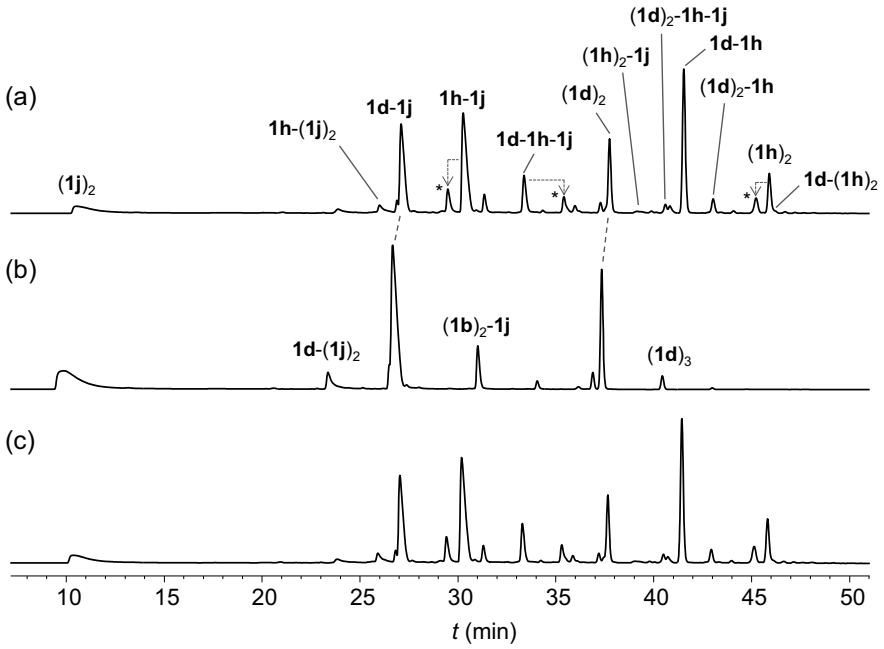


Figure 2.5. HPLC-UV traces (254 nm) of the equilibrated mixtures (2 mM of each BB in aqueous buffer (pH 4.5) with 25% (v/v) DMSO) of: (a) **1d**+**1h**+**1j**, (b) **1d**+**1j**, and (c) reaction (b) + **1h** and re-equilibration. Peaks marked with an asterisk correspond to dehydration side-products.

2.4. Conclusions

The use of DMSO as a co-solvent for the preparation of disulfide-based DCLs has proved to have interesting and beneficial effects on the thiol-disulfide dynamic covalent chemistry. These effects are summarized in the following four points:

i) The incorporation of relatively small amounts of DMSO increases the solubility of the BBs, allowing the use of water insoluble non-charged BBs. The neutral dithiol **1d** and the corresponding products of oxidation have proved to remain fully soluble at different pH values when just 10% (v/v) DMSO is used in aqueous mixtures.

ii) The presence of DMSO promotes the thiol oxidation reaction, highly reducing the reaction time for the disulfides formation. The addition of just 10% (v/v) DMSO causes a tenfold increase in the oxidation rate, allowing the practically full oxidation of the thiols within 24 h. The pH value between 2.5 and 6.5 practically does not affect the reaction rate, and the presence of just 2% (v/v) DMSO already enables the oxidation process even at markedly acidic pH. The presence of either negatively or positively charged groups in the microenvironment around the thiol group has little effect on the thiols oxidation rate.

iii) The incorporation of DMSO also accelerates the disulfide exchange reaction, allowing the systems to fully equilibrate even at slightly acidic pH values. In the presence of 10-25% (v/v) DMSO, regardless the charge of the BBs, the libraries have proved to reach the thermodynamic equilibrium at pH $\geq$ 4.5.

iv) The limit experimental conditions for a thermodynamically controlled system, *i.e.* 25% (v/v) DMSO in water at pH 4.5, have showed to be suitable for the generation of large and diverse DCLs of macrocyclic disulfides from the mixture of **1d**, **1h** and **1j**. The composition of these libraries is mainly dominated by the presence of the corresponding dimers together with small amount of trimers.

2.5. Experimental section

2.5.1 General methods

Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka and Merck) and were used without further purification. pH measurements were performed at room temperature on a Crison GLP21 pH-meter with the electrode Crison 50 14T. Absorbance measurements were performed on a Molecular Devices SpectraMax M5 microplate reader, at room temperature, and the monitoring wavelength was set at 412 nm. The 96 well microplates, PS, F-bottom, 655101 were used to place the samples and the microplate reader was set to shake the samples for 5 seconds before each measurement.

2.5.2. HPLC and MS analyses

PR-HPLC analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using a reversed-phase kromaphase C_{18} (25 x 0.46 cm, 5μm) column and (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures were used as mobile phase. The eluent used for those libraries not containing the BB **1j** was: 2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 40% CH_3CN over 48 min (constant flow set at 1 $mL \cdot min^{-1}$); and the eluent used for those libraries containing the BB **1j** was: 10 min at 2% CH_3CN in H_2O, then linear gradient from 2% to 40% CH_3CN over 62 min (constant flow set at 1 $mL \cdot min^{-1}$). Exceptionally, only for the HPLC analyses of the disulfides exchange evaluation (Figure 2.4 in section 2.3.5) the eluent used was: 2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 33% CH_3CN over 28 min (constant flow set at 1 $mL \cdot min^{-1}$). The monitoring wavelength was set at 254 nm and the temperature of the column was set at 25 °C. The HPLC samples were prepared by dilution of the DCLs with an acidic solution of 89% H_2O, 10% CH_3CN and 1% TFA.

HRMS analyses were carried out at the IQAC Mass Spectrometry Facility, using a UPLC-ESI-TOF equipment: [Acquity UPLC® BEH C_{18} 1.7 mm, 2.1x100 mm, LCT Premier Xe, Waters]. (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures at 0.3 mL/min were used as mobile phase. The eluent used for those libraries not containing the BB **1j** was: 2.5 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 50% CH_3CN over 27.5 min; and the eluent used for those libraries containing the

BB **1j** was: 2.5 min at 2% CH_3CN in H_2O, then linear gradient from 2% to 40% CH_3CN over 27.5 min. The temperature of the column was set at 25 °C. The HPLC samples were prepared by dilution of the DCLs with an acidic solution of 89% H_2O, 10% CH_3CN and 1% TFA. The HRMS analyses of the DCLs are shown in the electronic Annex.

2.5.3. Preparation of buffered water/DMSO mixtures

Six McIlvaine buffer solutions (pH 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5) were prepared by dissolving different amounts of Na_2HPO_4 and citric acid in milli-Q water, and incorporating different amounts of sodium chloride as an inert salt in order to fix the ionic strength of the six solutions at 0.5 M.[60] For the buffers prepared at pH 6.5 and 7.5 no citrate was used to avoid precipitation when mixing the buffer solutions with DMSO. Instead, these two buffers were prepared dissolving different amounts of Na_2HPO_4 and NaH_2PO_4. The %DMSO of the water/DMSO mixtures was directly calculated from the volumes of buffer and DMSO mixed, ignoring the small decrease in total volume that results from such mixing.[72]

2.5.4. Kinetic evaluation of the oxidation process

2.5.4.1. General procedure for the preparation of the oxidation samples

For the preparation of the oxidation samples containing 10% and 25% DMSO, *individual stocks* of the BBs **1d,h,j** (Conc. 1) were prepared in pure DMSO. From these, the oxidation samples (2 mM) were prepared by adding Vol. 1 of the individual stock to Vol. 2 of a buffer solution (pH 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5 separately). This addition was considered to be the starting time of the oxidation process.

Table 2.4. Concentrations (mM) and volumes (μL) of the general procedure for the preparation of the oxidation samples.

%DMSO	Conc. 1	Vol. 1	Vol. 2
10	20	30	270
25	8	75	225

For the preparation of the oxidation samples containing 0% DMSO, 20 μL of a 20 mM *individual stock* of **1h,j** prepared in milli-Q water were added to 180 μL of the corresponding buffer solution. For the oxidation samples containing 2% DMSO, 20 μL of a 20 mM *individual stock* prepared in milli-Q water were added to 176 μL of the corresponding buffer solution containing 4 μL of DMSO. For the oxidation samples

containing 100% DMSO, 50 µL of an 8 mM *individual stock* prepared in pure DMSO were added to 150 µL of DMSO. All the oxidation reactions were carried out at room temperature, in capped vials and without any stirring.

2.5.4.2. General procedure for the Ellman's test

A 0.80 mM solution of 5,5'-Dithiobis(2-nitrobenzoic acid) (DTNB or Ellman's reagent) was prepared by dissolving 8.0 mg in 25 mL of a 100 mM phosphate buffer (pH 8.0). Then, 200 µL of this freshly prepared solution were placed in a well of a 96-well plate and 10 µL of the oxidation sample were also added in the well and mixed. After 2 min of incubation at room temperature, the absorbance at 412 nm was measured. Simultaneously, for all the batches and reaction times, the absorbance of a blank was also measured. The blanks were prepared by adding 10 µL of a water/DMSO mixture (same %DMSO as the corresponding sample) to 200 µL of the DTNB solution. The net absorbance was calculated by subtracting the absorbance of the corresponding blank.

Finally, the free thiols concentration was obtained by means of a calibration curve (Figure 2.6). To obtain the calibration curve, a 6.21 mM stock solution of Cys was prepared by dissolving 37.59 mg in 50 mL of milli-Q water. From this, a set of stock solutions (5.17, 4.14, 3.10, 2.07 and 1.03 mM) was prepared by dilution with more milli-Q water. The net absorbance of each of the freshly prepared stocks was represented in front of the concentration and a least square regression line was fitted. The data showed a very good linear behaviour within the range of working concentrations ($R^2 = 1.00$).

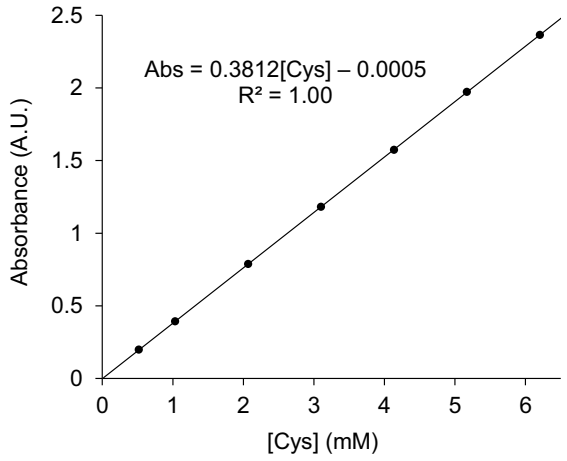


Figure 2.6. Linear least square calibration curve of the Ellman's test.

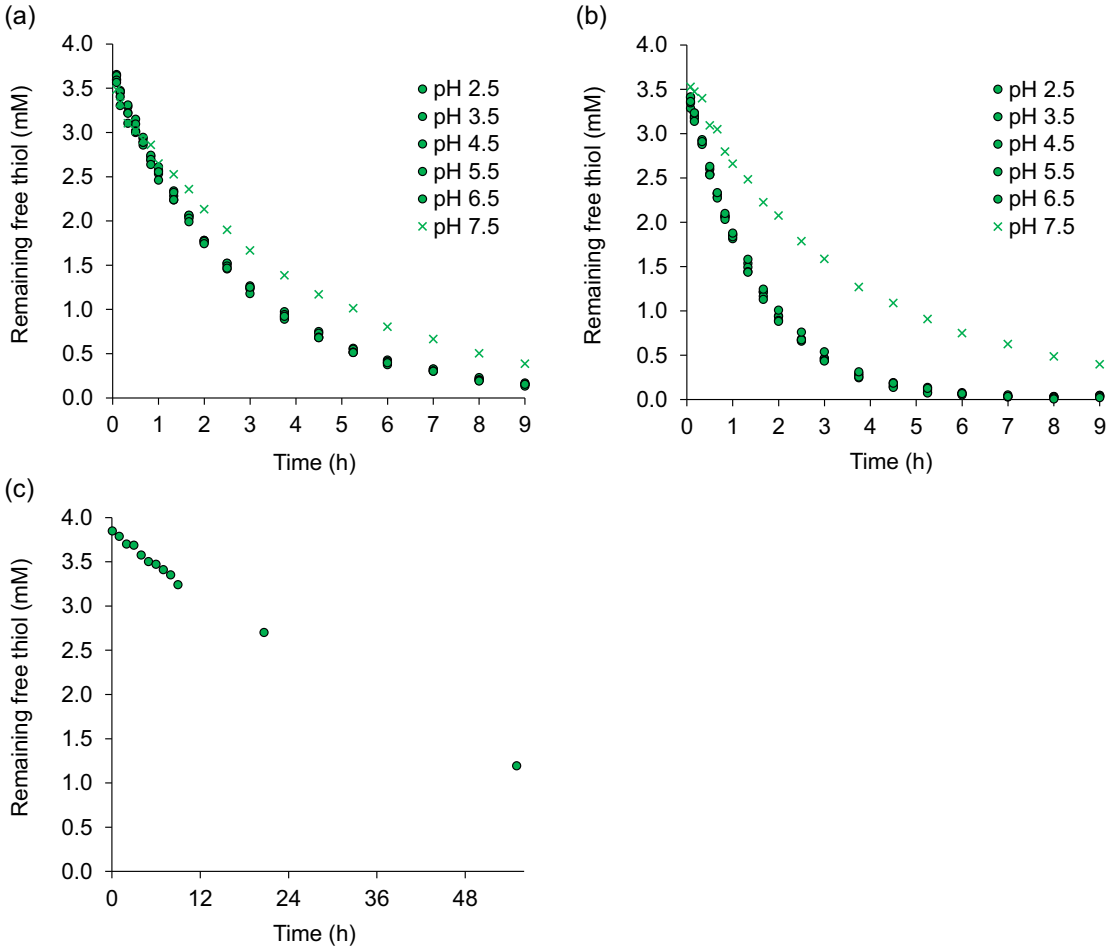


Figure 2.7. Representation of the remaining free thiol (mM) in front of the oxidation time (h) for 2 mM **1d** in: (a) aqueous buffer at pH 2.5-7.5 with 10% (v/v) DMSO, (b) aqueous buffer at pH 2.5-7.5 with 25% (v/v) DMSO, and (c) pure DMSO.

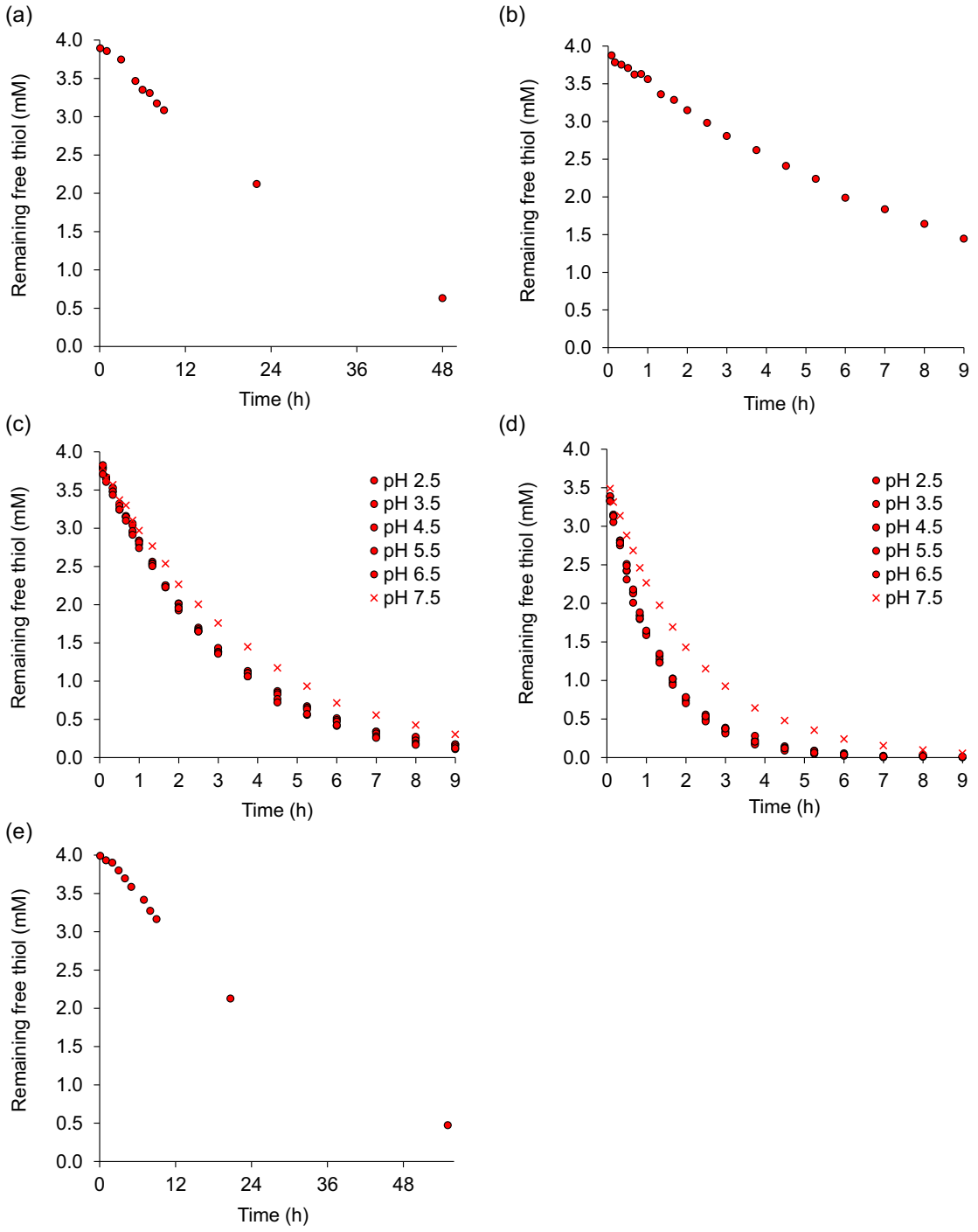


Figure 2.8. Representation of the remaining free thiol (mM) in front of the oxidation time (h) for 2 mM **1h** in: (a) aqueous buffer at pH 7.5, (b) aqueous buffer at pH 7.5 with 2% (v/v) DSMO, (c) aqueous buffer at pH 2.5-7.5 with 10% (v/v) DMSO, (d) aqueous buffer at pH 2.5-7.5 with 25% (v/v) DMSO, and (e) pure DMSO.

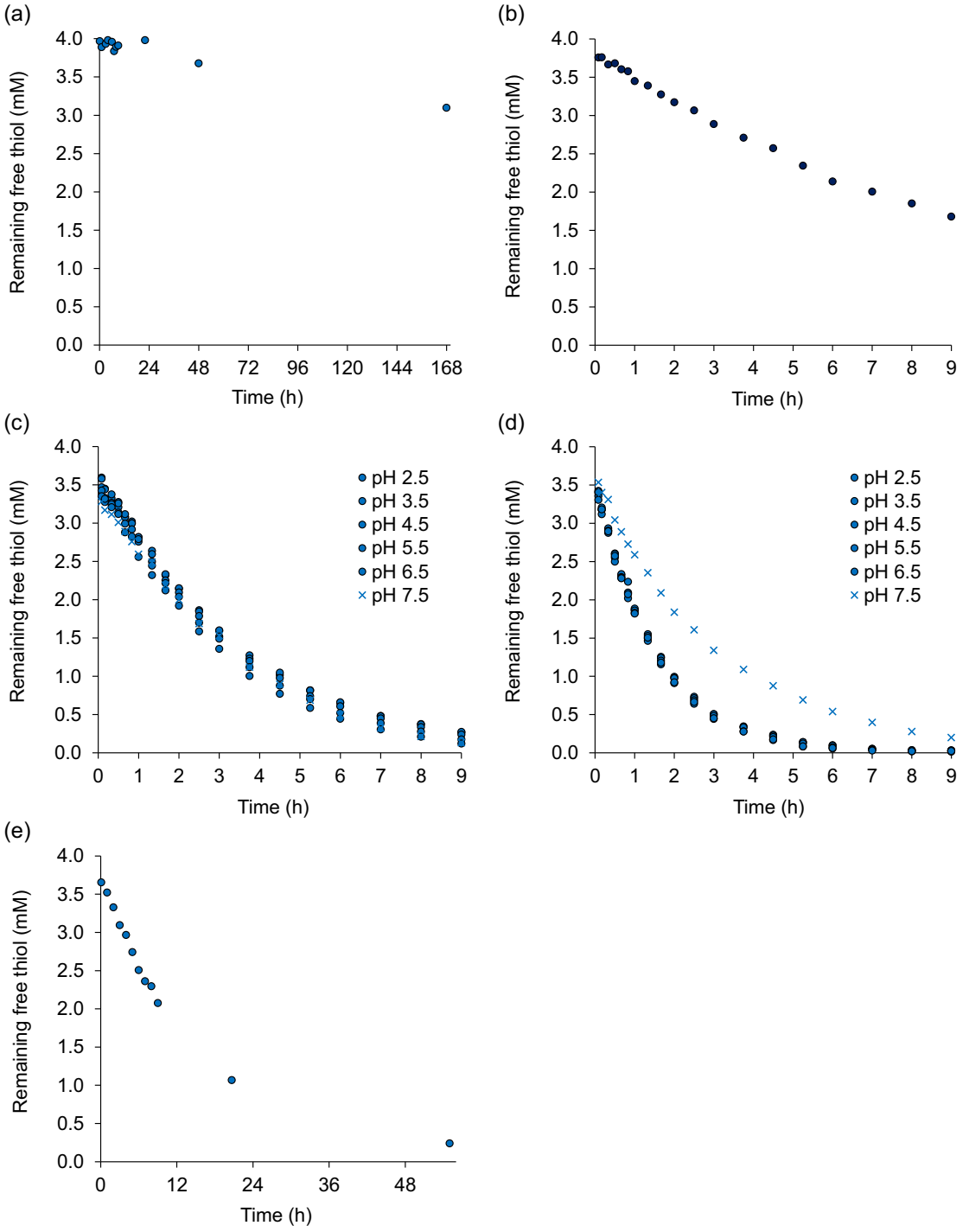


Figure 2.9. Representation of the remaining free thiol (mM) in front of the oxidation time (h) for 2 mM **1j** in: (a) aqueous buffer at pH 7.5, (b) aqueous buffer at pH 7.5 with 2% (v/v) DSMO, (c) aqueous buffer at pH 2.5-7.5 with 10% (v/v) DMSO, (d) aqueous buffer at pH 2.5-7.5 with 25% (v/v) DMSO, and (e) pure DMSO.

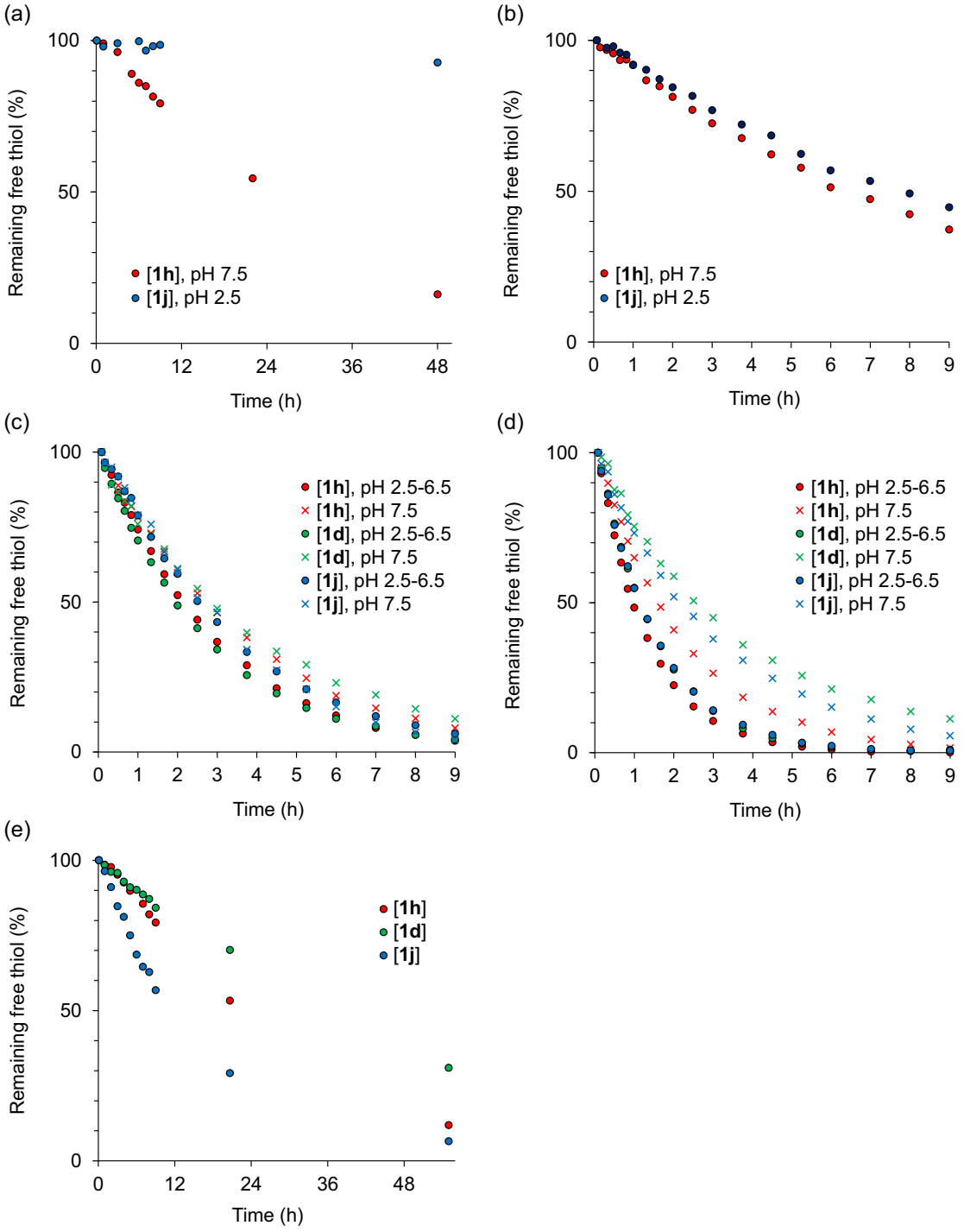


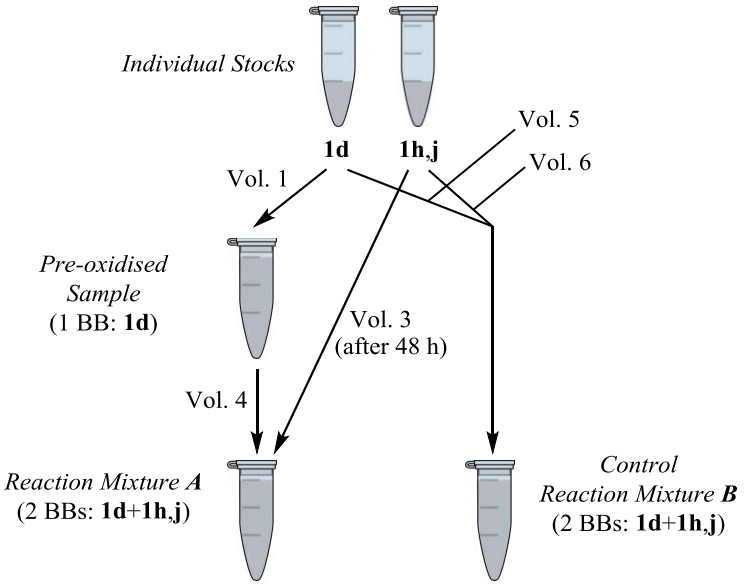
Figure 2.10. Comparison of the oxidation processes for **1d**,**h**,**j**. Representation of the remaining free thiol (%) in front of the oxidation time (h) for 2 mM **1d**,**h**,**j** in: (a) aqueous buffer at pH 7.5, (b) aqueous buffer at pH 7.5 with 2% (v/v) DSMO, (c) aqueous buffer at pH 2.5-7.5 with 10% (v/v) DMSO, (d) aqueous buffer at pH 2.5-7.5 with 25% (v/v) DMSO, and (e) pure DMSO.

2.5.4.3. Processing of the kinetic data

For each oxidation time, the real remaining free thiol concentration was calculated by means of the calibration curve. The points corresponding to the first hour of oxidation (the 2 first hours for the samples with 2% DMSO and the 6 first hours for the samples with 0% and 100% DMSO) were used to adjust a regression line (linear least square method). The slope of the line was taken as minus the initial rate ($-v_0$). For the calculation of the half-life time ($t_{1/2}$), the equation of the straight line containing the two closest points to the 50% concentration, *i.e.* the one just above and the one just below, was used. The time value that fulfilled the equation for 50% of remaining free thiol was taken as the half-life time. All the experiments were performed at least twice, observing no significant differences within the experimental error.

2.5.5. Reversibility tests

Individual stocks of the two BBs **1h**,**j** (Conc. 1) and **1d** (Conc. 2), were prepared in pure DMSO (Scheme 2.4). From these, a *pre-oxidised sample* was prepared by adding the *individual stock* of **1d** (Vol. 1) to a buffer solution (Vol. 2; pH 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5 separately). The *individual stock* of **1h**,**j** was stored at –80 ºC. After 48 hours, a 2 mM of each BB mixture, the *reaction mixture **A***, was prepared by adding the *individual stock* of **1h**,**j** (Vol. 3) to the *pre-oxidised sample* (Vol. 4). Only for the samples with a final 10% DMSO, 130 μL of the corresponding buffer solution (pH 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5 separately) were also added to the *mixture **A***.



Scheme 2.4. Preparation of the mixtures of the binary reversibility tests.

Simultaneously, the *control reaction mixtures* ***B*** (2 mM of each of the two BBs), was prepared by mixing the two *individual stocks*, Vol. 5 of **1d** and Vol. 6 of **1h,j**, with Vol. 7 of a buffer solution (pH 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5 separately). After 48 hours, the *reaction mixture* ***A***, the *control reaction mixture* ***B*** and the *pre-oxidised sample* were analysed by HPLC-UV (selected examples are shown in Figures 2.11-2.14).

Table 2.5. Concentrations (mM) and volumes (μL) of the general procedure for the binary reversibility tests.

%DMSO	Conc. 1	Conc. 2	Vol. 1	Vol. 2	Vol. 3	Vol. 4	Vol. 5	Vol. 6	Vol. 7
10	40	40	15	75	10	60	10	10	180
25	40	10	48	180	10	190	40	10	150

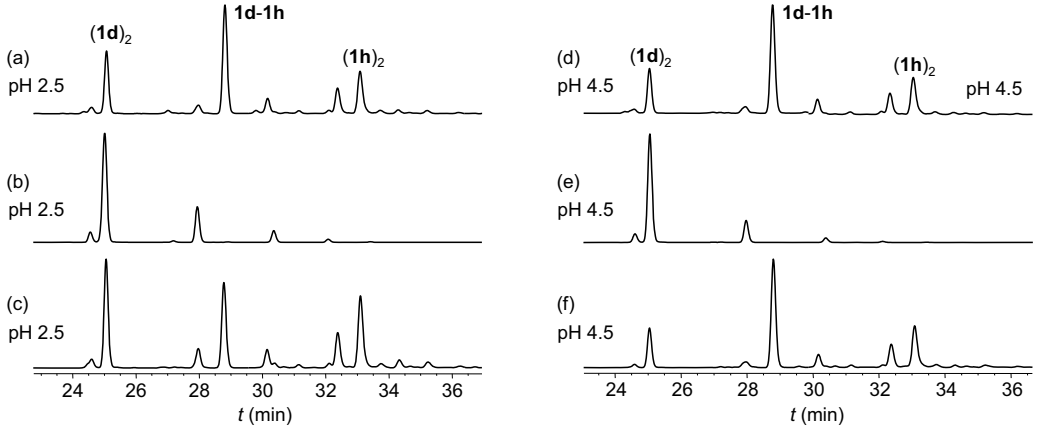


Figure 2.11. HPLC-UV traces (254 nm) of the *control reaction mixture* ***B*** (a,d), the *pre-oxidised sample* (b,e) and the *reaction mixture* ***A*** (c,f) of **1d+1h** in aqueous buffer with **10%** (v/v) DMSO, at pH 2.5 (a-d) and pH 4.5 (d-f).

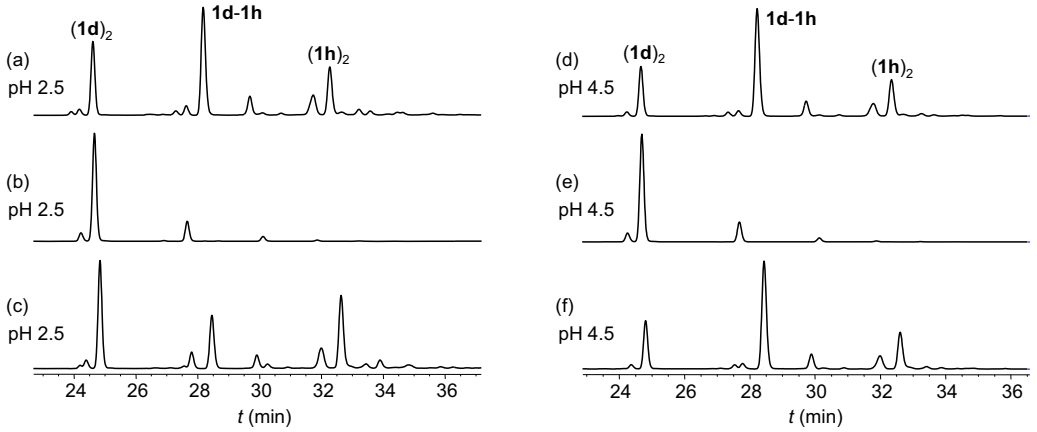


Figure 2.12. HPLC-UV traces (254 nm) of the *control reaction mixture* ***B*** (a,d), the *pre-oxidised sample* (b,e) and the *reaction mixture* ***A*** (c,f) of **1d+1h** in aqueous buffer with **25%** (v/v) DMSO, at pH 2.5 (a-d) and pH 4.5 (d-f).

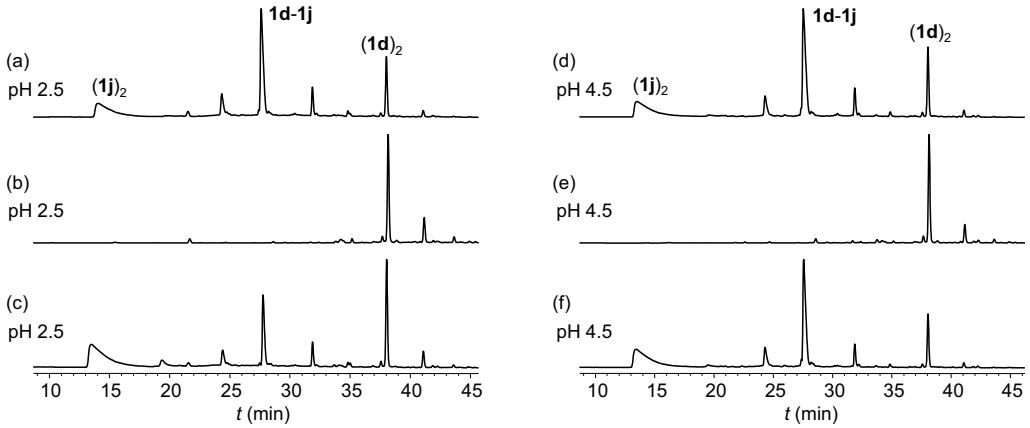


Figure 2.13. HPLC-UV traces (254 nm) of the *control reaction mixture* ***B*** (a,d), the *pre-oxidised sample* (b,e) and the *reaction mixture* ***A*** (c,f) of **1d+1j** in aqueous buffer with **10%** (v/v) DMSO, at pH 2.5 (a-d) and pH 4.5 (d-f).

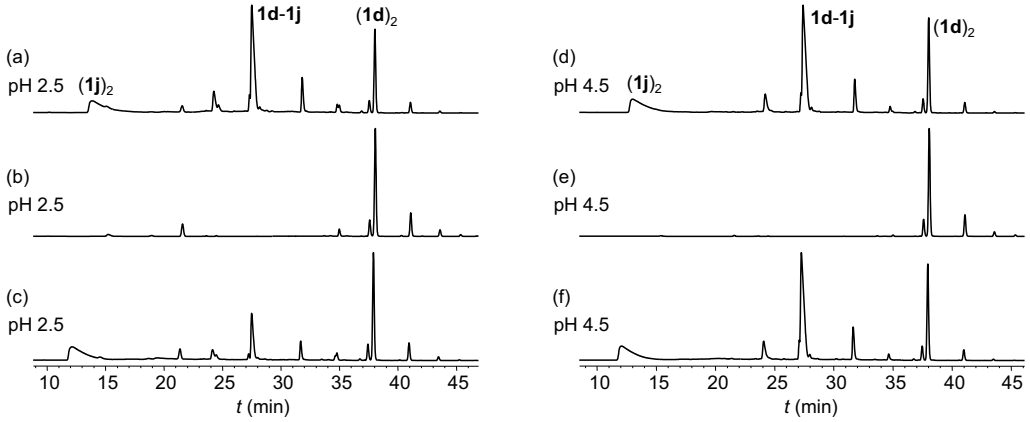


Figure 2.14. HPLC-UV traces (254 nm) of the *control reaction mixture* ***B*** (a,d), the *pre-oxidised sample* (b,e) and the *reaction mixture* ***A*** (c,f) of **1d+1j** in aqueous buffer with **25%** (v/v) DMSO, at pH 2.5 (a-d) and pH 4.5 (d-f).

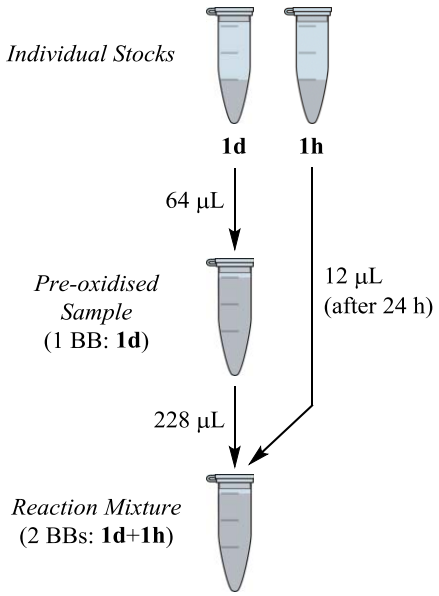
For the calculation of the exchange constant $K_{[\mathbf{A},\mathbf{B}]}$, the same extinction coefficient at 254 nm was assumed for the three BBs **1b**, **1h** and **1j**, as they are equipped with the same chromophore. Thus, the exchange constants were directly calculated by means of the corresponding HPLC areas as shown in Equation 2.3.

$$K_{[\mathbf{A},\mathbf{B}]} = \frac{[\mathbf{AA}]\cdot[\mathbf{BB}]}{[\mathbf{AB}]^2} \approx \frac{\text{Area}_{(\mathbf{AA})}\cdot\text{Area}_{(\mathbf{BB})}}{\text{Area}^2_{(\mathbf{AB})}} \qquad \textbf{(Equation 2.3)}$$

2.5.6. Study of the exchange process

A 40 mM *individual stock* of **1d** was prepared in DMSO (Scheme 2.5). From this, a *pre-oxidised sample* was prepared by adding 64 µL of the *individual stock* to 240 µL of a buffer solution (pH 2.5 and 4.5 separately).

After 24 hours the *pre-oxidised sample* was analysed by HPLC in order to confirm the total oxidation (Figure 2.4a,d in section 2.3.5). Then a 10 mM *individual stock* of **1h** was also prepared in DMSO. From this, a 2 mM of each BB mixture containing 25% DMSO, the *reaction mixture*, was prepared by adding 12 μL of the *individual stock* of **1h** to 228 μL of the *pre-oxidised sample*. This addition was set as the starting time. At the times of 40 min and >24 h the *reaction mixture* was analysed by HPLC (Figures 2.4b-c and 2.4e-f in section 2.3.5).

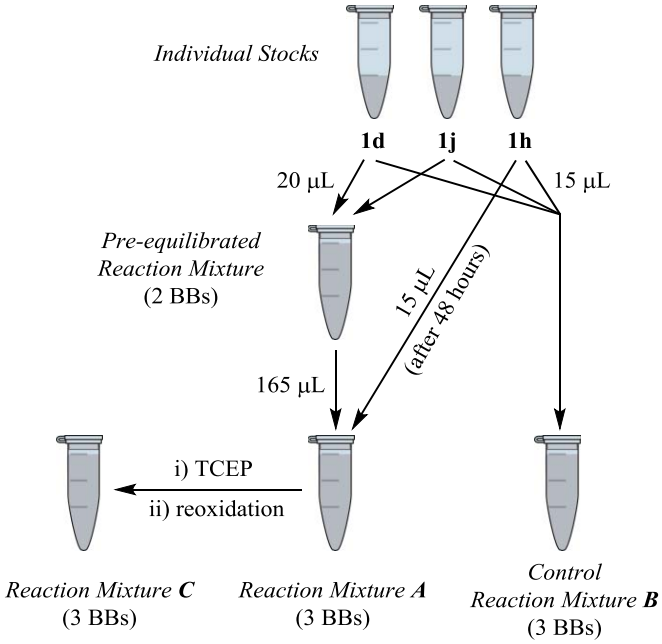


Scheme 2.5. Preparation of the solutions of the evaluation of the exchange rate.

2.5.7. Larger DCL in the limit experimental conditions

Individual stocks of each BB **1d**,**h**,**j** (24 mM) were prepared in pure DMSO (Scheme 2.6). From these, a *pre-equilibrated reaction mixture* was prepared by adding 20 μL of **1d** and **1j** to 180 μL of a pH 4.5 buffer solution. The *individual stock* of **1h** was stored at –80 ºC. After 48 hours, a 2 mM of each BB mixture, the *reaction mixture **A***, was prepared by adding 15 μL of the *individual stock* of **1h** to 165 μL of the *pre-equilibrated reaction mixture*.

Simultaneously, a 2 mM mixture of each BB, the *control reaction mixture **B***, was prepared by mixing 15 μL of each *individual stock* with 135 μL of a pH 4.5 buffer solution. After 48 hours, the *reaction mixture **A***, the *control reaction mixture **B*** and the *pre-equilibrated reaction mixture* were analysed by HPLC (Figure 2.5 in section 2.3.6).



Scheme 2.6. Preparation of the solutions of the ternary reversibility test.

Finally, the *reaction mixture* ***C*** was prepared by adding 0.35 equivalents of Tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl)[73] to the completely oxidised *reaction mixture* ***A***. The substoichiometric amount of TCEP allowed the partial reduction of the disulphides present in the mixture. After the reoxidation of the generated free thiols, 48 hours later, the *reaction mixture* ***C*** was analysed by HPLC (Figure 2.15).

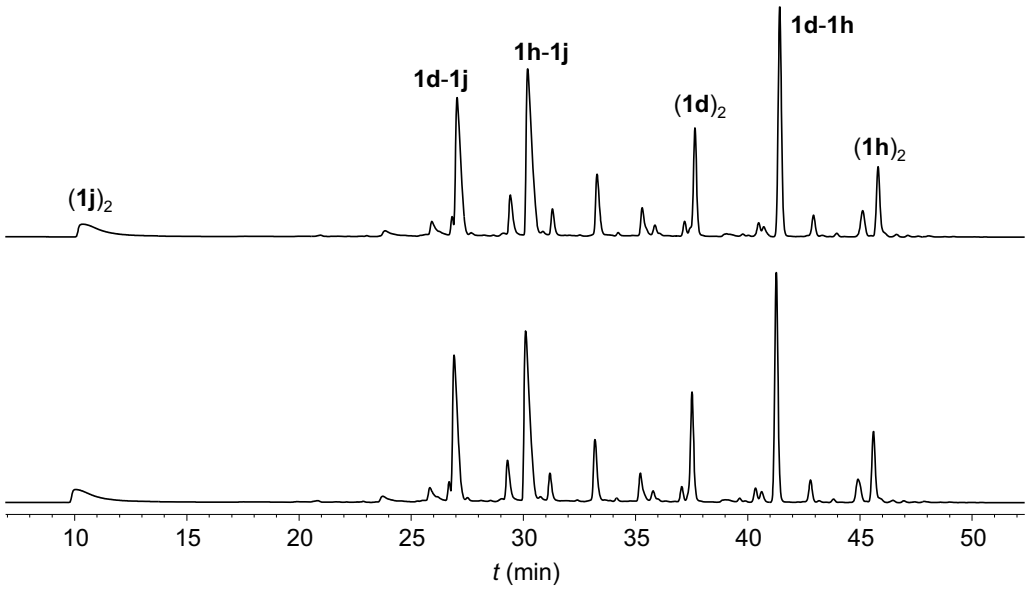


Figure 2.15. HPLC-UV traces (254 nm) of the *reaction mixtures* **A** and **C**.

After the addition of **1h** to the *pre-equilibrated mixture* of **1d+1j**, the mixture evolved to the same final situation as when the three BBs are left to oxidise together. Moreover, after the partial reduction of the *reaction mixture* **A** by means of the TCEP, the subsequent reoxidation also led to the same final situation. These observations were quantitatively verified by means of the $K_{[\mathbf{A},\mathbf{B}]}$ constant, and no significant changes were observed between the constants calculated for the mixtures **A**, **B** and **C** (Table 2.6). Therefore, in the presence of 25% DMSO in water at pH 4.5, the mixture unambiguously proved to reach the thermodynamic equilibrium.

Table 2.6. Values of the $K_{[\mathbf{A},\mathbf{B}]}$ constants calculated for the *mixtures* **A**, **B** and **C**.

	Mixture **A**	*Mixture* **B**	*Mixture* **C**
$K_{[\mathbf{1d},\mathbf{1h}]}$	0.141	0.142	0.144
$K_{[\mathbf{1h},\mathbf{1j}]}$	0.0911	0.0902	0.0909
$K_{[\mathbf{1d},\mathbf{1j}]}$	0.235	0.234	0.232

2.6. References

(1) Singh, R.; Whitesides, G. M. In *Sulphur-Containing Functional Groups (1993)*; John Wiley & Sons, Inc.: 2010, p 633.
(2) Dixon, D. A.; Zeroka, D. J.; Wendoloski, J. J.; Wasserman, Z. R. *J. Phys. Chem.* **1985**, *89*, 5334.
(3) Sevier, C. S.; Kaiser, C. A. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 836.
(4) Zhou, N. E.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1993**, *32*, 3178.
(5) Golik, J.; Clardy, J.; Dubay, G.; Groenewold, G.; Kawaguchi, H.; Konishi, M.; Krishnan, B.; Ohkuma, H.; Saitoh, K.; Doyle, T. W. *J. Am. Chem. Soc.* **1987**, *109*, 3461.
(6) Noiva, R. *Protein Expres. Purif.* **1994**, *5*, 1.
(7) Bardwell, J. C.; Lee, J. O.; Jander, G.; Martin, N.; Belin, D.; Beckwith, J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1038.
(8) Black, S. P.; Sanders, J. K. M.; Stefankiewicz, A. R. *Chem. Soc. Rev.* **2014**, *43*, 1861.
(9) Tam, J. P.; Wu, C. R.; Liu, W.; Zhang, J. W. *J. Am. Chem. Soc.* **1991**, *113*, 6657.
(10) Postma, T. M.; Albericio, F. *Eur. J. Org. Chem.* **2014**, *2014*, 3519.
(11) Cline, D. J.; Thorpe, C.; Schneider, J. P. *Anal. Biochem.* **2004**, *335*, 168.
(12) Chen, L.; Annis, I.; Barany, G. In *Current Protocols in Protein Science*; John Wiley & Sons, Inc.: 2001.
(13) DeCollo, T. V.; Lees, W. J. *J. Org. Chem.* **2001**, *66*, 4244.
(14) Wedemeyer, W. J.; Welker, E.; Narayan, M.; Scheraga, H. A. *Biochemistry* **2000**, *39*, 4207.
(15) Fernandes, P. A.; Ramos, M. J. *Chem. Eur. J.* **2004**, *10*, 257.
(16) Houk, J.; Singh, R.; Whitesides, G. M. In *Methods Enzymol.*; William B. Jakoby, O. W. G., Ed.; Academic Press: 1987; Vol. Volume 143, p 129.
(17) Davidson, S. M. K.; Regen, S. L. *Chem. Rev.* **1997**, *97*, 1269.
(18) Hioki, H.; Still, W. C. *J. Org. Chem.* **1998**, *63*, 904.
(19) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2000**, *122*, 12063.
(20) Stefankiewicz, A. R.; Sanders, J. K. M. *Chem. Commun.* **2013**, *49*, 5820.
(21) Ulatowski, F.; Sadowska-Kuziola, A.; Jurczak, J. *J. Org. Chem.* **2014**, *79*, 9762.
(22) Solà, J.; Lafuente, M.; Atcher, J.; Alfonso, I. *Chem. Commun.* **2014**, *50*, 4564.
(23) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Science* **2002**, *297*, 590.
(24) Fanlo-Virgós, H.; Alba, A.-N. R.; Hamieh, S.; Colomb-Delsuc, M.; Otto, S. *Angew. Chem. Int. Ed.* **2014**, *53*, 11346.
(25) Vial, L.; Sanders, J. K. M.; Otto, S. *New J. Chem.* **2005**, *29*, 1001.
(26) Brisig, B.; Sanders, J. K. M.; Otto, S. *Angew. Chem. Int. Ed.* **2003**, *42*, 1270.
(27) Stefankiewicz, A. R.; Sambrook, M. R.; Sanders, J. K. M. *Chem. Sci.* **2012**, *3*, 2326.
(28) Cougnon, F. B. L.; Jenkins, N. A.; Pantoş, G. D.; Sanders, J. K. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 1443.
(29) Cougnon, F. B. L.; Au-Yeung, H. Y.; Pantoş, G. D.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2011**, *133*, 3198.
(30) Au-Yeung, H. Y.; Pantoş, G. D.; Sanders, J. K. M. *J. Org. Chem.* **2011**, *76*, 1257.
(31) Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. *Science* **2005**, *308*, 667.
(32) Ponnuswamy, N.; Cougnon, F. B. L.; Pantoş, G. D.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2014**, *136*, 8243.

(33) Ponnuswamy, N.; Cougnon, F. B. L.; Clough, J. M.; Pantoş, G. D.; Sanders, J. K. M. *Science* **2012**, *338*, 783.
(34) Malakoutikhah, M.; Peyralans, J. J. P.; Colomb-Delsuc, M.; Fanlo-Virgós, H.; Stuart, M. C. A.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 18406.
(35) Otto, S. *Acc. Chem. Res.* **2012**, *45*, 2200.
(36) Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S. *Science* **2010**, *327*, 1502.
(37) Ura, Y.; Beierle, J. M.; Leman, L. J.; Orgel, L. E.; Ghadiri, M. R. *Science* **2009**, *325*, 73.
(38) Chaudhary, A.; Nagaich, U.; Gulati, N.; Sharma, V.; Khosa, R.; Partapur, M. U. *JAPER* **2012**, *2*, 32.
(39) Smallwood, I. *Handbook of organic solvent properties*; Butterworth-Heinemann, 1996.
(40) Yiannios, C. N.; Karabinos, J. V. *J. Org. Chem.* **1963**, *28*, 3246.
(41) Snow, J. T.; Finley, J. W.; Friedman, M. *Biochem. Biophys. Res. Commun.* **1975**, *64*, 441.
(42) Singh, R.; Whitesides, G. M. *J. Am. Chem. Soc.* **1990**, *112*, 6304.
(43) Singh, R.; Whitesides, G. M. *J. Am. Chem. Soc.* **1990**, *112*, 1190.
(44) Gromova, A. V.; Ciszewski, J. M.; Miller, B. L. *Chem. Commun.* **2012**, *48*, 2131.
(45) Klein, J. M.; Saggiomo, V.; Reck, L.; Luning, U.; Sanders, J. K. M. *Org. Biomol. Chem.* **2012**, *10*, 60.
(46) Cougnon, F. B. L.; Ponnuswamy, N.; Jenkins, N. A.; Pantoş, G. D.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2012**, *134*, 19129.
(47) Rauschenberg, M.; Bomke, S.; Karst, U.; Ravoo, B. J. *Angew. Chem. Int. Ed.* **2010**, *49*, 7340.
(48) Perez-Fernandez, R.; Pittelkow, M.; Belenguer, A. M.; Lane, L. A.; Robinson, C. V.; Sanders, J. K. M. *Chem. Commun.* **2009**, 3708.
(49) Corbett, P. T.; Tong, L. H.; Sanders, J. K. M.; Otto, S. *J. Am. Chem. Soc.* **2005**, *127*, 8902.
(50) Otto, S.; Engberts, J. B. F. N. *Org. Biomol. Chem.* **2003**, *1*, 2809.
(51) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2001**, *106*, 521.
(52) Blokzijl, W.; Engberts, J. B. F. N. *Angew. Chem. Int. Ed.* **1993**, *32*, 1545.
(53) Hafezi, N.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 12861.
(54) Rodriguez-Docampo, Z.; Eugenieva-Ilieva, E.; Reyheller, C.; Belenguer, A. M.; Kubik, S.; Otto, S. *Chem. Commun.* **2011**, *47*, 9798.
(55) Sarma, R. J.; Otto, S.; Nitschke, J. R. *Chem. Eur. J.* **2007**, *13*, 9542.
(56) Saur, I.; Scopelliti, R.; Severin, K. *Chem. Eur. J.* **2006**, *12*, 1058.
(57) Otto, S.; Kubik, S. *J. Am. Chem. Soc.* **2003**, *125*, 7804.
(58) Mukerjee, P.; Ostrow, J. D. *Tetrahedron Lett.* **1998**, *39*, 423.
(59) McIlvaine, T. *J. Biol. Chem.* **1921**, *49*, 183.
(60) Elving, P. J.; Markowitz, J. M.; Rosenthal, I. *Anal. Chem.* **1956**, *28*, 1179.
(61) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70.
(62) Wu, C.; Leroux, J.-C.; Gauthier, M. A. *Nat. Chem.* **2012**, *4*, 1044.
(63) Wu, C.; Belenda, C.; Leroux, J.-C.; Gauthier, M. A. *Chem. Eur. J.* **2011**, *17*, 10064.
(64) Beeren, S. R.; Pittelkow, M.; Sanders, J. K. M. *Chem. Commun.* **2011**, *47*, 7359.
(65) Mindell, J. A. *Annu. Rev. Physiol.* **2012**, *74*, 69.
(66) Capasso, S. *Thermochim. Acta* **1996**, *286*, 41.
(67) Wright, H. T. *Protein Eng.* **1991**, *4*, 283.
(68) Geiger, T.; Clarke, S. *J. Biol. Chem.* **1987**, *262*, 785.

(69) Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S. *J. Chem. Soc., Chem. Commun.* **1992**, 919.
(70) Radkiewicz, J. L.; Zipse, H.; Clarke, S.; Houk, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 3499.
(71) Brennan, T. V.; Clarke, S. *Int. J. Pept. Protein Res.* **1995**, *45*, 547.
(72) Tôrres, R. B.; Marchiore, A. C. M.; Volpe, P. L. O. *J. Chem. Thermodyn.* **2006**, *38*, 526.
(73) Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M. *J. Org. Chem.* **1991**, *56*, 2648.

CHAPTER 3

The emergence of halophilic evolutionary patterns from a dynamic combinatorial library of macrocyclic pseudopeptides

The studies comprehended in this Chapter resulted in the following publication:

Atcher, J.; Moure, A.; Alfonso, I. *Chem. Commun.* **2013**, 49, 487.

3.1. Introduction

3.1.1. Precedents of bio-inspired dynamic covalent systems

The study of the adaptive behavior of dynamic chemical libraries becomes especially interesting if the members of the systems have a conceptual resemblance with the biochemical world. Thus, the structural and chemical information that is expressed during the adaptation of bio-inspired dynamic systems can have a fundamental relevance for the better understanding of a specific molecular evolutionary process. The following lines shortly comment on two examples of dynamic covalent systems with a bio-inspired design: self-replication and molecular walkers.

3.1.1.1. Self-replication

A relevant field in which the adaptive chemistry has tremendously important (potential) applications is the prebiotic chemistry[1] and the investigation of new perspectives for the origin of life.[2-3] In this regard, a property that is undoubtedly essential for the emergence of life is the ability to self-replicate and many theories place replication before metabolism as the initially emergent process of life.[4] With the aim to investigate self-replication, some scientists have used DCLs as suitable chemical environments in which self-replication phenomena can take place. The essence of self-replication, like in most of the processes typically studied in DCC, is a templated synthesis, which intrinsically entails a recognition event.[5] The differential characteristic of the replicating systems is that in self-replication the product of the templated synthesis is the template for its own formation. In the context of a DCL where a mixture of library members is in rapid exchange, this autocatalytic process can shift the equilibrium of the whole molecular network towards the formation of the self-replicating members at the expenses of the other species of the library.[6]

Some remarkable examples have recently appeared in the literature exploring self-replication processes in DCC. Thus, Sadownik and Philp coupled a replicating system to a small DCL of competing reactants and demonstrated that a replicating template is capable of exploiting and dominating an exchanging pool of reagents in order to amplify its own formation,[7] Giuseppone described a dynamic replicator system where reversible formation of an amphiphilic molecule is able to promote the formation of itself by accelerating dissolution of the starting material,[8] and Ulijn reported an enzyme-

mediated peptide synthesis in which gelation drives formation of a single product from a dynamic mixture.[9-10]

Otto and co-workers reported self-replicating pseudopeptidic macrocycles that emerge from a small DCL and compete for a common feedstock.[6,11] In this remarkable example, a dynamic network was generated by the oxidation of the dithiol building block shown in Figure 3.1a, rendering a library of oligomers initially dominated by the corresponding trimeric and tetrameric macrocyclic species (Figure 3.1b). Very interestingly, after some days the hexameric and heptameric species suddenly emerged, growing exponentially to quickly dominate the library (Figure 3.1c). These two species were found to aggregate into thin micron-long fibers (Figure 3.1d). Whether it is the hexamer or the heptamer that dominates the composition of the library depends upon the applied mechanical stimulus: shaking gives predominantly the hexamer while stirring favors the heptamer. The authors were able to give a satisfactory explanation for such an unexpected behavior, and additionally examined the effect of different peptide chains[12] and solvent compositions.[13]

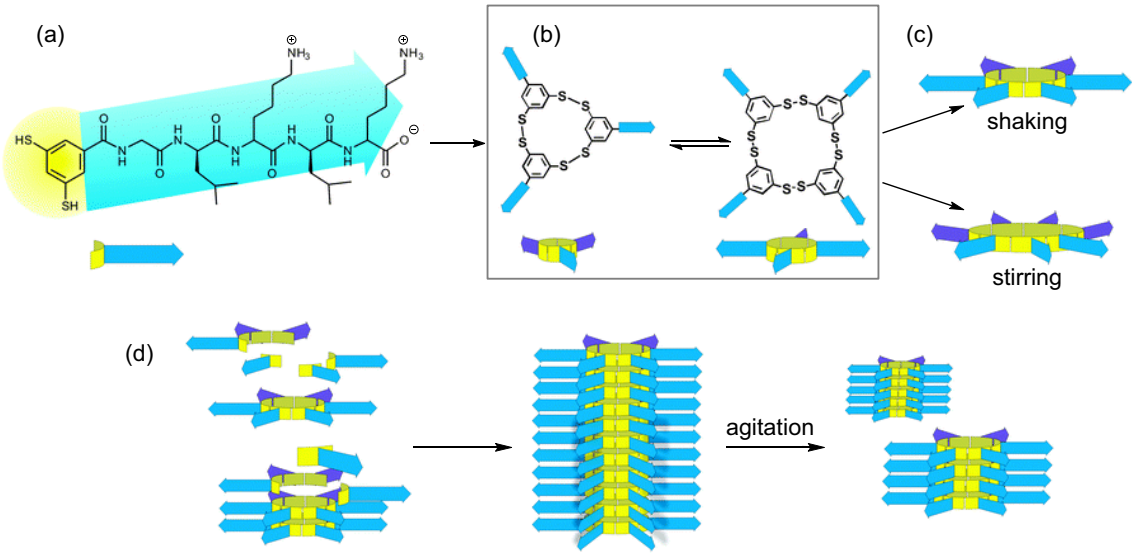


Figure 3.1. (a) Peptide-functionalized dithiol building block, that initially gives rise to (b) a mixture dominated by cyclic trimer and tetramer. (c) Upon agitation a conversion to larger macrocycles takes place. (d) Schematic representation of the stacks of the hexamer, held together by β-sheets formed by the peptide chains. Breaking the fibers by agitation increases the number of ends and promotes fiber growth (Figure modified from reference[11]).

Replicating systems based on the self-assembly of coiled-coil peptides have been studied by Ghadiri and other scientists, often using the irreversible native chemical ligation (NCL) to drive the replication network.[14-19] Recently, however, Ashkenasy and co-workers reported the use of the coiled-coil structural motif for generating replicating

networks under partial thermodynamic control.[20-23] In order to allow reversibility, the authors implemented the transesterification reaction for the ligation process, thus introducing a thioester bond in the peptide sequence. When comparing the high-resolution crystal structure of the thioester coiled-coil protein with the native analogue, minimal differences were observed along the entire structure, except for some deformation around the thioester bond.[22] The reversible ligation reaction was put into practice to generate replicating binary networks under partial thermodynamic control, demonstrating that the outcome of the replication could be influenced by both chemical and physical inputs (templates and light respectively).[23] Additionally, in a larger network consisting of the simultaneous replication of six coiled-coil peptide mutants, the authors were able to illustrate that product formation in such a complex system is governed by the interplay of competitive and cooperative intermolecular relationships.[20]

3.1.1.2. Molecular walkers

Another important field in which scientists have used Nature as a source of bio-inspiration is the development of synthetic molecular walkers. Movement is one of the crucial biological tasks in living organisms and, at the molecular level, this task is performed by biological molecular motors or machines.[24] Among these molecular motors, myosin, kinesin and dynein are "walking proteins" that convert chemical energy[a] into mechanical work ("walking motion" in this case). These proteins have two feet that step forward by repeated cycles of conformational changes in multiple protein domains.[24] This impressive bio-molecular system has been the source of inspiration for the development of synthetic molecular walkers within the frame of DCvC.[25-26]

The first example of a molecule in which a molecular fragment can walk along a track was reported by Leigh and co-workers[27-28] and consisted of a small compound containing two legs with two distinct feet: one hydrazide foot and one thiol foot. The corresponding track was composed of alternating aldehyde and disulfide footholds, and the motion was achieved by the sequential performance of the reversible hydrazone and disulfide chemistries. In this way, one of the legs is always attached to the track, while the other can move freely. By oscillating equilibration conditions, it is possible to cycle between movements of both legs (Figure 3.2). Unfortunately, the movement of this walker is fully random and the final distribution is determined by the relative free

[a] Walking proteins are powered by the hydrolysis of adenosine triphosphate (ATP).

energy of the different states. In order to overcome this limitation, a photoswitchable azobenzene group was incorporated in the middle of the track.[29] By switching the azobenzene at the appropriate part of the cycle, certain control over the direction of the movement was possible.

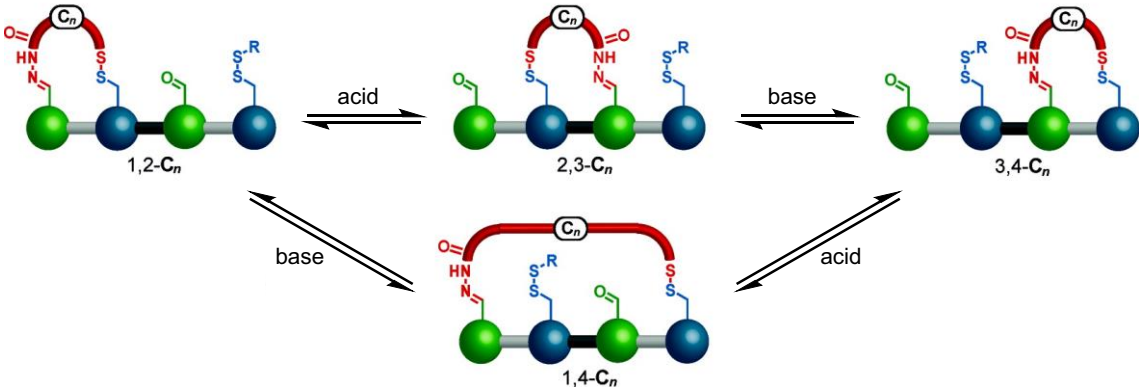


Figure 3.2. Schematic representation of the reversible reactions that allow the "walking motion". Cycling between acidic and basic conditions enables the walker to make steps by means of the orthogonal hydrazone and disulfide reversible linkages respectively (Figure modified from reference[27]).

Leigh's research group has also explored the Michael and retro-Michael reactions to prepare a system able to walk along a track without external intervention.[30] A similar behavior has also been observed by Kovaříček and Lehn in a system composed of a polyamine track and an aldehyde walker capable of movement by means of the imine reversible chemistry.[31] In both examples the motion was achieved without the sequential addition of chemical reagents.

3.1.2. The halophilic proteins

Among all the evolutionary processes observed in Nature, the one suffered by extremophiles is especially appealing, since it has allowed them to survive under very extreme conditions.[32] Halophilic archea are extremophiles that thrice in highly saline environments[b] such as natural salt lakes and marine salterns, *e.g.* the Great Salt Lake and the Dead Sea.[33] To counterbalance the external osmotic pressure, these microorganisms accumulate in their cytoplasm salt concentrations close to saturation, reaching intracellular concentrations of up to 4 M.[34] Accordingly, their bio-molecular machinery has evolved to be stable and fully functional at very high concentrations of salt.[35-39]

[b] Halophilic archea grow optimally at ~4.5 M NaCl.

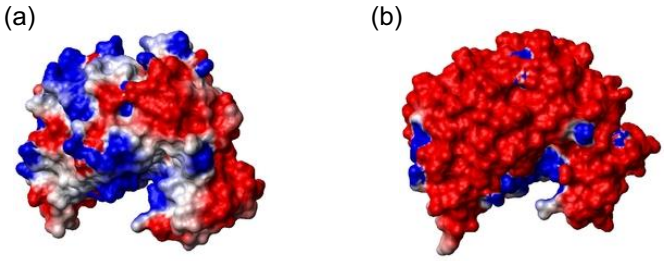


Figure 3.3. Difference in surface charge distribution between the non-halophilic[40] (a) and halophilic[41] (b) malate dehydrogenases (PDB IDs 1GUZ and 1O6Z respectively). Acidic and basic residues are marked in red and blue, respectively.

Proteins from the halophilic organisms, the so-called halophilic proteins, have evolved to maximize stability and activity at hypersaline media.[36,42] Whereas non-halophilic proteins tend to unfold and/or aggregate under such conditions, the halophilic counterparts not only tolerate but also are actually stabilized by high salinity. In recent years, comparative analyses between the proteomes of halophilic and non-halophilic organisms have recognized a characteristic signature in the amino acid composition of proteins with hypersaline adaptation.[35,43] Salt dependency is conferred exclusively by surface residues[43-48] (Figure 3.3) and the characteristic features of the halophilic proteins include four main changes. The first, and most characteristic one, is a large increase in acidic amino acids.[33,35,43-44] Glutamic acids and, more frequently, aspartic acids are known to increase their presence on the surface of the halophilic proteins, resulting in a large negative net charge at neutral pH. The second characteristic feature is the drastic drop in the number of lysines, often replaced by arginines.[44,49] The third feature is a decrease in the overall hydrophobic content.[35,44,50] The fourth and last characteristic associated to the halophilic proteins is a decrease in the accessible surface area (ASA).[44]

Although the effect of these structural changes is still not fully understood, it has been suggested that haloadaptation is primarily due to the binding of networks of hydrates ions with the high surface density of carboxylate groups.[51-52] These interactions with the surface residues would be the responsible for the stabilization of the folded conformation, being the accessible surface area a key parameter.[44] However, this widely held belief has recently been challenged by some authors,[53] and several investigations are currently focused on this topic. Understanding the haloadaptation mechanism is of particular interest given the influence of salt on function, folding, oligomerization and solubility of proteins, and has important potential applications in the biotechnological industry.

3.2. Objectives and hypothesis

The main objective of the present Chapter is to investigate the potential ability of a bio-inspired dynamic combinatorial library to show adaptive trends in parallel with the much more complex evolution of biological systems. We hypothesized that, when exposed to the same external stimulus, simple DCLs of pseudopeptidic macrocycles would display evolutionary trends in parallel with those observed for the biological evolution of the halophilic proteins.

This main objective can be divided into the following three specific aims:

i) To design and prepare a bio-inspired minimalistic DCL with pertinent peptide-like information.

ii) To study the adaptation of this dynamic library to the addition of salt as a bio-inspired external stimulus. The analysis of the salt-induced changes on the composition of the library would allow identifying which species are favored by the increase of the salt concentration.

iii) To stablish a parallelism between the adaptation of this small DCL and the natural evolution of the halophilic proteins. With this aim, the behavior of the library members would be evaluated depending on different factors, including the content of acidic residues and the accessible surface area.

3.3. Results and discussion

3.3.1. Adaptation in DCC *vs.* biological evolution

The case of the halophilic proteins is especially appealing for the experimental modeling of natural evolutionary processes by means of a DCL, mainly for two reasons. Firstly, a very simple external stimulus is known to be the only driving force of the whole evolutionary process. Secondly, the adaptive changes of the halophilic proteins are well established and characterized (section 3.1.2).

But, does it really make any sense to try to establish a parallelism between the adaptive trends of a DCL and the evolutionary changes described for the halophilic proteins? DCLs adapt their composition by means of a very simple expression of "chemical evolution" based on selection and amplification mechanisms,[54] while the evolution of halophilic proteins implies much more complex genetic processes.[35] Hence, their adaptive/evolutionary pathways are radically different. Moreover, in essence, DCLs are designed to work under thermodynamic control, whereas life is a far-from-equilibrium system. In view of the foregoing, the present study would seem not to make any sense. However, there is something that makes the two processes quite similar in a way: both are driven by the increase in stability of the whole system. This is always true for the DCLs operating under thermodynamic control; and also for the halophilic proteins, that have evolved in order to increase their stability in hypersaline media, thus retaining their folding and function.

Although there are several reported examples of dynamic chemical systems designed with a clear biological source of inspiration (section 3.1.1), to the best of our knowledge this is the first study in which a DCL is specifically designed to be able to show adaptive trends in parallel with the evolution of a biological system.

3.3.2. Design and composition of the DCL

For generating a DCL with the potential ability to display adaptive trends in parallel with the evolution of the halophilic proteins, it is crucial to make sure that the corresponding BBs have the suitable structural information. In this regard, the BBs synthesized within the present thesis are all provided with pertinent "peptide-like" information, as they are equipped with two units of an amino acid moiety.

Considering the behavior of the representative BBs **1d**, **1h** and **1j** in terms of the nature of the macrocyclic disulfides generated upon oxidation (Chapter 2), the dynamic library of the present study was also envisioned to be mainly composed by the corresponding macrocyclic dimers. Taking this consideration into account, a minimalistic DCL was conceived as the mixture of only three dithiols. This is the minimum number of BBs needed for the generation a sufficiently diverse dimer-based DCL, able to show characteristic amplification patterns for the selection/stabilization of a single member.[c]

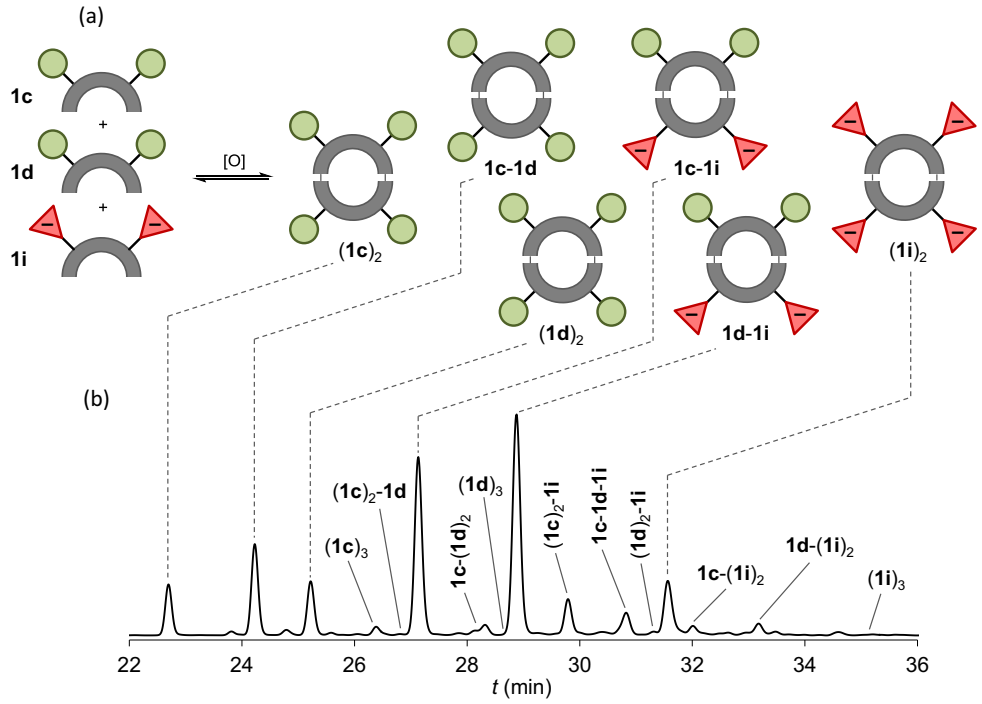


Figure 3.4. Schematic representation (a) and HPLC-UV trace (254 nm) (b) of the DCL generated by the oxidation of **1c**+**1d**+**1i** (2 mM each) in aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO. For simplicity, only the dimers (the predominant species) are depicted in the schematic representation.

The three chosen BBs were the one based on Glu (**1i**), used as a probe containing an anionic amino acid at neutral pH; the one based on Gln (**1c**), as an isostructural non-charged congener; and the one based on Ser (**1d**), as an innocent non-charged derivative. The oxidation of these three dithiols was performed at 2 mM each in 20 mM[d]

[c] In the case of a smaller system, *i.e.* a dynamic mixture generated from only two BBs (**A** and **B**), the effect of the mass balance does not allow distinguishing different changes in relative stability of the two corresponding homodimers (**AA** and **BB**). On the contrary, a ternary system allows the unambiguous identification of the selection/stabilization of a single member (*e.g.* section 3.3.3).

[d] The concentration of the buffer is quite low, at least compared with the McIlvine buffer previously used in Chapter 2. An initially lower concentration of electrolytes should allow observing larger effects caused by the addition of salt (section 3.3.3).

aqueous phosphate buffer (pH 7.5) containing 25% (v/v) DMSO. Once the system reached the equilibrium composition, it was subsequently analyzed by HPLC-UV and UPLC-MS, allowing the unambiguous assignment[e] of the species formed in the DCL (Figure 3.4). All possible macrocyclic homo- and heterodimers (six) were identified as the major species of the library. Additionally, all possible trimers (ten) were also detected in relatively small amounts.

3.3.3. Adaptation of the DCL to the increase of the salt concentration

Once the DCL was properly characterized, we investigated the effect of the addition of salt as a bio-inspired external stimulus. In the frame of DCC, this stimulus has previously been studied with completely different purposes (section 4.1.1). Thus, we performed the same DCL as the one represented in Figure 3.4, now in the presence of increasing concentrations of salt (0.5-2.0 M NaCl). Very interestingly, the subsequent HPLC-UV analyses showed that the composition of the library is notably altered by the presence of the electrolyte (Figure 3.5a). Moreover, the pertinent reversibility tests (section 3.5.4) confirmed that the salt was operating under thermodynamic control and, thus, the observed variations in concentration can be directly attributed to salt-induced changes in the relative stability of the members.

In order to quantitatively evaluate the concentration changes observed in Figure 3.5a, the amplification factor (AF)[55-58] was calculated for the six dimers of the library and for each of the tested concentrations of salt. The AF is typically defined as the ratio between the concentration[f] of a specific library member in the presence of the external stimulus and its concentration in the absence of the external stimulus. However, in this particular case a slight variation was introduced: the concentrations used were relative to the overall population of dimers (Equation 3.2 in section 3.5.3.2). By calculating the AF in this way, any change in the overall trimers/dimers concentrations ratio would not effect on the calculus. Thus, the obtained AF values exclusively depend on the changes in the relative concentration of the six dimers.

[e] The MS assignment was further corroborated by dynamic deconvolution experiments (section 3.3.4).
[f] For calculating the AF the actual concentration is not needed. The concentrations ratio (with and without the external stimulus) is the same as the readily available ratio between the corresponding HPLC peak areas.

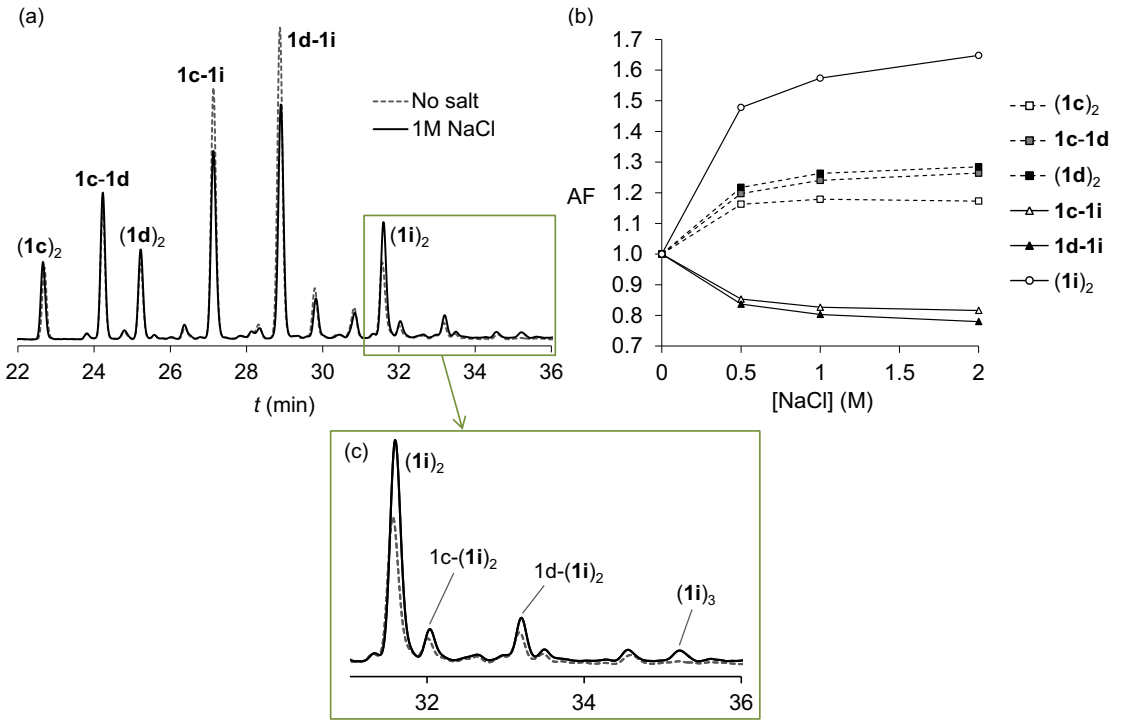


Figure 3.5. (a) HPLC-UV traces (254 nm) of the DCL generated by the oxidation of **1c**+**1d**+**1i** (2 mM each) in aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO in the absence of salt (dashed gray line) and in the presence of 1.0 M NaCl (continuous black line); (b) representation of the amplification factor of the six dimers as a function of the [NaCl] (M); and (c) zoom of (a).

The representation of the AF of each dimer as a function of the salt concentration unambiguously showed that, as the salt content increases, the library progressively adapts its composition following a certain pattern[56] (Figure 3.5b). Comparing the library performed without salt and the library containing 2.0 M NaCl, it is obvious that the salt induces a drastic increase in the presence of the homodimer (**1i**)$_2$ (up to ~60% increase, circles in Figure 3.5b), a decrease in the concentration of the heterodimers containing **1i** (~20% decrease, triangles in Figure 3.5b) and an increase in the presence of the dimers exclusively made of **1c** and/or **1d** (~20% increase, squares in Figure 3.5b). Very interestingly, this pattern corresponds to the effect of the single amplification of (**1i**)$_2$. As this homodimer is favored, and hence amplified in the mixture, the whole dynamic library rearranges its distribution in order to fulfill the overall mass balance. Thus, as **1i** is consumed for the production of (**1i**)$_2$, the two heterodimers containing **1i**, *i.e.* **1c-1i** and **1d-1i**, are consequently disfavored. The resulting decrease in their concentration liberates BBs **1c** and **1d**, causing the increase in the presence of those dimers exclusively made of them, *i.e.* (**1c**)$_2$, **1c-1d** and (**1d**)$_2$. The latter are opportunistically favored species, as they increase their concentration in the mixture without being

directly favored by the external stimulus, but as a consequence of the whole rearrangement of concentrations.

Regarding the trimers, their distribution was also found to be affected by the external stimulus. The addition of salt induced the increase of those trimeric macrocycles containing more than one unit of the BB **1i** (Figure 3.5c). The trimers **x**-(**1i**)$_2$ (with **x** = **1c** or **1d**) showed a 40-60% increase in concentration, while (**1i**)$_3$, which was barely detected in the absence of salt, showed a fourfold increase of its concentration in 2.0 M NaCl.

In summary, the DCL generated by the mixture of **1c**+**1d**+**1i** adapts its composition to the increase of the salt content following an amplification pattern in agreement with the selection of the **1i**-richest members. The salt induces the stabilization of those library members containing more than one unit of the Glu-based BB **1i**, *i.e.* the dimer (**1i**)$_2$ and the trimers **1c**-(**1i**)$_2$, **1d**-(**1i**)$_2$ and (**1i**)$_3$. Very remarkably, the hypersaline environment induces the adaptation of the DCL towards the amplification of the macrocycles concentrating the larger number of acidic residues, in a clear parallelism with the biological evolution of the halophilic proteins.

3.3.4. Dynamic deconvolution experiments

In order to corroborate the stabilization of (**1i**)$_2$ as the cause of the salt-induced amplifications pattern observed for the six dimers (Figure 3.5b), we decided to perform dynamic deconvolution[59-60] experiments. The three possible sublibraries containing only two BBs, *i.e.* binary mixtures, were performed in the absence of salt and in the presence of 1.0 M NaCl. The corresponding HPLC-UV analyses showed that the salt alters the composition of the two binary mixtures containing **1i**, increasing the concentration of the two corresponding homodimers (Figure 3.6a-d); whereas the sublibrary made of **1c**+**1d** was insensitive to the salt content (Figure 3.6e-f). Hence, the behavior of the three smaller binary mixtures is in perfect agreement with our proposal, confirming the salt-induced stabilization of (**1i**)$_2$ as the ultimate cause of the amplifications pattern observed for the six dimers of the ternary mixture (Figure 3.5).

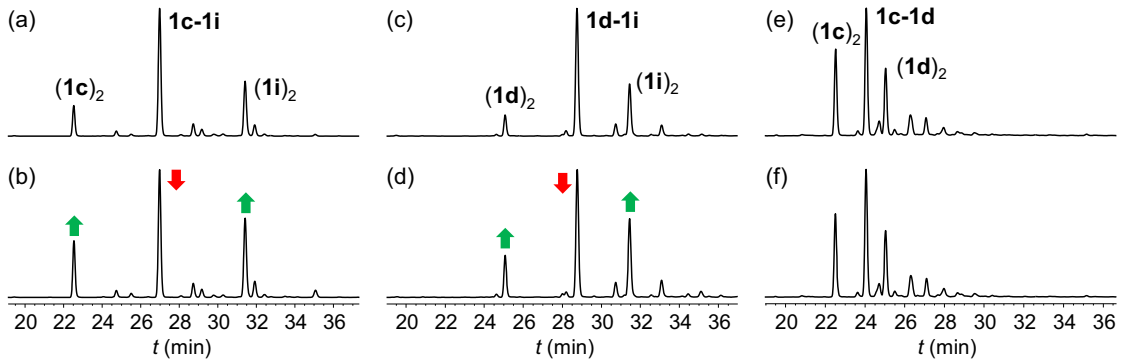


Figure 3.6. HPLC-UV traces (254 nm) of the binary mixtures performed at 2 mM of each BB in 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO. Mixture of **1c**+**1i** in the absence (a) and presence of 1.0 M NaCl (b); **1d**+**1i** in the absence (c) and presence of 1.0 M NaCl (d); and **1c**+**1d** in the absence (e) and presence of 1.0 M NaCl (f).

Additionally, once confirmed the stabilization of (**1i**)$_2$, the two binary mixtures containing **1i** (Figure 3.6a-d) also allowed to get some valuable quantitative information of the corresponding increase in stability. We calculated that the dimer (**1i**)$_2$ was stabilized in 1.0 M NaCl by about $\Delta\Delta G° = -2.6$ kJ·mol^{-1}.[g]

3.3.5. Additional experiments

3.3.5.1. Effect of the salt concentration on the (**1i**)$_3$/(**1i**)$_2$ proportion

A simple experiment was envisioned to confirm the salt-induced stabilization of those library members concentrating the largest number of the Glu-based BB. Dithiol **1i** (2 mM) was oxidized alone in a 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO. Once the equilibrium composition was reached, the HPLC-UV and UPLC-MS analyses confirmed the dimer (**1i**)$_2$ as the major member of the library (peak at 31.9 min in Figure 3.7a), and the trimer (**1i**)$_3$ as a very minor species (peak at 35.5 min in Figure 3.7a, <3% of the overall HPLC area).

Interestingly, when performing the same DCL in the presence of increasing concentrations of salt (0.5-2.0 M NaCl, Figure 3.7b-d), the library showed a clear increase in the presence of the trimer. This trend was quantitatively evaluated by means of the corresponding trimer/dimer exchange constant ($K_{T/D}$, Equation 3.1), which exhibited a tenfold increase in 2.0 M NaCl (Figure 3.7).

[g] The calculation details can be found in section 3.5.5.2.

$$3\ (\mathbf{1i})_2 \rightleftharpoons 2\ (\mathbf{1i})_3 \quad K_{T/D} = \frac{[(\mathbf{1i})_3]^2}{[(\mathbf{1i})_2]^3} \qquad \textbf{(Equation 3.1)}$$

As previously commented in section 3.3.3, both the dimer $(\mathbf{1i})_2$ and the trimer $(\mathbf{1i})_3$ are stabilized (in absolute terms) by the external stimulus. However, in this particular experiment, as there are no other species in the library apart from these two, their changes in concentration are intimately related in a clearly competitive fashion: only one of them can be amplified in the mixture at the expense of the other.[h] Within the field of DCC, macrocyclizations are postulated to maximize the number of species in solution increasing the entropy due to greater disorder. In a DCL like the one presented here, where the species compete for a scarce BB (**1i**), library members that use relatively small numbers of that BB have a competitive advantage over those members that contain larger number of the particular BB.[55] Even so, in this case the larger of the two present oligomers, *i.e.* the trimeric species, is the "winner" of the adaptive competing process. This behavior indicates a much better salt-adaptation of the trimer with respect to the dimer. In summary, this experiment allowed confirming the salt-induced stabilization of those members concentrating the largest number of the Glu-based BB.

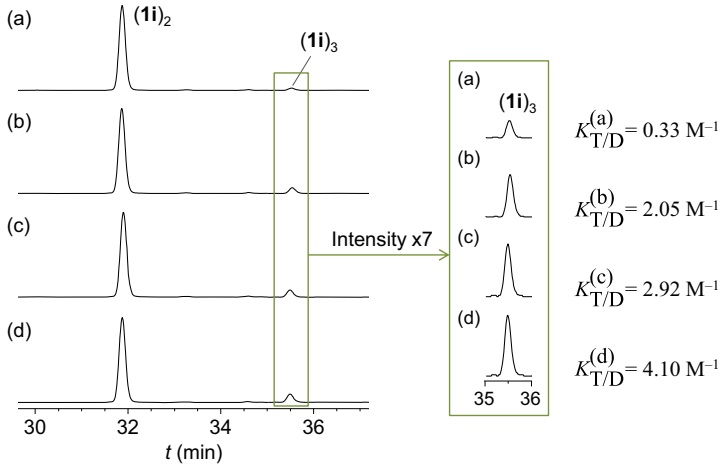


Figure 3.7. HPLC-UV traces (254 nm) of the DCLs generated by the oxidation of 2 mM **1i** in 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO, in the absence of salt (a), and in the presence of 0.5 M NaCl (b), 1.0 M NaCl (c) and 2.0 M NaCl (d). Zoom at the peak of the trimer $(\mathbf{1i})_3$ (sevenfold increase in the intensity) and values of the corresponding $K_{T/D}$ constants (M^{-1}).

[h] Amplifications in a DCL depend on changes in the relative stability of the members.

The same experiment performed for **1c** also led to a DCL consisting of the corresponding dimer $(\mathbf{1c})_2$ and trimer $(\mathbf{1c})_3$, with a ~91% and ~9% of the overall HPLC area respectively. However, in this case the addition of salt did not induce any variation of the trimer/dimer proportion, suggesting the need of charged groups for the salt-induced amplification process, as confirmed in the following section 3.3.5.2.

3.3.5.2. Performance of the DCL at acidic pH

The same DCL as the one represented in Figure 3.4 was also performed at acidic pH. Dithiols **1c**+**1d**+**1i** (2 mM each) were dissolved in 20 mM aqueous phosphate buffer (pH 2.5) with 25% (v/v) DMSO. Once the oxidation of the dithiols was complete, the mixture was found to be composed of the six possible dimers, the ten possible trimers and a little amount of the products of the corresponding intramolecular cyclizations, *i.e.* the cyclic monomeric disulfides *c*-**1c**, *c*-**1d** and *c*-**1i** (Figure 3.8a). We hypothesized that, at such a low pH, the slow disulfides exchange would have kinetically trapped the products of the intramolecular cyclization, probably intermediate species of the oxidation process.

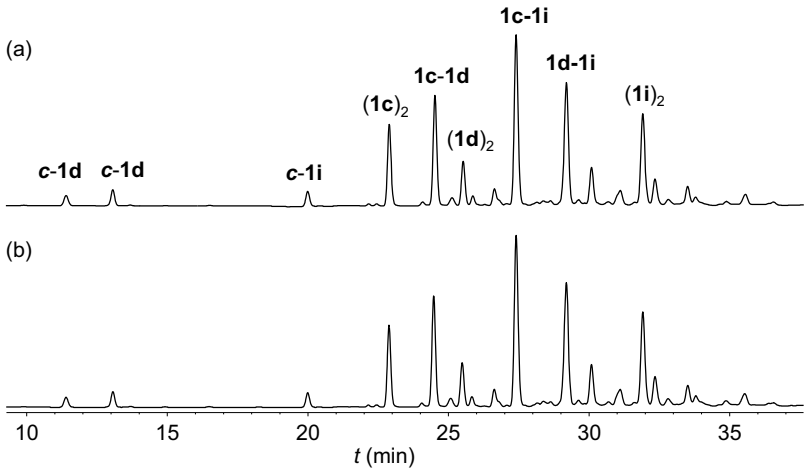


Figure 3.8. HPLC-UV traces (254 nm) of the DCL generated by the oxidation of **1c**+**1d**+**1i** (2 mM each) in aqueous phosphate buffer (pH 2.5) with 25% (v/v) DMSO in the absence of salt (a) and in the presence of 2.0 M NaCl (b).

In spite of the difficulty of reaching the thermodynamic equilibrium at acidic pH (section 2.3.4), the libraries performed in the absence and presence of 2.0 M NaCl showed an identical composition (Figure 3.8a *vs.* 3.8b). Thus, at pH 2.5 the generated DCL was insensitive to the salt content, seeming evident that the salt-induced amplification of the **1i**-richest species requires the carboxylate form of the side chains.

3.3.5.3. The effect of different salts

The last additional experiment consisted in performing again the DCL generated by the mixture of **1c**+**1d**+**1i** but using different salts as external stimuli. The addition of KCl allowed investigating the effect of a different cation, while the addition of $NaNO_3$ allowed investigating the effect of a different anion. In both cases the mixtures rendered practically identical amplification factors than NaCl (*e.g.* NaCl *vs.* KCl in Figure 3.9), implying no specific effects of either the cation or the anion in the amplification process. Thus, this result suggests the increase in ionic strength as the driving force of the whole adaptive process.

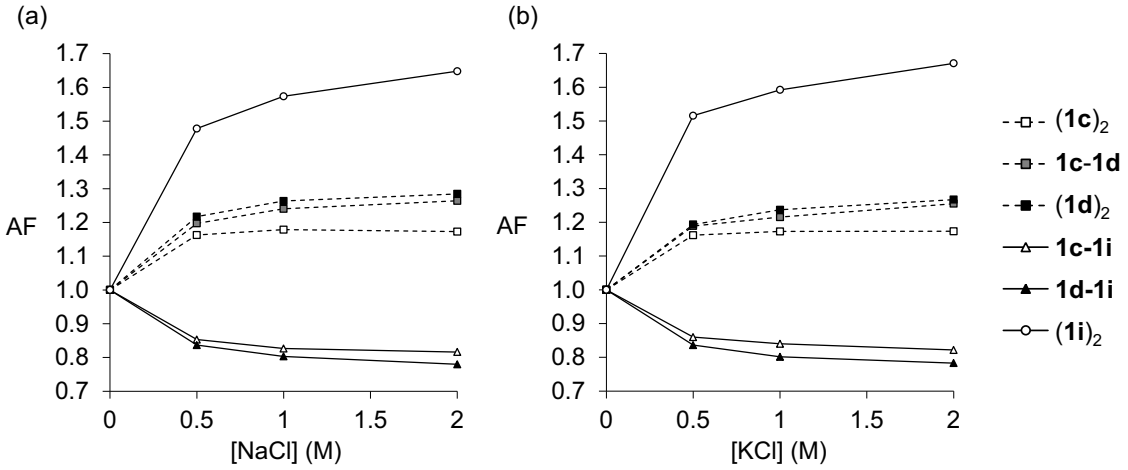


Figure 3.9. Representation of the AF of the six dimers as a function of [NaCl] (M) (a) and [KCl] (M) (b), for the DCL generated by the oxidation of **1c**+**1d**+**1i** (2 mM each) in aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO. The equivalent graphic for $NaNO_3$ (data not shown) also presents an identical amplification pattern.

3.3.6. Structural studies on the dimer (**1i**)$_2$

3.3.6.1. Molecular modeling

In order to better understand the process at the molecular level, the structural characterization of dimer (**1i**)$_2$ was performed, both for the tetraacidic and the fully deprotonated forms. The large flexibility of macrocycles of this size prompted us to initially carry out a Monte Carlo conformational search (10000-20000 geometries without restrictions), followed by a minimization using the MMFFaq[i] force field.[61] Very interestingly, a highly symmetric folded geometry was obtained for the global minimum

[i] This version of the force field (Spartan 06 software) takes into account water solvent as a continuum medium.

of the tetraanionic structure (Figure 3.10c), whereas the tetraacid showed an unfolded flexible conformation (Figures 3.10a-b).

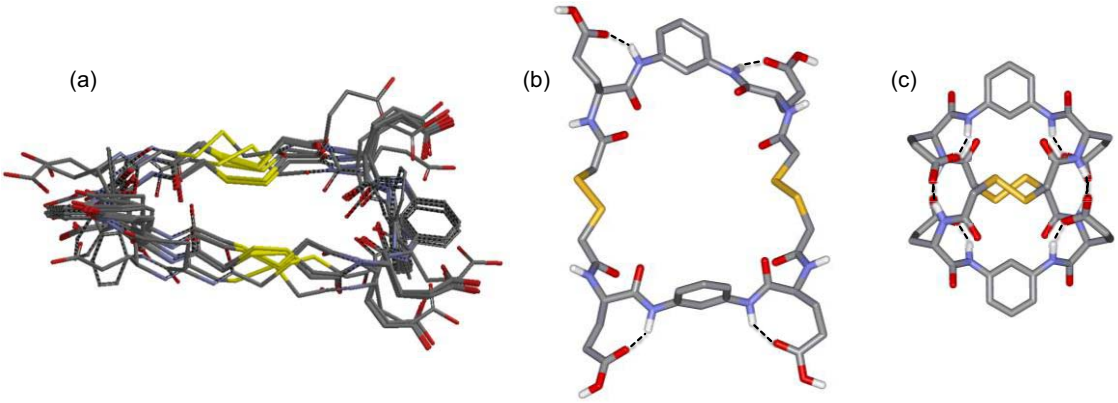


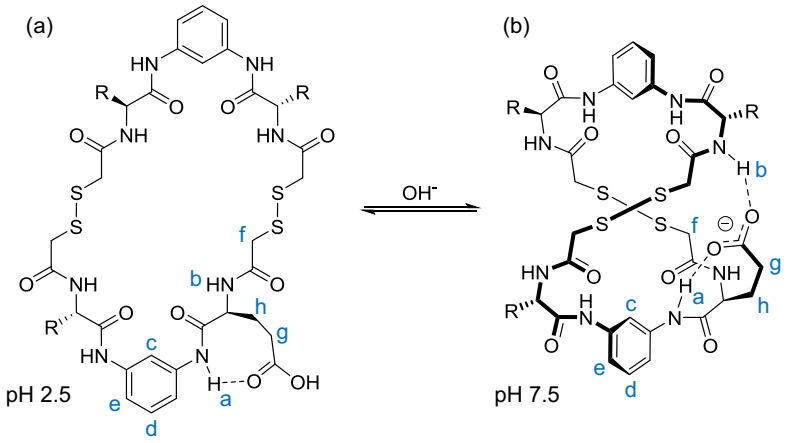
Figure 3.10. (a) Side view of the superposition of the accessible local minima for the tetraacid form of (**1i**)$_2$, showing the high flexibility of the macrocyclic framework. Minimized geometries for (**1i**)$_2$ in the neutral (b) and tetraanionic (c) forms.

This difference can be explained considering the H-bonding patterns. The folded conformation sets eight H-bonds: each carboxylate binds one aromatic and one aliphatic amide, stablishing a figure-eight folding.[62-65] On the other hand, the unfolded conformation only has four H-bonds between each carboxylic acid and the neighboring aromatic amide. Another important structural difference arises from the disposition of the aromatic *m*-diamidophenyl moiety. The folded conformation forces the aromatic amido NH group to point inwards and the corresponding carbonyls outwards. In the flexible acidic form, the rotation of these amide groups set the carbonyls preferentially inwards. Thus, a conformational rearrangement would follow the deprotonation of the side chains, as confirmed in the following section 3.3.6.2.

Later on, in the next Chapter, the conformational preferences of the tetraanionic form of (**1i**)$_2$ are more exhaustively studied by means of molecular dynamics simulations in explicit solvent molecules (section 4.3.5.2).

3.3.6.2. Nuclear magnetic resonance

Three samples were prepared by dissolving **1i** (2 mM) in different conditions: (a) 20 mM phosphate buffer (pH 7.5) with 1.0 M NaCl and 25% (v/v) DMSO-d_6; (b) 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6; and (c) 20 mM phosphate buffer (pH 2.5) with 25% (v/v) DMSO-d_6. The *in situ* oxidation of BB **1i** led to the formation of dimer (**1i**)$_2$ as the main species in the three samples, as observed in the corresponding (H)WET-^{1}H NMR spectra (Figure 3.11a-c).



Scheme 3.1. Proposed conformational rearrangement of $(\mathbf{1i})_2$ produced by a pH change.

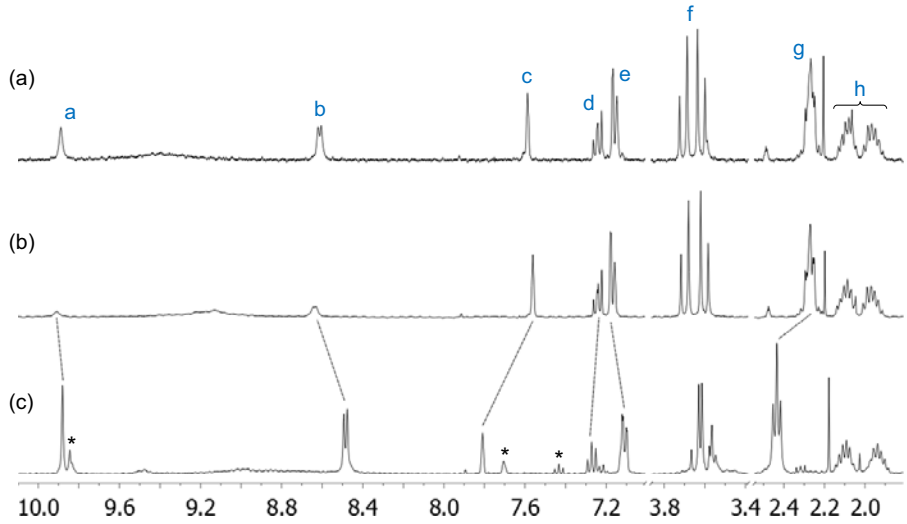


Figure 3.11. Partial (H)WET-^{1}H NMR spectra of oxidized **1i** (2 mM) in: (a) 20 mM phosphate buffer (pH 7.5) with 1.0 M NaCl and 25% (v/v) DMSO-d_6, (b) 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6, and (c) 20 mM phosphate buffer (pH 2.5) with 25% (v/v) DMSO-d_6. The symbol (*) identifies signals of the cyclic monomeric disulfide ***c*-1i**.

Several signals changed upon increasing the pH from 2.5 to 7.5, due to the deprotonation of the side chains (Figure 3.11c *vs.* 3.11b, corresponding to the structures represented in Scheme 3.1). Thus, the signal g, *i.e.* the methylene in α to the carboxylic group, shifts upfield in agreement with the carboxylate formation. Concomitantly, the aliphatic amido NH (b) moves downfield as a result of the establishment of an intramolecular H-bond with the carboxylate, which is geometrically favorable in the folded conformation. Interestingly, the aromatic NH (a) does not significantly change its chemical shift, which supports that this proton is implicated in similar H-bonding at both acidic and neutral pH, as observed in the modeling. Furthermore, the changes observed for the aromatic protons (c, d and e) are also of note, since c and d shift upfield and e downfield by increasing the pH. This behavior can be explained by considering

the rotation of the aromatic amide, which changes the relative disposition of the carbonyl group and the aromatic ring, as observed in the models. Finally, the larger anisochrony of several methylene signals (g and, specially, f) suggests a more rigid conformation of the cycle in the anionic form.

Regarding the spectrum of (**1i**)$_2$ acquired at pH 7.5 in a saline medium, it displayed no differences with respect to the one acquired in the absence of salt (Figure 3.11b *vs.* 3.11a). This interesting observation clearly indicates that the salt is not inducing the folding but stabilizing the already folded structure present at neutral pH.[j] Additionally, the equivalent NMR spectra of the isostructural non-charged Gln-based homodimer (**1c**)$_2$ showed no changes with the pH (Figure 3.12), supporting our proposal for the conformational behavior of (**1i**)$_2$.

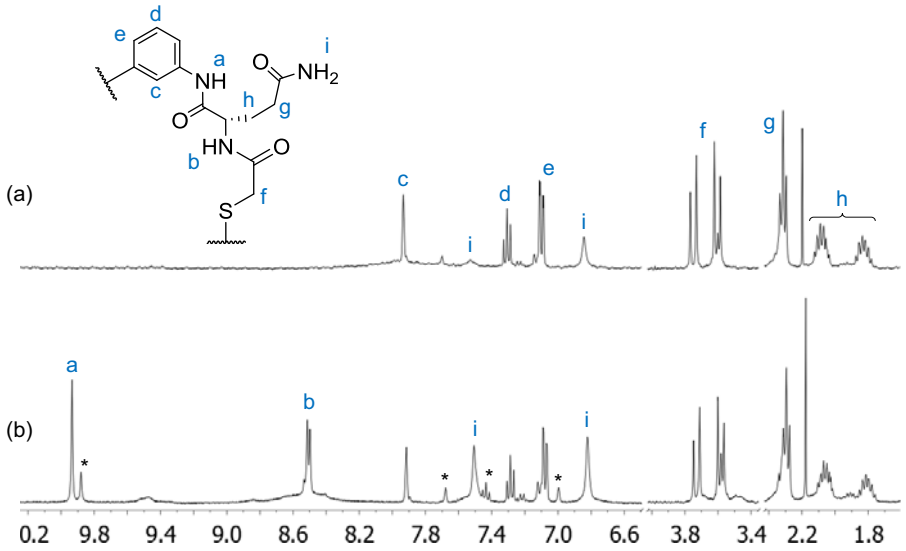


Figure 3.12. Partial (H)WET-^{1}H NMR spectra of oxidized **1c** (2 mM) in: (a) 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6, and (b) 20 mM phosphate buffer (pH 2.5) with 25% (v/v) DMSO-d_6. The symbol (*) identifies signals of the cyclic monomeric disulfide ***c*-1c**.

3.3.6.3. Circular dichroism

Circular dichroism (CD) is a valuable spectroscopic technique for studying the structure of proteins,[66] peptides[67] and pseudopeptides[68-70] in solution because many common conformational motifs have characteristic UV CD spectra. Hence, the CD spectrum of (**1i**)$_2$ was recorded at the three different experimental conditions previously studied by NMR: pH 7.5 in the absence and presence of 1.0 M NaCl, and pH 2.5

[j] In several cases inorganic salts are known to induce partially structured states in proteins and contribute to the stability of the corresponding folded conformation.

(Figure 3.13a; blue, green and red lines respectively). Remarkably, the spectra recorded at different pH values showed significantly different signatures both at ~220 nm and ~254 nm, wavelengths where the amide bond and the aromatic system absorb respectively,[71] implying a different conformation of the macrocycle at acidic and neutral pH. On the contrary, the CD spectrum acquired at pH 7.5 in a saline medium displayed no differences with respect to the one recorded in the absence of salt. These observations are clearly in agreement with the NMR experiments, confirming that the salt does not change the conformation of the dimer (**1i**)$_2$ but stabilizes the already folded conformation present at pH 7.5.

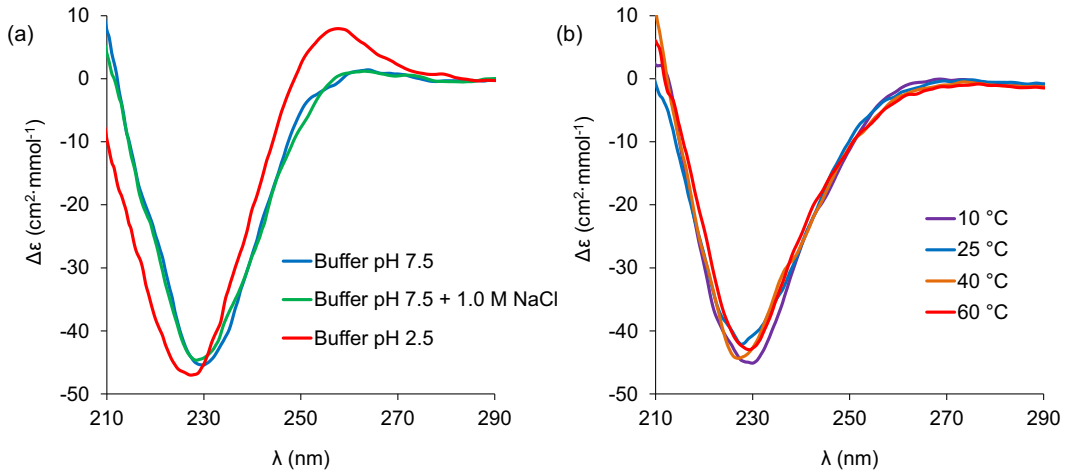


Figure 3.13. (a) CD spectra (Δε *vs.* λ) of (**1i**)$_2$ in aqueous phosphate buffer with 25% (v/v) MeOH at pH 7.5 in the absence of salt (—), at pH 7.5 in the presence of 1.0 M NaCl (—), and at pH 2.5 in the absence of salt (—); (b) CD spectra (Δε *vs.* λ) of (**1i**)$_2$ in aqueous phosphate buffer (pH 7.5) with 25% (v/v) MeOH at different temperatures (10-60 °C).

One of the major applications of CD today is the determination of the thermodynamics of protein folding.[66,72] The use of CD to follow protein denaturation depends on the fact that the change in ellipticity (θ) is directly proportional to the change in concentration of native and denatured forms. Similarly, we envisioned a thermal unfolding experiment for the homodimer (**1i**)$_2$, which would provide us with information about the thermodynamics of the corresponding folding. Consequently, the spectrum of the sample prepared at pH 7.5 in the absence of salt was recorded at different temperatures, from 10 to 60 °C (Figure 3.13b). Unfortunately, the CD signal did not display any significant variation with the temperature and the experiment did not provide us with any useful information regarding the stability of the proposed folding for the tetraanionic form of (**1i**)$_2$.

3.3.7. Incorporation of BB 1h to the DCL

Having characterized the salt-induced amplification trends of the DCL formed by the mixture of the three BBs **1c**+**1d**+**1i**, we wondered about the performance of a more complex dynamic library. We decided to evaluate the effect of the presence of an additional acidic derivative by introducing the Asp-based BB **1h**. Consequently, the four BBs **1c**+**1d**+**1h**+**1i** were mixed at 2 mM each in the same solvent conditions as before, and also in the presence of different concentrations of NaCl (0.0-2.0 M). Once again, as observed by HPLC-UV (Figure 3.14a) and UPLC-MS, the generated DCL mainly consisted of all the possible homo- and heterodimers (ten in total).

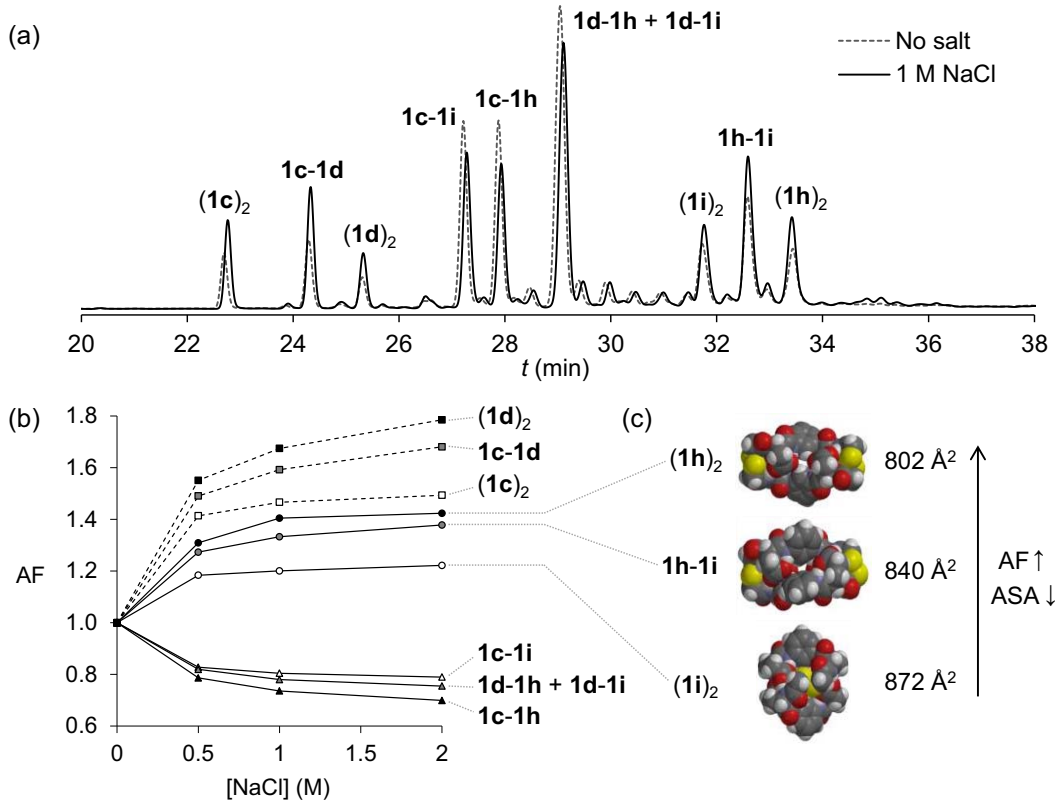


Figure 3.14. (a) HPLC-UV traces (254 nm) of the DCL generated by the oxidation of **1c**+**1d**+**1h**+**1i** (2 mM each) in aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO in the absence of salt (dashed grey line) and in the presence of 1.0 M NaCl (continuous black line); (b) representation of the amplification factor of the ten dimers as a function of the [NaCl] (M); and (c) CPK models for the global minimum calculated for $(\mathbf{1h})_2$, **1h-1i** and $(\mathbf{1i})_2$ with the corresponding ASA values (Å^2).

Regarding the effect of the salt on this larger DCL, the increase of the ionic strength produced a pattern of concentration changes corresponding to the effect of the amplification of all dimers assembling **1h** and/or **1i**, *i.e.* the Glu- and Asp-based BBs (circles in Figure 3.14b). The salt-induced amplification of $(\mathbf{1i})_2$ was somehow

palliated[k] by the simultaneous amplification of the other anionic dimers (**1h**)$_2$ and **1h-1i**, in a clearly competitive process.[l] Moreover, the corresponding amplification factors increase in the order (**1i**)$_2$ < **1h-1i** < (**1h**)$_2$, implying a better salt-adaptation of the Asp derivatives.

This evolutionary trend has been also reported for the halophilic proteins and was related to the smaller accessible surface area (ASA) for the shortest side chain of the Asp amino acid, which permits a more compact folding.[44] Molecular models of the anionic macrocyclic dimers (**1i**)$_2$, **1h-1i** and (**1h**)$_2$ showed a folding stabilized by carboxylate-amide intamolecular H-bonding patterns similar to that found in (**1i**)$_2$ (Figure 3.14c). The corresponding CPK areas follow the trend (**1i**)$_2$ > **1h-1i** > (**1h**)$_2$, and thus, the salt-adaptation of the members of this DCL also increase as the ASA decreases.

[k] In the absence of **1h** the AF of (**1i**)$_2$ in 2.0 M NaCl was ~1.6 (section 3.3.3), while in the presence of **1h** it is ~1.2.

[l] See the following Chapter 4 (section 4.3.6) for a deeper discussion on the competitive co-adaptive relationship.

3.4. Conclusions

A minimalistic DCL of macrocyclic pseudopeptides has been successfully used to reproduce some of the evolutionary trends described for the much more complex biological evolution of the halophilic protein. This is evidenced in the following five conclusions:

i) The addition of salt to a simple DCL of macrocyclic pseudopeptides induces the amplification of the library members concentrating the largest number of acidic side chains, *i.e.* the **1h**/**1i**-richest species. This adaptive process has been studied by the analysis of the amplifications pattern, and has a remarkable resemblance with the natural evolution of the halophilic proteins.

ii) The salt-induced selection and amplification of the **1h**/**1i-**richest species requires the carboxylate form of the side chains.

iii) No specific effects of either the cation or the anion of the added salt are responsible for the observed amplifications, implying the increase in ionic strength as the driving force of the whole adaptive process.

iv) Structural studies, including molecular modeling, NMR and CD experiments, suggest a folded conformation for the amplified dimers $(\mathbf{1h})_2$, **1h-1i** and $(\mathbf{1i})_2$. This is not directly related to the salt-induced stabilization of these three species.

v) The Asp derivatives present a better salt-adaptation than the Glu derivatives. This trend has also a clear similitude with the evolution of the halophilic proteins and is related to the smaller ASA for the shorter side chain of the Asp amino acid.

Overall, these findings show the utility of DCLs for the experimental modeling of the molecular evolutionary trends observed in Nature, with foreseen implications in the prebiotic chemistry and in the understanding of the origin of life.

3.5. Experimental section

3.5.1 General methods

Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka and Merck) and were used without further purification. pH measurements were performed at room temperature on a Crison GLP21 pHmeter with the electrode Crison 50 14T. NMR spectroscopic experiments were carried out on a Varian Mercury 400 instrument (400 MHz for ^{1}H). UV-Vis and CD measurements were carried out at the Scientific and Technologic Center of UB (CCiTUB) using a JASCO J-810 spectropolarimeter (all spectra were recorded at room temperature unless otherwise specified).

3.5.2. HPLC and MS analyses

The HPLC and MS analyses were performed as specified in Chapter 2 (section 2.5.2). The HRMS analyses of the DCLs are shown in the electronic Annex.

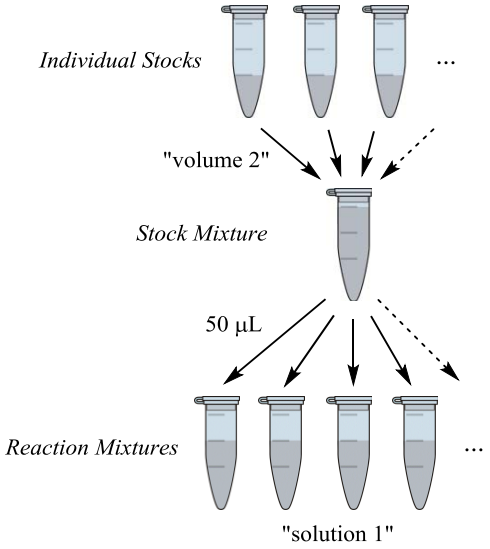
3.5.3. Preparation and evaluation of the DCLs

3.5.3.1. General procedure for the preparation of the DCLs

Individual stocks of each BB ("concentration 1") were prepared by dissolving separately "weight 1" of the corresponding BB in "volume 1" of DMSO (Scheme 3.2). Then a *stock mixture* was prepared by mixing "volume 2" of each *individual stock*. The *reaction mixtures* (2 mM of each BB) were prepared by adding 50 μL of the *stock mixture* to 150 μL of the buffered "solution 1". After 48 hours, the fully oxidized mixtures were quantitatively analyzed by HPLC and UPLC-MS. All the DCLs were performed at least twice, observing no significant differences within the experimental error.

For the preparation of the DCLs containing three BBs, "concentration 1" is 24 mM; "weight 1" is 2.21 mg of **1c**, 1.83 mg of **1d** and 2.20 mg of **1i**; "volume 1" is 170 μL; "volume 2" is 150 μL; and "solution 1" is a 26.7 mM aqueous phosphate buffer (pH 7.5 or 2.5) containing 0.00, 0.67, 1.33 or 2.67 M NaCl/KCl/$NaNO_3$. For the preparation of the DCLs containing four BBs: "concentration 1" is 32 mM; "weight 1" is 2.29 mg of **1c**, 1.92 mg of **1d**, 2.18 mg of **1h** and 2.30 mg of **1i**; "volume 1" is 140 μL; "volume 2"

is 130 µL; and "solution 1" is a 26.7 mM aqueous phosphate buffer (pH 7.5) containing 0.00, 0.67, 1.33 or 2.67 M NaCl.



Scheme 3.2. Procedure for the preparation of the DCLs.

3.5.3.2. Calculation of the amplification factor

For a given dimer "i" with an area A_i in a mixture "j" with an area summation of all dimers A_T, the amplification factor (AF) of "i" in "j" was calculated as shown in Equation 3.2. The use of relative areas (A_i/A_T) instead of the absolute values (A_i) is particularly appropriate in this case because the molar absorptivity of the BBs is known to be the same regardless the other BBs with which they combine (section 2.3.3). Therefore, A_T can only vary if the overall trimers/dimers concentrations ratio changes.

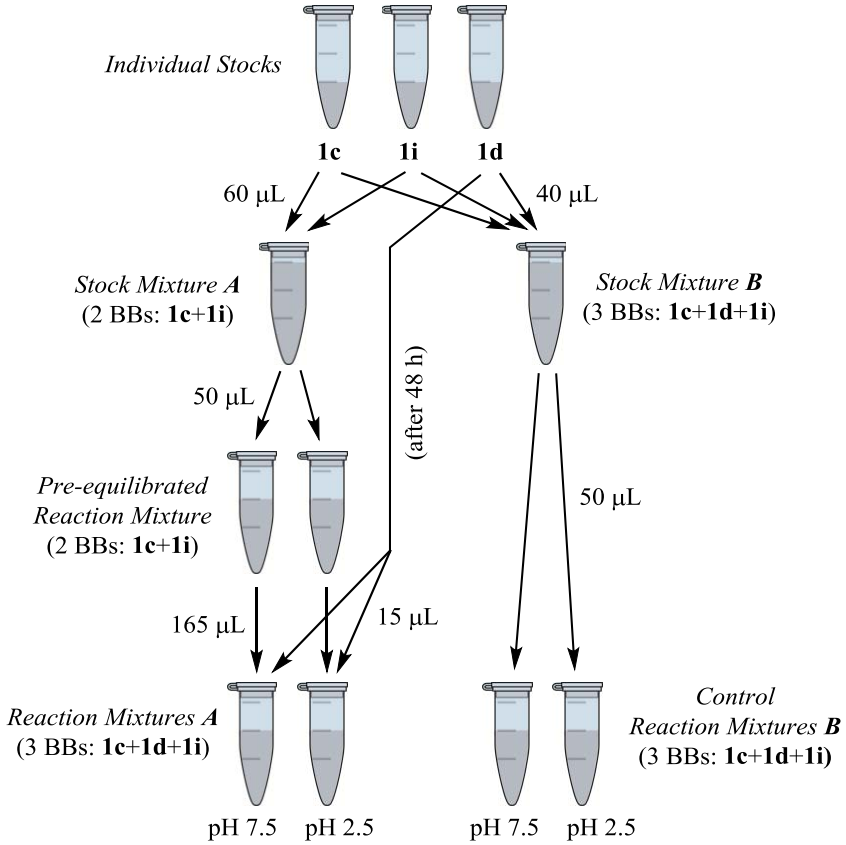
$$AF_{i,j} = \frac{\left(A_i/A_T\right)_j}{\left(A_i/A_T\right)_{No\ salt}} \qquad \textbf{(Equation 3.2)}$$

3.5.4. Reversibility tests

3.5.4.1. Reversibility tests at pH 7.5 and 2.5 in the absence of salt

Individual stocks (24 mM) of each BB **1c**,**d**,**i** were prepared by dissolving separately, 1.47 mg of **1c**, 1.25 mg of **1d** and 1.49 mg of **1i** in 120 µL of DMSO (Scheme 3.3). From these, the *stock mixture **A*** containing the BBs **1c** and **1i** was prepared by mixing 60 µL of the corresponding *individual stocks*. Then, two *pre-equilibrated reaction mixtures* were prepared by mixing separately 50 µL of the *stock mixture **A*** with 225 µL

of a 26.7 mM phosphate buffer (pH 7.5) and a 26.7 mM phosphate buffer (pH 2.5). The *individual stock* of **1d** was stored at –80 ºC. After 48 hours, the *reaction mixtures **A*** (2 mM of each BB) were prepared by adding 15 µL of the *individual stock* of **1d** to 165 µL of each of the *pre-equilibrated reaction mixture*s.



Scheme 3.3. Preparation of the solutions of the reversibility test.

Simultaneously, a stock mixture of the three BBs (*stock mixture **B***) was prepared by mixing 40 µL of the three *individual stocks*. From this, two *control reaction mixtures **B*** (2 mM of each BB) were prepared by mixing separately 50 µL with 150 µL of a 26.7 mM phosphate buffer (pH 7.5) and a 26.7 mM phosphate buffer (pH 2.5). After 48 hours, the *pre-equilibrated reaction mixtures*, the *reaction mixtures **A*** and the *control reaction mixtures **B*** were analyzed by HPLC (Figure 3.15).

At pH 7.5, after the addition of **1d** to the *pre-equilibrated mixture* of **1c**+**1i**, the mixture evolved to the same final situation as when the three BBs are left to oxidize together. Therefore, at slightly basic pH the mixture proved to reach the thermodynamic equilibrium. On the contrary, at pH 2.5 the oxidation rate is faster than the exchange reaction, and thus, **1d** is forced to react with itself and only small amounts of the

heterodimers containing **1d** were formed. Therefore, at pH 2.5 the mixture is not under equilibrium conditions.

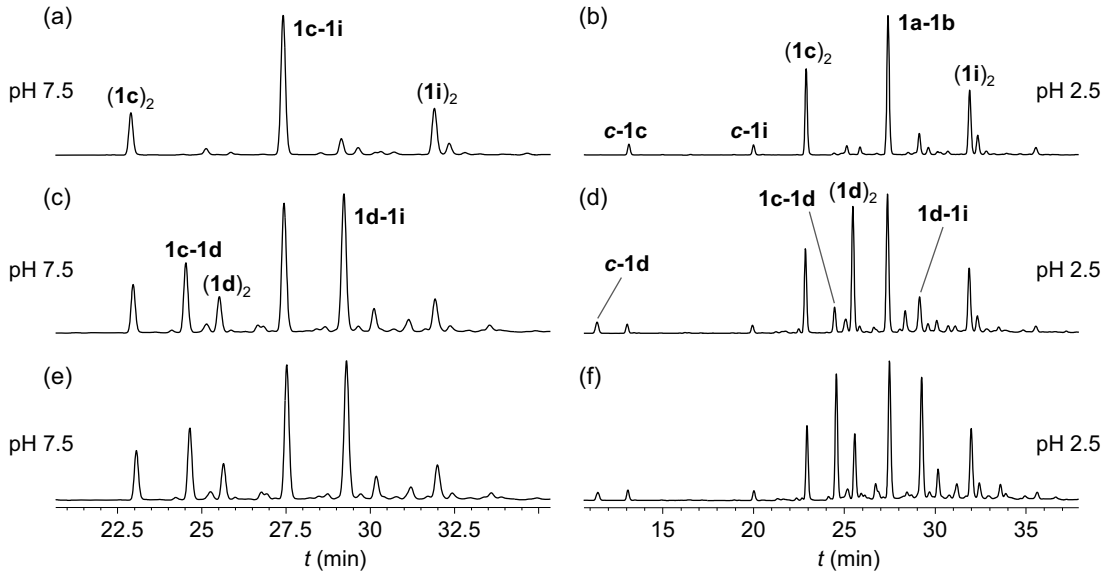


Figure 3.15. HPLC-UV traces (254 nm) of the *pre-equilibrated reaction mixtures* (a-b), the *reaction mixture **A*** (c-d) and the *control reaction mixture **B*** (e-f); performed at pH 7.5 (a,c,e) and pH 2.5 (b,d,f).

Additionally, 0.35 equivalents of tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl)[73] were added to both *reaction mixtures **A***. After the re-oxidation, their HPLC traces showed that the composition remained unchanged at pH 7.5 but not at pH 2.5, corroborating that the library reaches the equilibrium at pH 7.5 but not at pH 2.5 (Figure 3.16).

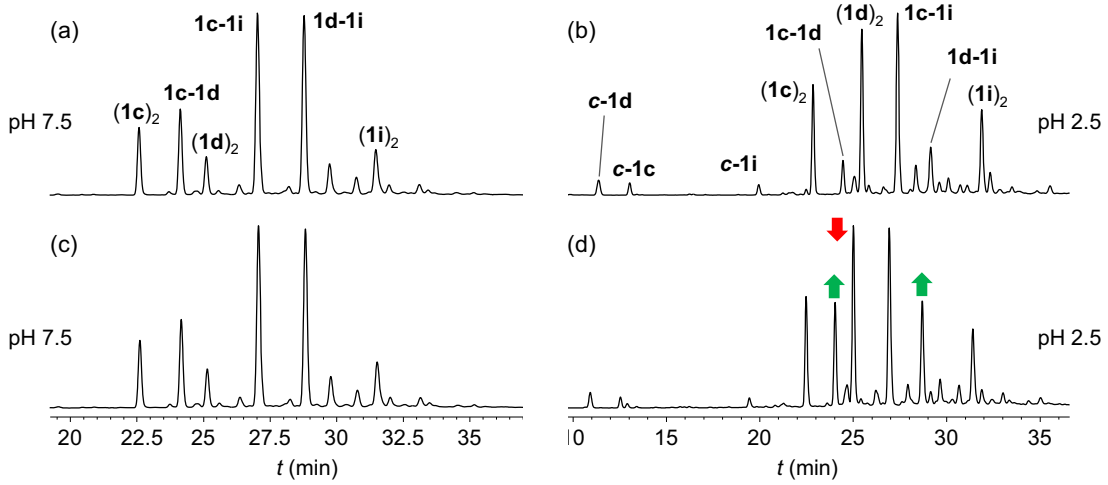


Figure 3.16. HPLC-UV traces (254 nm) of the *reaction mixture **A*** performed at pH 7.5 (a,c) and pH 2.5 (b,d); before (a-b) and after (c-d) the addition of 0.35 equivalents of TCEP·HCl.

3.5.4.2. Reversibility tests at pH 7.5 in the presence of salt

The three *individual stocks*, the *stock mixtures* ***A*** and ***B***, the *pre-equilibrated reaction mixture*, the *reaction mixture* ***A*** and the *control reaction mixture* ***B*** were prepared using the previously explained methodology shown in Scheme 3.3 (Figure 3.17a-c). However in this case the 26.7 mM phosphate buffer (pH 7.5) used to prepare the samples contained 1.33 M of NaCl.

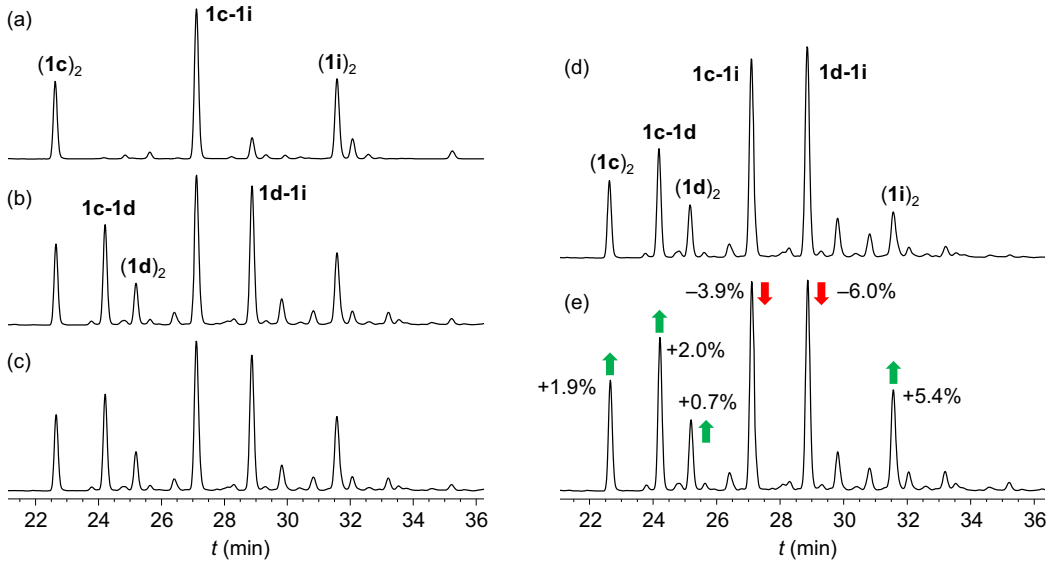


Figure 3.17. HPLC-UV traces (254 nm) of the *pre-equilibrated reaction mixtures* (a), the *reaction mixture* ***A*** (b) and the *control reaction mixture* ***B*** (c), performed at pH 7.5 in the presence of 1.0 M NaCl. HPLC-UV traces (254 nm) of the DCL generated from **1c**+**1d**+**1i** in 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO, before (d) and after (e) the addition of NaCl.

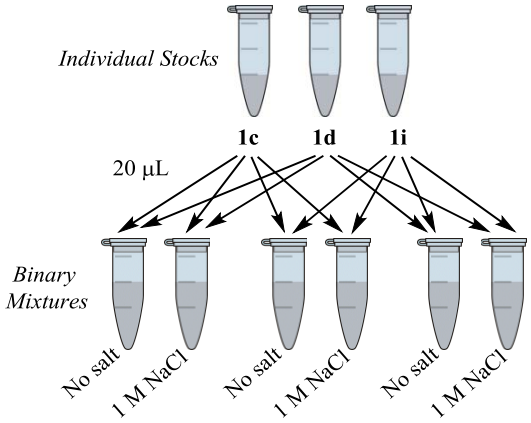
After the addition of **1d** to the *pre-equilibrated mixture* of **1c+1i**, the mixture evolved to the same final situation as when the three BBs are left to oxidize together. Therefore, at slightly basic pH and containing a high salt concentration, the mixture proved to reach the thermodynamic equilibrium.

Additionally, 45 µL of the *stock mixture* ***B*** were added to 100 µL of a 26.7 mM phosphate buffer (pH 7.5). After 48 hours, 44 µL of a 3.84 M NaCl solution and 5 µL of the *stock mixture* ***B*** were added to 120 µL of the previously oxidized mixture. The resulting solution was let to re-equilibrate and, after 24 hours, the corresponding HPLC analysis showed that the library evolved to the same final situation as when the salt was in the library from the beginning (Figure 3.17d-e).

3.5.5. Binary mixtures

3.5.5.1. Preparation of the binary mixtures

Individual stocks (16 mM) of each BB **1c,d,i** were prepared in DMSO (Scheme 3.4). From these, two *binary mixtures* (2 mM of each BB) were prepared by mixing 20 µL of each of the two corresponding *individual stocks* with 120 µL of a 26.7 mM aqueous phosphate buffer (pH 7.5) or a 26.7 mM aqueous phosphate buffer (pH 7.5) containing 1.33 M NaCl. After 24 hours, the six fully oxidized mixtures were analyzed by HPLC-UV (Figure 3.6 in section 3.3.4).



Scheme 3.4. Preparation of the binary mixtures.

3.5.5.2. Calculation of the stabilization energy for (**1i**)$_2$

For a binary mixture generated from BBs **A** and **B**, the formation constant (K_f) of the three corresponding dimers, *i.e.* **AA**, **BB** and **AB**, is defined as shown in Equations 3.3, 3.4 and 3.5. These three formation constants can be combined in a new expression where the exchange constant $K_{[\mathbf{A},\mathbf{B}]}$ (previously defined in Chapter 2, Equation 2.1) is introduced and the concentrations [**A**] and [**B**] are cancelled (Equation 3.6).

$$K_f(\mathbf{AA}) = [\mathbf{AA}]/[\mathbf{A}]^2 \qquad \textbf{(Equation 3.3)}$$

$$K_f(\mathbf{BB}) = [\mathbf{BB}]/[\mathbf{B}]^2 \qquad \textbf{(Equation 3.4)}$$

$$K_f(\mathbf{AB}) = [\mathbf{AB}]/([\mathbf{A}]\cdot[\mathbf{B}]) \qquad \textbf{(Equation 3.5)}$$

$$K_f(\mathbf{AA})\cdot K_f(\mathbf{BB}) = K_f^2(\mathbf{AB})\cdot K_{[\mathbf{A},\mathbf{B}]} \qquad \textbf{(Equation 3.6)}$$

For the binary mixture generated from **1c**+**1i**, the ratio between the formation constant of $(\mathbf{1i})_2$ in the presence an absence of salt was calculated as shown in Equation 3.7. This expression was obtained by isolating $K_f(\mathbf{AA})$ in Equation 3.6. The constant $K_f((\mathbf{1c})_2)$ is known to be insensitive to the salt concentration and the constant $K_f((\mathbf{1c\text{-}1i})$, for practical purposes, can also be considered not to depend on the concentration of salt.[m] Accordingly, the calculation of the ratio of formation constants is simplified and can be directly obtained by dividing the corresponding exchange constants[n] (Equation 3.7).The same calculation was also done for the binary mixture generated from **1d**+**1i**, obtaining a very similar result (Equation 3.8). Using 2.9 as the average of the values obtained from Equations 3.7 and 3.8, the salt-induced stabilization of $(\mathbf{1i})_2$ was calculated by means of Equation 3.9.

$$\frac{K_f^{\text{NaCl}}((\mathbf{1i})_2)}{K_f^{\text{No salt}}((\mathbf{1i})_2)} = \frac{\left(\dfrac{K_f^2(\mathbf{1c\text{-}1i})\cdot K_{[\mathbf{1c},\mathbf{1i}]}}{K_f((\mathbf{1c})_2)}\right)_{\text{NaCl}}}{\left(\dfrac{K_f^2(\mathbf{1c\text{-}1i})\cdot K_{[\mathbf{1c},\mathbf{1i}]}}{K_f((\mathbf{1c})_2)}\right)_{\text{No salt}}} \simeq \frac{K_{[\mathbf{1c},\mathbf{1i}]}^{\text{NaCl}}}{K_{[\mathbf{1c},\mathbf{1i}]}^{\text{No salt}}} = 2.8 \qquad \textbf{(Equation 3.7)}$$

$$\frac{K_f^{\text{NaCl}}((\mathbf{1i})_2)}{K_f^{\text{No salt}}((\mathbf{1i})_2)} = \ldots \simeq \frac{K_{[\mathbf{1d},\mathbf{1i}]}^{\text{NaCl}}}{K_{[\mathbf{1d},\mathbf{1i}]}^{\text{No salt}}} = 3.0 \qquad \textbf{(Equation 3.8)}$$

$$\Delta\Delta G^{\circ} = -RT\ln\left(\frac{K_f^{\text{NaCl}}((\mathbf{1i})_2)}{K_f^{\text{No salt}}((\mathbf{1i})_2)}\right) = -2.6\ \text{kJ}\cdot\text{mol}^{-1} \qquad \textbf{(Equation 3.9)}$$

3.5.6. Study of the $(\mathbf{1i})_3/(\mathbf{1i})_2$ proportion

A *stock solution* (8.54 mM) of **1i** was prepared by dissolving 2.24 mg in 510 µL of DMSO. From this, four *reaction mixtures* (2.13 mM) were prepared by mixing 50 µL with 150 µL of the following solutions: i) 26.7 mM aqueous phosphate buffer (pH 7.5), ii) 26.7 mM aqueous phosphate buffer (pH 7.5) with 0.67 M NaCl, iii) 26.7 mM aqueous phosphate buffer (pH 7.5) with 1.33 M NaCl, and iv) 26.7 mM aqueous phosphate buffer (pH 7.5) with 2.66 M NaCl. After 2 days, the fully oxidized *reaction mixtures* were analyzed by HPLC.

[m] In the ternary system it is equivalent to consider the stabilization of $(\mathbf{1i})_2$ or the destabilization of both **1c-1i** and **1d-1i.** The observed amplification of $(\mathbf{1i})_2$ is the result of an increase in its relative stability.

[n] The exchange constants calculated for the mixtures of **1c**+**1i** and **1d**+**1i** in the presence and absence of 1.0 M NaCl are: $K_{[\mathbf{1c},\mathbf{1i}]}^{\text{NaCl}} = 0.28$, $K_{[\mathbf{1c},\mathbf{1i}]}^{\text{No salt}} = 0.10$, $K_{[\mathbf{1d},\mathbf{1i}]}^{\text{NaCl}} = 0.20$ and $K_{[\mathbf{1d},\mathbf{1i}]}^{\text{No salt}} = 0.065$.

The molar absorptivity at 254 nm of BB **1i** was found to be the same whether as free dithiol or as part of any of the generated macrocyclic disulfides (in a similar experiment as the one explained in section 2.3.3). Therefore, considering that the initial concentration of **1i** was $2.13 \cdot 10^{-3}$ M, and the fact that the libraries are composed by only $(\mathbf{1i})_2$ and $(\mathbf{1i})_3$, the corresponding concentrations were obtained by means of the Equations 3.10 and 3.11, where $A_{(\mathbf{1i})_2}$ and $A_{(\mathbf{1i})_3}$ are the HPLC areas of the dimer and trimer of **1i** respectively, and $A_T = A_{(\mathbf{1i})_2} + A_{(\mathbf{1i})_3}$. These concentrations were used to calculate the $K_{T/D}$ by means of Equation 3.1.

$$[(\mathbf{1i})_2] = \left(A_{(\mathbf{1i})_2} \cdot \frac{2.13 \cdot 10^{-3}\ \mathrm{M}}{A_T} \right) \Big/ 2 \qquad \textbf{(Equation 3.10)}$$

$$[(\mathbf{1i})_3] = \left(A_{(\mathbf{1i})_3} \cdot \frac{2.13 \cdot 10^{-3}\ \mathrm{M}}{A_T} \right) \Big/ 3 \qquad \textbf{(Equation 3.11)}$$

3.5.7. Structural studies

3.5.7.1. Molecular modeling

All the theoretical calculations were performed with Spartan 06 software operating in a Dell workstation. Monte Carlo conformation searches were performed without restrictions by generating 10000-20000 geometries, which were minimized subsequently using the MMFFaq force field. This version of the force field takes into account water solvent as a continuum medium and has proved to be the most suitable for the conformational analysis of pseudopeptide and peptoid molecules.[61,74-76] The obtained local minima were ordered following the corresponding MMFFaq energies. The process was repeated several times staring from different initial geometries to ensure mapping all the conformational space. The corresponding conformational searches of the same molecule starting from different geometries rendered identical results. This fact ensures the fidelity and reliability of the results from this conformational analysis. The corresponding Boltzmann distribution and CPK areas were calculated using the same software.

3.5.7.2. Nuclear magnetic resonance

An 8.3 mM *stock solution* of **1i** was prepared by dissolving 3.22 mg in 750 μL of DMSO-d_6. From this, three NMR samples (2.1 mM each) were prepared by mixing 210

μL of the *stock solution* with 630 μL of the following solutions: (a) a 26.7 mM aqueous phosphate buffer (pH 7.5) with 1.33 M NaCl, (b) a 26.7 mM aqueous phosphate buffer (pH 7.5), and (c) a 26.7 mM aqueous phosphate buffer (pH 2.5). Similarly, a 7.1 mM *stock solution* of **1c** was prepared by dissolving 1.60 mg in 440 μL of DMSO-d_6. From this, two NMR samples (1.8 mM each) were prepared by mixing 210 μL of the *stock solution* with 630 μL of the following solutions: (d) a 26.7 mM phosphate buffer (pH 7.5), and (e) a 26.7 mM phosphate buffer (pH 2.5). Finally, 5 μL of a 20 mM solution of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS)[77] in H_2O were added to all the NMR samples as an internal reference. After 24 hours, the five fully oxidized NMR samples were analyzed by HPLC (Figure 3.18).

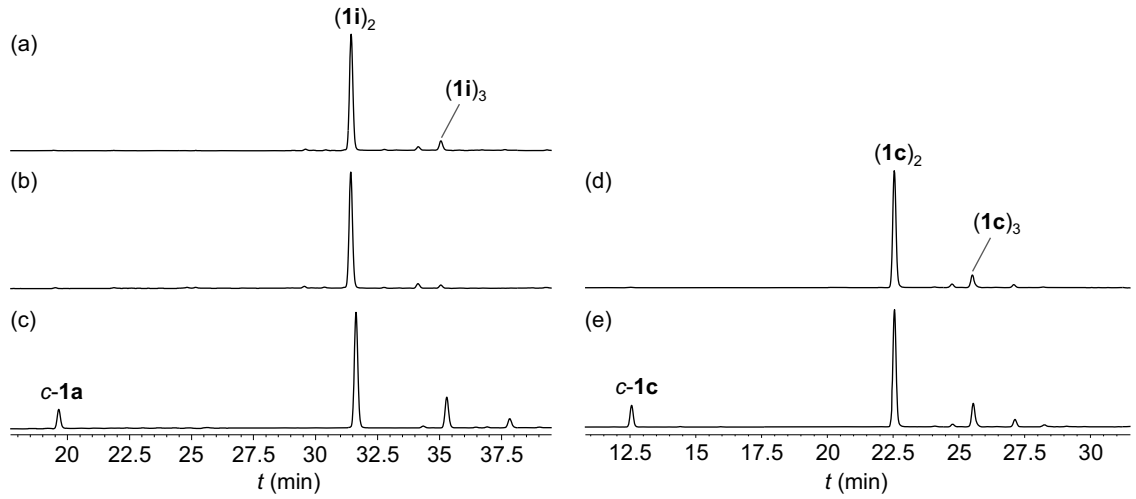


Figure 3.18. HPLC-UV traces (254 nm) of oxidized **1i** in: (a) 20 mM phosphate buffer (pH 7.5) with 1.0 M NaCl and 25% (v/v) DMSO-d_6, (b) 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6, and (c) 20 mM phosphate buffer (pH 2.5) with 25% (v/v) DMSO-d_6. HPLC-UV traces (254 nm) of oxidized **1c** in: (d) 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6, and (e) 20 mM phosphate buffer (pH 2.5) with 25% (v/v) DMSO-d_6.

The NMR samples were analyzed using the (H)WET (Water suppression Enhanced through T_1 effects) pulse sequence.[78] The methyl group of DSS was used as a reference at 0.00 ppm. Additionally, a (H)WET-COSY experiment was performed for the NMR samples of $(\mathbf{1i})_2$ and $(\mathbf{1c})_2$ at pH 7.5 and 2.5 (Figure 3.19) in order to unambiguously assign the proton signals.

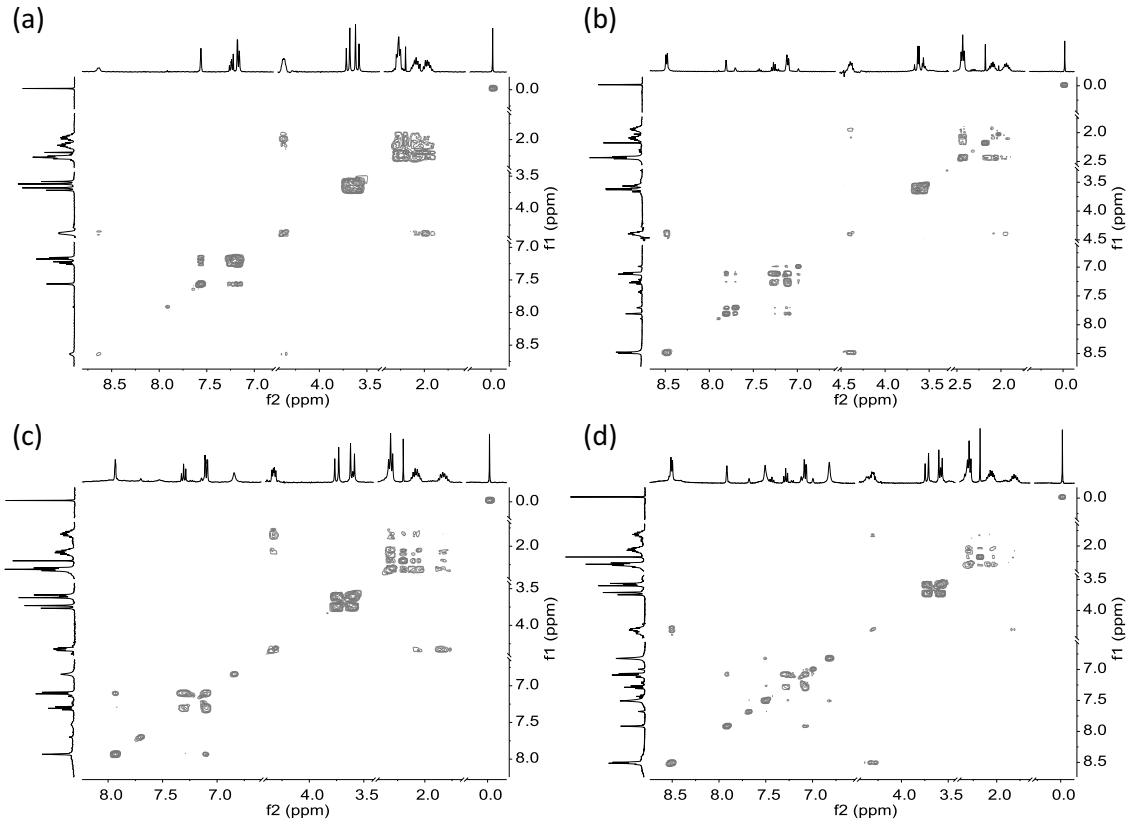


Figure 3.19. Partial (H)WET-COSY spectra of: (a) (**1i**)$_2$ in 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6, (b) (**1i**)$_2$ in 20 mM phosphate buffer (pH 2.5) with 25% (v/v) DMSO-d_6, (c) (**1c**)$_2$ in 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6, and (d) (**1c**)$_2$ in 20 mM phosphate buffer (pH 2.5) with 25% (v/v) DMSO-d_6.

3.5.7.3. Circular dichroism

BB **1i** (29.5 mg) was dissolved in a 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) MeOH. The oxidation process was monitored for 10 days, until the total consumption of dithiol **1i** was observed by HPLC. The solution was first evaporated under vacuum to remove the MeOH and then lyophilized to remove the H_2O. The resulting solid was purified using preparative reversed-phase chromatography (gradient: from 5% to 30% CH_3CN in H_2O) and 20.1 mg of pure (**1i**)$_2$ were obtained as a white solid (Figure 3.20).

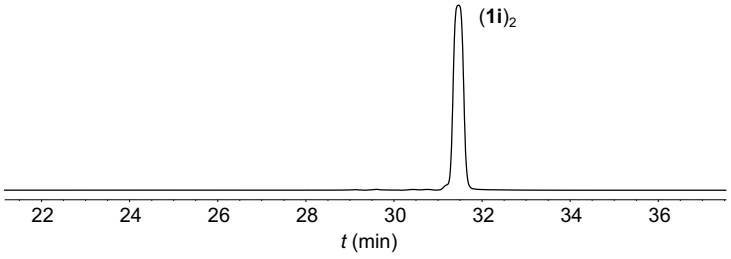


Figure 3.20. HPLC-UV trace (254 nm) of dimer (**1i**)$_2$ after purification.

A 1.07 mM solution of pure dimer $(\mathbf{1i})_2$ was prepared by dissolving 0.44 mg in 400 μL of a mixture of H_2O with 25% (v/v) MeOH. Aliquots of this solution (5 μL each) were consecutively added to the following blanks: i) 1.0 mL of a 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) MeOH, ii) 1.0 mL of a 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) MeOH containing 1.0 M NaCl, and iii) 1.0 mL of a 20 mM aqueous phosphate buffer (pH 2.5) with 25% (v/v) MeOH. The corresponding absorbance and CD signals were observed to be proportional to the concentration, at least up to $2.10 \cdot 10^{-5}$ M of $(\mathbf{1i})_2$. The spectra were analyzed with the Spectra Manager software (JASCO Corporation) v. 1.53.01 and a Means-Movement smoothing with a convolution width of 5 was applied. For the representation shown in Figure 3.13a (section 3.3.6.3), the molar absorption ($\Delta\varepsilon$, $cm^2 \cdot mmol^{-1}$) was calculated by means of Equation 3.12, where θ is the ellipticity (mdeg), C is the concentration (M) and l is the cell path length (cm). No changes were observed between the normalized spectra at different concentrations.

$$\Delta\varepsilon = \frac{\theta}{32980 \cdot C \cdot l} \qquad \textbf{(Equation 3.12)}$$

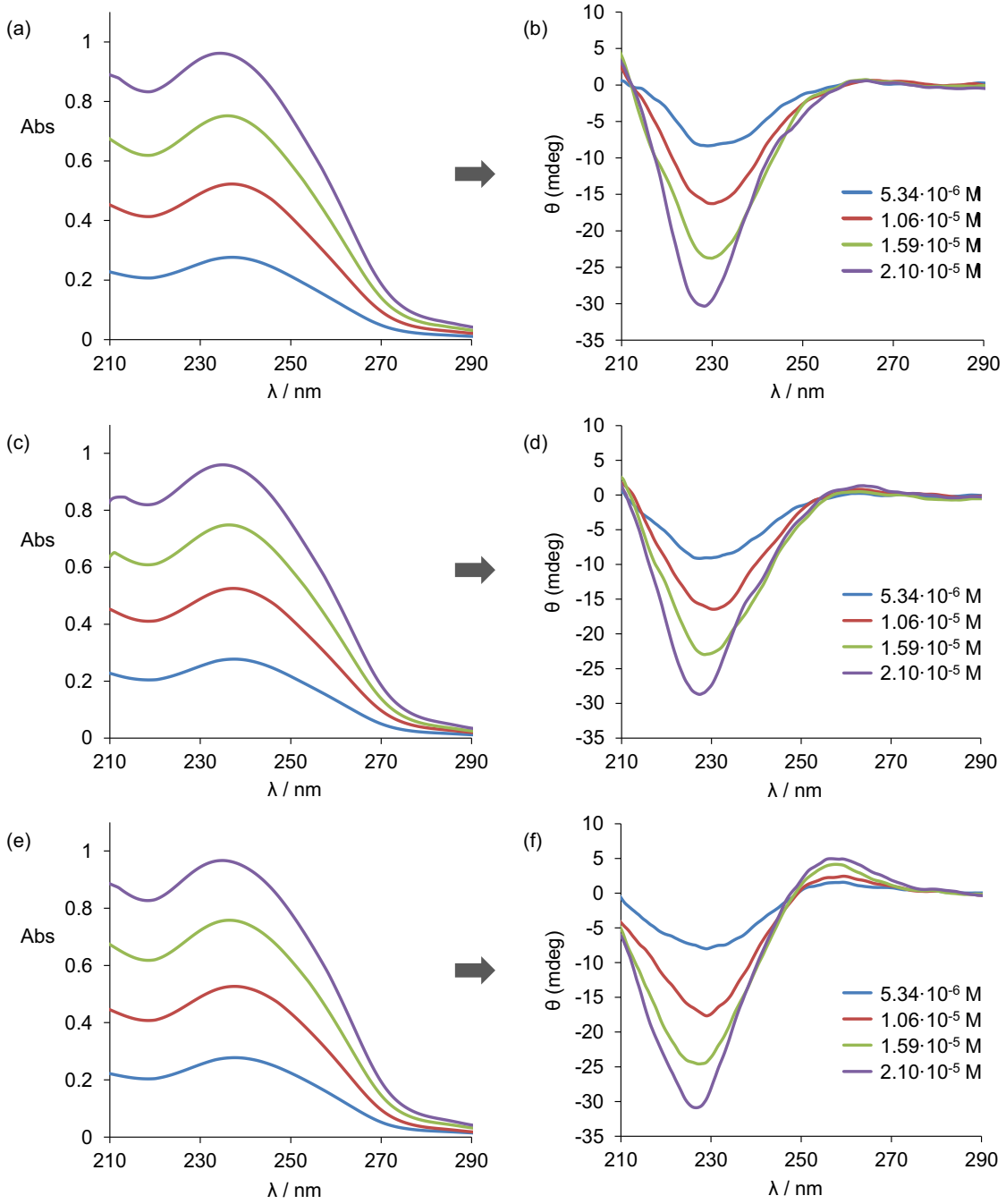


Figure 3.21. Absorbance (a,c,e) and circular dichroism (b,d,g) spectra of (**1i**)$_2$ in: (a-b) 20 mM phosphate buffer (pH 7.5) with 25% (v/v) MeOH, (c-d) 20 mM phosphate buffer (pH 7.5) with 25% (v/v) MeOH and 1.0 M NaCl, and (e-f) 20 mM phosphate buffer (pH 2.5) with 25% (v/v) MeOH.

3.6. References

(1) Bada, J. L. *Chem. Soc. Rev.* **2013**, *42*, 2186.
(2) Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. *Chem. Rev.* **2013**, *114*, 285.
(3) Szostak, J. W. *Nature* **2009**, *459*, 171.
(4) Pross, A. *Origins Life Evol. B.* **2004**, *34*, 307.
(5) Rebek, J. *J. Mol. Recogn.* **1992**, *5*, 83.
(6) Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S. *Science* **2010**, *327*, 1502.
(7) Sadownik, J. W.; Philp, D. *Angew. Chem. Int. Ed.* **2008**, *47*, 9965.
(8) Nguyen, R.; Allouche, L.; Buhler, E.; Giuseppone, N. *Angew. Chem. Int. Ed.* **2009**, *48*, 1093.
(9) Maeda, Y.; Javid, N.; Duncan, K.; Birchall, L.; Gibson, K. F.; Cannon, D.; Kanetsuki, Y.; Knapp, C.; Tuttle, T.; Ulijn, R. V.; Matsui, H. *J. Am. Chem. Soc.* **2014**, *136*, 15893.
(10) Williams, R. J.; Smith, A. M.; Collins, R.; Hodson, N.; Das, A. K.; Ulijn, R. V. *Nat. Nanotechnol.* **2009**, *4*, 19.
(11) Otto, S. *Acc. Chem. Res.* **2012**, *45*, 2200.
(12) Malakoutikhah, M.; Peyralans, J. J. P.; Colomb-Delsuc, M.; Fanlo-Virgós, H.; Stuart, M. C. A.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 18406.
(13) Leonetti, G.; Otto, S. *J. Am. Chem. Soc.* **2015**, *137*, 2067.
(14) Ashkenasy, G.; Jagasia, R.; Yadav, M.; Ghadiri, M. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10872.
(15) Saghatelian, A.; Yokobayashi, Y.; Soltani, K.; Ghadiri, M. R. *Nature* **2001**, *409*, 797.
(16) Kennan, A. J.; Haridas, V.; Severin, K.; Lee, D. H.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2001**, *123*, 1797.
(17) Severin, K.; Lee, D. H.; Kennan, A. J.; Ghadiri, M. R. *Nature* **1997**, *389*, 706.
(18) Lee, D. H.; Severin, K.; Yokobayashi, Y.; Ghadiri, M. R. *Nature* **1997**, *390*, 591.
(19) Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. *Nature* **1996**, *382*, 525.
(20) Dadon, Z.; Wagner, N.; Alasibi, S.; Samiappan, M.; Mukherjee, R.; Ashkenasy, G. *Chem. Eur. J.* **2015**, *21*, 648.
(21) Eisenberg, M.; Shumacher, I.; Cohen-Luria, R.; Ashkenasy, G. *Biorg. Med. Chem.* **2013**, *21*, 3450.
(22) Dadon, Z.; Samiappan, M.; Shahar, A.; Zarivach, R.; Ashkenasy, G. *Angew. Chem. Int. Ed.* **2013**, *52*, 9944.
(23) Dadon, Z.; Samiappan, M.; Wagner, N.; Ashkenasy, G. *Chem. Commun.* **2012**, *48*, 1419.
(24) Schliwa, M. *Molecular motors*; Springer, 2006.
(25) Qu, D.-H.; Tian, H. *Chem. Sci.* **2013**, *4*, 3031.
(26) von Delius, M.; Leigh, D. A. *Chem. Soc. Rev.* **2011**, *40*, 3656.
(27) von Delius, M.; Geertsema, E. M.; Leigh, D. A.; Tang, D.-T. D. *J. Am. Chem. Soc.* **2010**, *132*, 16134.
(28) von Delius, M.; Geertsema, E. M.; Leigh, D. A. *Nat. Chem.* **2010**, *2*, 96.
(29) Barrell, M. J.; Campaña, A. G.; von Delius, M.; Geertsema, E. M.; Leigh, D. A. *Angew. Chem. Int. Ed.* **2011**, *50*, 285.
(30) Campaña, A. G.; Carlone, A.; Chen, K.; Dryden, D. T. F.; Leigh, D. A.; Lewandowska, U.; Mullen, K. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 5480.
(31) Kovaříček, P.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 9446.

(32) Valentine, D. L. *Nat. Rev. Microbiol.* **2007**, *5*, 316.
(33) Lanyi, J. K. *Bacteriol. Rev.* **1974**, *38*, 272.
(34) Pieper, U.; Kapadia, G.; Mevarech, M.; Herzberg, O. *Structure* **1998**, *6*, 75.
(35) Paul, S.; Bag, S.; Das, S.; Harvill, E.; Dutta, C. *Genome Biol.* **2008**, *9*, 1.
(36) Madern, D.; Ebel, C.; Zaccai, G. *Extremophiles* **2000**, *4*, 91.
(37) Eisenberg, H. *Arch. Biochem. Biophys.* **1995**, *318*, 1.
(38) Galinski, E. A.; Trüper, H. G. *FEMS Microbiol. Rev.* **1994**, *15*, 95.
(39) Eisenberg, H.; Mevarech, M.; Zaccai, G. *Adv. Protein Chem.* **1992**, *43*, 1.
(40) Dalhus, B.; Saarinen, M.; Sauer, U. H.; Eklund, P.; Johansson, K.; Karlsson, A.; Ramaswamy, S.; Bjørk, A.; Synstad, B.; Naterstad, K.; Sirevåg, R.; Eklund, H. *J. Mol. Biol.* **2002**, *318*, 707.
(41) Irimia, A.; Ebel, C.; Madern, D.; Richard, S. B.; Cosenza, L. W.; Zaccaï, G.; Vellieux, F. M. D. *J. Mol. Biol.* **2003**, *326*, 859.
(42) Mevarech, M.; Frolow, F.; Gloss, L. M. *Biophys. Chem.* **2000**, *86*, 155.
(43) Fukuchi, S.; Yoshimune, K.; Wakayama, M.; Moriguchi, M.; Nishikawa, K. *J. Mol. Biol.* **2003**, *327*, 347.
(44) Tadeo, X.; López-Méndez, B.; Trigueros, T.; Laín, A.; Castaño, D.; Millet, O. *PLoS Biol.* **2009**, *7*, e1000257.
(45) Premkumar, L.; Greenblatt, H. M.; Bageshwar, U. K.; Savchenko, T.; Gokhman, I.; Sussman, J. L.; Zamir, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7493.
(46) Bieger, B.; Essen, L.-O.; Oesterhelt, D. *Structure* **2003**, *11*, 375.
(47) Richard, S. B.; Madern, D.; Garcin, E.; Zaccai, G. *Biochemistry* **2000**, *39*, 992.
(48) Dym, O.; Mevarech, M.; Sussman, J. *Science* **1995**, 1344.
(49) Kastritis, P. L.; Papandreou, N. C.; Hamodrakas, S. J. *Int. J. Biol. Macromol.* **2007**, *41*, 447.
(50) Britton, K. L.; Baker, P. J.; Fisher, M.; Ruzheinikov, S.; Gilmour, D. J.; Bonete, M.-J.; Ferrer, J.; Pire, C.; Esclapez, J.; Rice, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4846
(51) Ebel, C.; Costenaro, L.; Pascu, M.; Faou, P.; Kernel, B.; Proust-De Martin, F.; Zaccai, G. *Biochemistry* **2002**, *41*, 13234.
(52) Zaccai, G.; Cendrin, F.; Haik, Y.; Borochov, N.; Eisenberg, H. *J. Mol. Biol.* **1989**, *208*, 491.
(53) Qvist, J.; Ortega, G.; Tadeo, X.; Millet, O.; Halle, B. *J. Phys. Chem. B* **2012**, *116*, 3436.
(54) Lehn, J.-M. *Chem. Soc. Rev.* **2007**, *36*, 151.
(55) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Chem. Eur. J.* **2008**, *14*, 2153.
(56) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Angew. Chem. Int. Ed.* **2007**, *46*, 8858.
(57) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *J. Am. Chem. Soc.* **2005**, *127*, 9390.
(58) Corbett, P. T.; Otto, S.; Sanders, J. K. M. *Chem. Eur. J.* **2004**, *10*, 3139.
(59) Ramström, O.; Lohmann, S.; Bunyapaiboonsri, T.; Lehn, J.-M. *Chem. Eur. J.* **2004**, *10*, 1711.
(60) Bunyapaiboonsri, T.; Ramström, O.; Lohmann, S.; Lehn, J.-M.; Peng, L.; Goeldner, M. *ChemBioChem* **2001**, *2*, 438.
(61) Brandt, W.; Herberg, T.; Wessjohann, L. *Peptide Science* **2011**, *96*, 651.
(62) Haridas, V.; Singh, H.; Sharma, Y.; Lal, K. *J. Chem. Sci.* **2007**, *119*, 219.
(63) Ranganathan, D.; Haridas, V.; Nagaraj, R.; Karle, I. L. *J. Org. Chem.* **2000**, *65*, 4415.
(64) Karle, I. L.; Ranganathan, D.; Haridas, V. *J. Am. Chem. Soc.* **1996**, *118*, 10916.
(65) Mascal, M.; Moody, C. J.; Morrell, A. I.; Slawin, A. M. Z.; Williams, D. J. *J. Am. Chem. Soc.* **1993**, *115*, 813.

(66) Greenfield, N. J. *TrAC, Trends Anal. Chem.* **1999**, *18*, 236.
(67) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Bio.* **1991**, *20*, 519.
(68) Alfonso, I.; Bru, M.; Burguete, M. I.; García-Verdugo, E.; Luis, S. V. *Chem. Eur. J.* **2010**, *16*, 1246.
(69) Alfonso, I.; Bolte, M.; Bru, M.; Burguete, M. I.; Luis, S. V.; Rubio, J. *J. Am. Chem. Soc.* **2008**, *130*, 6137.
(70) Bru, M.; Alfonso, I.; Burguete, M. I.; Luis, S. V. *Angew. Chem. Int. Ed.* **2006**, *45*, 6155.
(71) Kelly, S. M.; Price, N. C. *Curr. Pretein Pept. Sc.* **2000**, *1*, 349.
(72) Greenfield, N. J. *Nat. Protoc.* **2007**, *1*, 2527.
(73) Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M. *J. Org. Chem.* **1991**, *56*, 2648.
(74) Rodriquez, C. F.; Orlova, G.; Guo, Y.; Li, X.; Siu, C.-K.; Hopkinson, A. C.; Siu, K. W. M. *J. Phys. Chem. B* **2006**, *110*, 7528.
(75) Strittmatter, E. F.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 6069.
(76) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908.
(77) Tiers, G.; Coon, R. *J. Org. Chem.* **1961**, *26*, 2097.
(78) Ogg, R. J.; Kingsley, R. B.; Taylor, J. S. *J. Magn. Reson., Ser B* **1994**, *104*, 1.

CHAPTER 4

Salt-induced adaptation of a dynamic combinatorial library of pseudopeptidic macrocycles: unraveling the electrostatic effects in mixed aqueous media

The studies comprehended in this Chapter resulted in the following publication:

Atcher, J.; Moure, A.; Bujons, J.; Alfonso, I. *Chem. Eur. J.* **2015**, *21*, 6869 (selected as a hot paper).

4.1. Introduction

4.1.1. The ionic strength as an external stimulus

While the addition of a template is probably the most common way to perturb the product distribution of a dynamic library, literature also contains many examples of other chemical and physical external stimuli used for this purpose. These examples include changing pH,[1-5] temperature,[2,6-9] light,[10-13] gelation,[14-18] mechanical forces[19] and presence of an electric field.[20-21] In addition, in relatively few examples the increase of the ionic strength resulting from the addition of an inorganic salt has also shown a remarkable influence on the composition of DCLs generated in aqueous media.[7-8,22-31] The following lines briefly comment on the main reported examples of the use of this external stimulus.

The first use of the ionic strength as an external stimulus was reported for the preparation of mechanically interlocked molecular architectures in water. Sanders and co-workers have investigated the use of aromatic and hydrophobic interactions between library components made from naphthalene diimide (NDI) acceptor and dialkoxynaphthalene (DN) donor building blocks (*e.g.* Figure 4.1). Very interestingly, the study of these libraries, generated by disulfide exchange, led to the discovery of donor-acceptor [2]-catenanes,[8,27-32] [3]-catenanes[25-26] and even giant species (Figure 4.1),[25] showing the utility of DCC for preparing topologically complex molecules.

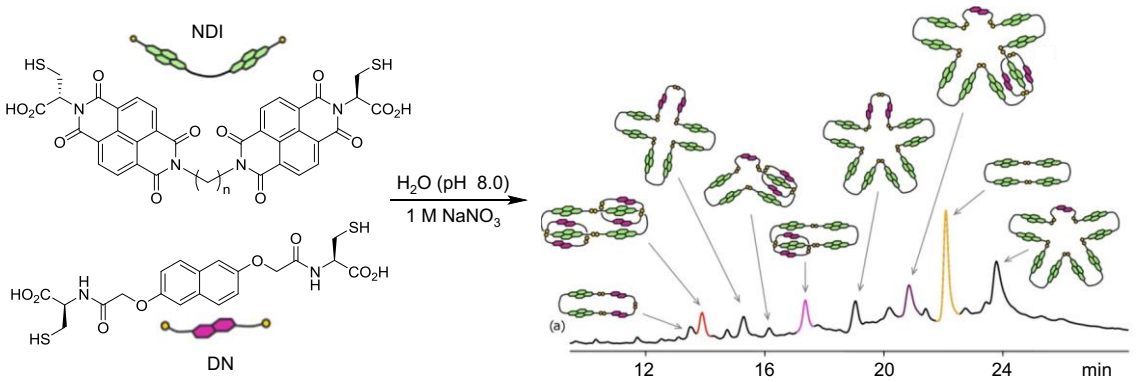


Figure 4.1. HPLC trace (383 nm) of a DCL generated from the mixture of a naphthalene diimide (NDI) acceptor and a dialkoxynaphthalene (DN) donor building blocks (2.5 mM each) in aqueous medium (pH 8.0) in the presence of 1 M $NaNO_3$ (Figure from reference[25]).

To achieve such impressive results, the authors had to control various structural parameters of the building blocks (linker length, flexibility, chirality) as well as some external factors. Among these external factors, the ionic strength was found to be a

crucial parameter. The strength of the hydrophobic interaction is known to be enhanced by increasing the concentration of the electrolytes, *i.e.* increasing the ionic strength.[33] Taking this into account, they realized that some of the library members, especially those with interlocked topologies, are very compact in solution, and thus have a reduced total area of exposed hydrophobic surfaces. Accordingly, these species were stabilized and amplified by increasing the ionic strength (*e.g.* addition of 1 M $NaNO_3$ in Figure 4.1). Overall, this example illustrates that controlling hydrophobic effects through solvent ionic strength is an effective approach for perturbing the product distribution of a DCL. Moreover, the addition of an inorganic salt proved to be particularly useful for the preparation of mechanically interlocked structures in aqueous media.

Another interesting example in which the addition of an inorganic salt is used as an external stimulus was reported by Hafezi and Lehn.[7] Their purpose was to investigate for the first time the adaptation of a dynamic library to phase separation. The separation of an organo-aqueous mixture into two separate phases leads to the formation of two distinct solvents with differentiated physicochemical properties, such as polarity and viscosity. The authors reasoned that, as a consequence of a phase separation, the composition of a DCL should be affected, leading to the amplification of the fittest compounds in each of the two generated environments. For demonstrating that, they used a mixture of four components: a hydrophilic aldehyde and amine, and a hydrophobic aldehyde and amine (Figure 4.2a). Dissolution of these four species in a single phase water/acetonitrile[a] mixture (2:3, v/v) generated a dynamic library consisting of the four possible imines, possessing either hydrophilic-hydrophilic, hydrophobic-hydrophobic or amphiphilic characteristics (Figure 4.2a). These four imines were present in almost statistical distribution.

Phase separation was accomplished by the addition of an inorganic salt (among other external stimuli).[b] Interestingly, the library distribution reorganized as follows: the hydrophobic-hydrophobic and hydrophilic-hydrophilic imines were strongly distributed into the organic and aqueous phases respectively, whereas the formation of both amphiphilic imines was disfavored (Figure 4.2b). Phase reunification was achieved by

[a] A detailed explanation of the properties of water/acetonitrile mixtures is given in the next Chapter (section 5.3.5.2).

[b] Other external factors investigated within the same study include either a physical stimulus (temperature) or the addition of chemical effectors other than inorganic salts (a hydrophilic organic molecule or a water immiscible organic liquid).

either warming the solution to 70 °C or precipitation,[c] leading to the reformation of the amphiphilic imines (Figure 4.2b).

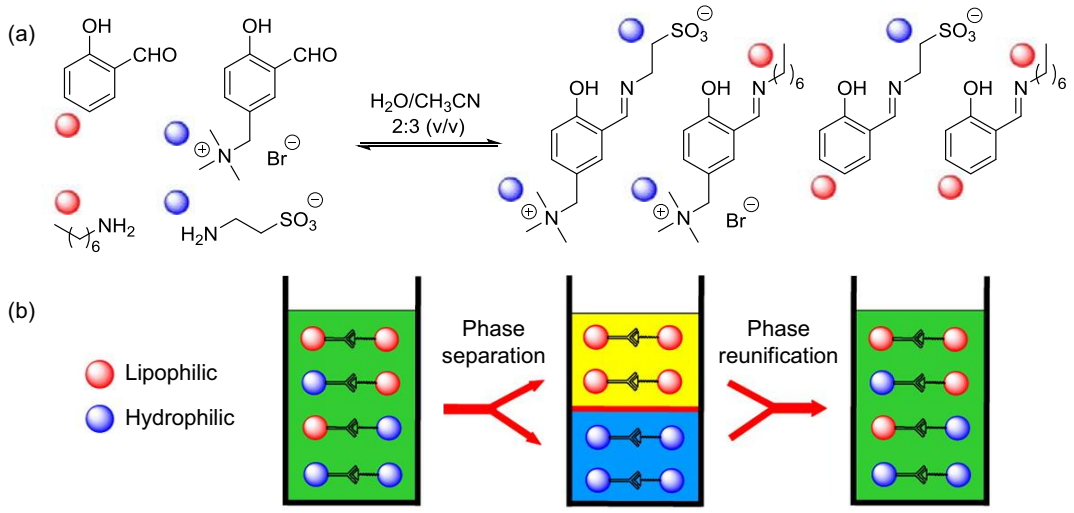


Figure 4.2. (a) Scheme of the imine-based DCL generated from a hydrophilic aldehyde and amine, and a hydrophobic aldehyde and amine. (b) Schematic representation of the component redistribution occurring in a DCL on phase separation and recombination of a binary organo-aqueous solvent mixture. The single homogeneous phase, the organic phase and the aqueous phase are represented in green, yellow and blue respectively (Figure modified from reference[7]).

In summary, the authors demonstrated that the addition of an inorganic salt can be used to induce the reorganization of a DCL distribution by liquid-liquid phase separation of a binary solvent mixture. The process is reversible as the biphasic mixture can be reunited.

The two examples commented above illustrate the main purposes for which high concentrations of an inorganic salt have been used so far in the field of DCC: enhancement of hydrophobic effects and phase separation of organo-aqueous mixtures. To the best of our knowledge, the studies comprehended in this Chapter, as well as in the previous Chapter 3, are the first examples of the use of this external factor as a bio-inspired stimulus.

4.1.2. Precedents of highly complex molecular networks in DCC

As previously mentioned in the General Introduction, complexity within the frame of DCC is the result from a combination of different interrelated features. Therefore, to evaluate the complexity level of a DCL is not an easy task. Taking this consideration into account, the modest aim of this section is to comment some examples of dynamic

[c] Precipitation was only used for the recombination of the KF-induced biphase. Treatment of a KF solution with solid $LiClO_4$ results in rapid precipitation of both LiF and $KClO_4$.

libraries with a highly complex composition. In particular, the following examples are presented to illustrate: i) DCLs consisting of a large number of library members, and ii) DCLs consisting of topologically complex molecules.

4.1.2.1. Large DCLs

Most of the DCLs found in the literature are limited to relatively small libraries, often made from only one or two building blocks forming less than ten library members. However, in few examples[34-37] large libraries have been successfully used for the discovery of several molecular receptors. Thus, Miller and co-workers reported the use of the resin-bound dynamic combinatorial chemistry method[38] to generate and screen a disulfide-based DCL with a theoretical diversity of more than 11000 library members.[35] This impressively complex system allowed identifying a selective and high-affinity ligand for an RNA fragment of HIV-1.

Another remarkable example was reported by Ludlow and Otto.[34] They reasoned that modern analytical techniques should be able to cope with large libraries in solution phase, and they prepared a large disulfide-based[d] DCL by mixing one tripodal building block and seven bipodal building blocks (Figure 4.3). The combination of BBs with different number of reacting sites is an additional source of chemical diversity that has scarcely been explored.[34,39-42] Considering the formation of cyclic oligomers up to tetramer, the authors estimated the presence of at least 9000 unique compounds. Then, they screened this huge library for binding to ephedrine (Figure 4.3). Using a combination of LC-MS and LC-MS-MS techniques, they identified different isomers of tetramers $\mathbf{AB}_3$ and $\mathbf{A}_2\mathbf{B}_2$ (incorporating BBs **A** and **B**, Figure 4.3) as the best binders to the template. Notably, these compounds were only detectable in the presence of the guest molecule. This demonstrates that it is not necessary for the library members to exist in detectable amounts in the untemplated library as long as their amplification is efficient enough when the template is present.

[d] Most, if not all, of the literature examples reporting very large DCLs are generated by means of the disulfide reversible chemistry. As mentioned in the general introduction, the symmetry of the disulfide bond allows generating large diversity.

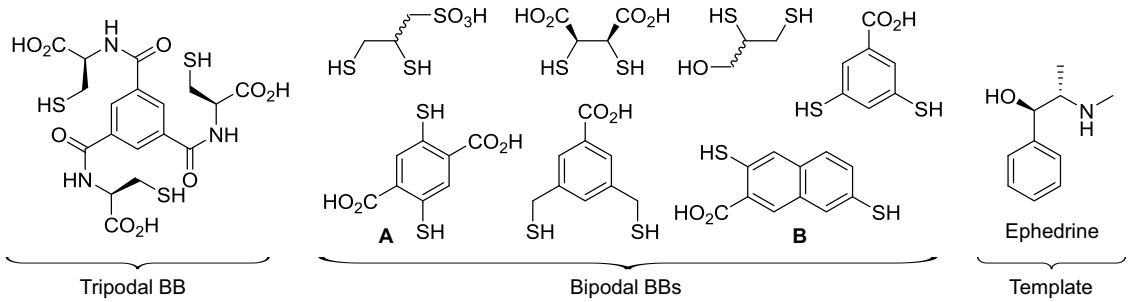


Figure 4.3. Structure of the eight building blocks used to generate a DCL of estimated 9000 components, and structure of the template molecule used to induce the amplification of different isomers of the tetramers $\mathbf{AB}_3$ and $\mathbf{A}_2\mathbf{B}_2$ (Figure modified from reference[34]).

On the other hand, structural diversity can also be generated without the need of mixing many different building blocks. As an example, Jurczak and co-workers recently described the use of a single building block to prepare a DCL consisting of numerous neutral macrocyclic disulfides (Figure 4.4).[43] Under kinetic control, the authors achieved to modify the composition of the library by changing the mode of adding the reagents. Thus, adding I_2 slowly to a concentrated solution of the starting dithiol (50 mM), facilitates the formation of larger oligomers and increases the structural diversity of the resulting mixture. Within this study, they also generated the same library under thermodynamic control and described the template effect of different organic anions.

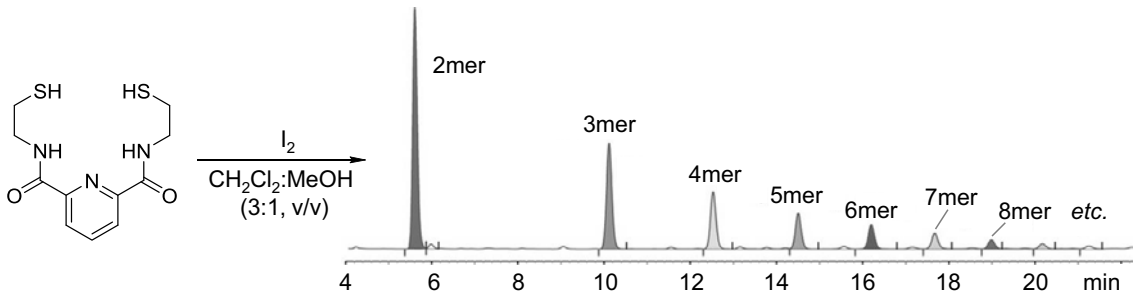


Figure 4.4. Structure of the neutral dithiol building block and HPLC-UV trace of the DCL of macrocyclic disulfides generated upon its kinetically controlled oxidation (Figure modified from reference[43]).

In theory, the probability of generating strong binders increases with library size, yet so does the probability that the concentration of individual library members drops below detection limits. Considering these factors, Ludlow and Otto studied the effect of library size on the probability of finding good molecular receptors, concluding that larger DCLs are likely to produce better binders.[44]

4.1.2.2. Topologically complex DCLs

In those DCLs where the template molecule is also part of the library, self-templating can be successfully exploited to create mechanical bonds and complex topologies. For this purpose, the reversibility of the reactions is crucial for achieving high selectivity, since it provides an error-correction mechanism that allows for the conversion of the misassembled kinetic products to the thermodynamic ones. The efficiency of this principle has been employed for the preparation of remarkably complex molecular topologies, including molecular walkers (section 3.1.1.2), catenanes (section 4.1.1), rotaxanes,[45-46] Borromean rings,[47] Solomon links,[48] pentafoil knots[49] and daisy chains.[50-51]

As an astonishing example, Sanders and co-workers have reported the discovery of an impressive trefoil knot in water, involving only purely organic building blocks.[23-24] The synthesis of molecular knots is particularly difficult since it is entropically much more demanding than topologically simpler macrocyclization or catenation processes. The authors used a single building block composed of three naphthalenediimide π-systems connected by flexible hydrophilic β-amino acids (Figure 4.5). A cysteine moiety attached at both ends provided the building block with two thiol groups for disulfide exchange. Very interestingly, dissolution of this dithiol in water (pH 8.0) in the presence of 1 M $NaNO_3$[e] rendered a DCL consisting almost exclusively of a trimeric species with a trefoil knot structure (Figure 4.5). Since this knot is topologically chiral, two different diasteromers could be obtained. However, only the right-handed form was observed, which implies that the synthesis of the knot is stereoselective. The authors reasoned that its formation is probably driven by hydrophobic effects, being the knot the smallest structure that minimizes the solvent-exposed hydrophobic surface.

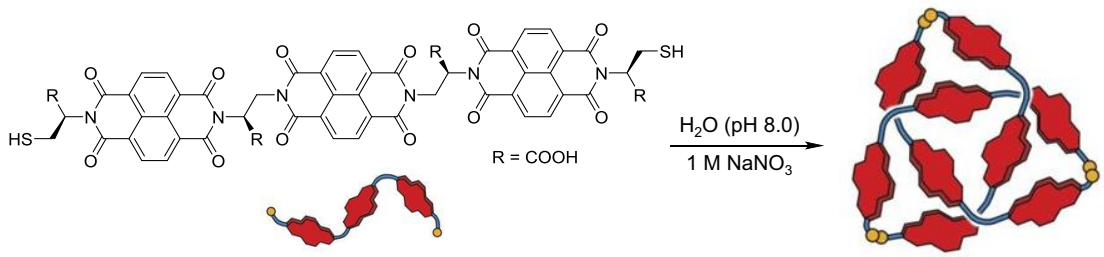


Figure 4.5. Structure of the dithiol building block composed of three NDI units connected by flexible amino acids, and cartoon representation of the right-handed trimeric knot generated upon its oxidation (Figure modified from reference[24]).

[e] A high concentration of salt was added in order to enhance the hydrophobic effects (section 4.1.1).

4.2. Objectives and hypothesis

The main objective of the present Chapter is to study the effect of increasing the ionic strength on the composition of a complex dynamic library containing differently charged macrocyclic species. With this purpose, we hypothesized that the top-down deciphering of the norms governing the entire network would allow the full understanding of the whole process.

This main objective can be divided into the following three specific aims:

i) To design and prepare a highly diverse DCL of macrocyclic pseudopeptides containing negative, neutral and positive charges.

ii) To study the adaptation of this complex library to the increase of the ionic strength. With this aim, deconvolution experiments and structural studies would allow characterizing the co-adaptive relationships present between the members of the complex library.

iii) To rationally design dynamic systems with bio-inspired tailored co-adaptive relationships.

4.3. Results and discussion

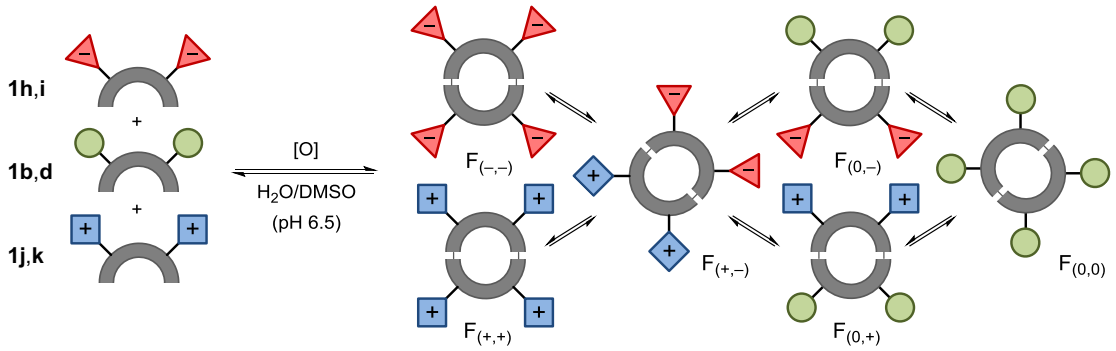
4.3.1. Motivation for the present study

In the previous Chapter of this thesis, we realized that the increase of the salt concentration in a DCL of pseudopeptidic macrocycles induced the amplification of the members concentrating acidic side chains (anionic at the pH of the experiments). This adaptive trend has a remarkable resemblance with the natural evolution of the halophilic proteins. Thus, we realized that our dynamic libraries comprised a good benchmark simple model to study these types of adaptive processes. However, a complete physicochemical explanation for the observed adaptive changes was missing.

In the present study we aimed to fully understand the salt-adaptation processes of our DCLs. With this purpose, we paradoxically envisioned a much more complex library containing differently charged species. Increasing complexity leads to the emergence of higher level features.[52] Thus, the adaptive trends of a more complex system contain more information regarding the effect of the corresponding external stimulus. The top-down deciphering of the norms governing the entire network by a minimalistic deconvolution was foreseen as the strategy to reach a satisfactory explanation for the salt-induced adaptation process. The deep analysis of the network behavior should allow us to unravel the chemical information stored in the members of the library that is expressed in the dynamic process, within a complex network of similar congeners and upon the presence of an external stimulus.

4.3.2. Design and composition of a large and diverse DCL

We designed a highly diverse DCL combining the 6 dithiols **1b**+**1d**+**1h**+**1i**+**1j**+**1k**. At pH 6.5, the carboxylates of **1h-i**, the polar groups of **1b**,**d** and the ammonium groups of **1j-k**, confer negative, neutral and positive charges respectively to the corresponding side chains. In addition, we implemented different lengths of the side chains for the BBs based on charged amino acids: Asp (**1h**) *vs.* Glu (**1i**) and Orn (**1j**) *vs.* Lys (**1k**). We retained the initial bio-inspiration of Chapter 3 in the fact that the chemical structures used herein also contain peptide-like information and the selected stimulus, *i.e.* the increase of the salt concentration (section 4.3.3), is present in Nature.



Scheme 4.1. Representation of the complex dynamic molecular network generated by the mixture of **1b**+**1d**+**1h**+**1i**+**1j**+**1k**.

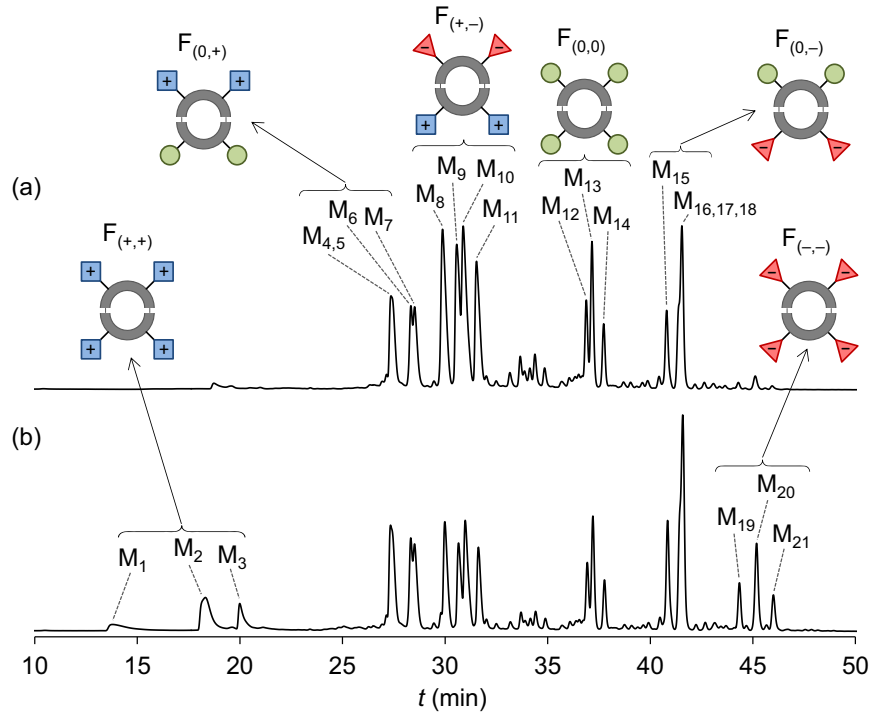
The 6 BBs were mixed at 0.5 mM each in 40 mM aqueous bis-Tris[f] buffer (pH 6.5) with 25% (v/v) DMSO (Scheme 4.1). These experimental conditions were carefully chosen with the following criteria. Firstly, the BBs were mixed in quite low concentration in order to favor the formation of dimers over trimers, as well as to prevent precipitation. The libraries were prepared by means of the Ellman's test (section 4.5.3), ensuring the exact 0.5 mM concentration of the 6 dithiols. Secondly, the bis-Tris buffer was used because its maximum buffering capacity is at pH 6.5, the appropriate slightly acidic conditions to ensure full ionization of the charged BBs (section 4.3.5.1). In this case, the control of the pH is of special importance because the suitable pH range is relatively narrow: either too acidic or too alkaline media would partially protonate the carboxylates of **1h-i** or deprotonate the ammonium groups of **1j-k**, respectively. Additionally, the pH of an aqueous bis-Tris buffer, in contrast with the McIlvine buffer (section 2.3.2), is expected not to significantly change upon the addition of DMSO,[53] making the control of the pH of the libraries easier.

Once the system generated from the 6 BBs reached the equilibrium composition,[g] the mixture was subsequently analyzed by HPLC-UV and UPLC-MS. The library was found to be dominated by 15 of the 21 possible dimers[h] and only small amounts of trimers[h] were detected (Figure 4.6a). At this point, to simplify the broad analysis of such a complex system, we opted to focus on the study of the predominant dimers ($M_{1\text{-}21}$) and to consider, as a first approximation, six families of dimers depending on the charge of

[f] Bis(2-hydroxyethyl)aminotris(hydroxymethyl)methane.
[g] See reversibility tests in section 4.5.4.
[h] The number of possible dimers (N_D) and trimers (N_T) for a system generated from "n" bipodal BBs with a C_2-symmetry, can be calculated as follows:

$$N_D = \frac{n \cdot (n+1)}{2};\ N_T = n \cdot \left(\frac{1}{6} n^2 + \frac{1}{2} n + \frac{1}{3} \right)$$

their constituent BBs (Scheme 4.1 and Figure 4.6a). Thus, $M_{1\text{-}3}$ constitute $F_{(+,+)}$, $M_{4\text{-}7}$ constitute $F_{(0,+)}$, $M_{8\text{-}11}$ constitute $F_{(+,-)}$, $M_{12\text{-}14}$ constitute $F_{(0,0)}$, $M_{15\text{-}18}$ constitute $F_{(0,-)}$ and $M_{19\text{-}21}$ constitute $F_{(-,-)}$. The only 6 dimers with very low concentrations were those of the two families combining BBs with charges of the same sign ($F_{(+,+)}$ and $F_{(-,-)}$).



$F_{(+,+)}$	$F_{(0,+)}$	$F_{(+,-)}$	$F_{(0,0)}$	$F_{(0,-)}$	$F_{(-,-)}$
M_1: **(1j)$_2$**	$M_{4,5}$: **1b-1j**	M_8: **1i-1j**	M_{12}: **(1b)$_2$**	M_{15}: **1b-1i**	M_{19}: **(1i)$_2$**
M_2: **1j-1k**	$M_{4,5}$: **1d-1j**	M_9: **1h-1j**	M_{13}: **1b-1d**	$M_{16\text{-}18}$: **1b-1h**	M_{20}: **1h-1i**
M_3: **(1k)$_2$**	M_6: **1b-1k**	M_{10}: **1i-1k**	M_{14}: **(1d)$_2$**	$M_{16\text{-}18}$: **1d-1h**	M_{21}: **(1h)$_2$**
	M_7: **1d-1k**	M_{11}: **1h-1k**		$M_{16\text{-}18}$: **1d-1i**	

Figure 4.6. HPLC-UV traces (254 nm) of the DCL generated by the mixture of the six BBs **1b**+**1d**+**1h**+**1i**+**1j**+**1k** at 0.5 mM each in 40 mM bis-Tris buffer (pH 6.5) with 25% (v/v) DMSO in the absence of salt (a) and in the presence of 1.0 M NaCl (b).

4.3.3. Adaptation of the large DCL to the increase of the ionic strength

Interestingly, when performing the same DCL in the presence of increasing concentrations of salt (0.5-2.0 M NaCl), the amplification trends of all the dimers were found to be very similar among the members of the same family, validating our initial classification (Figure 4.6). At a first glance, the most notable adaptive change in the composition of the library as the salt content reaches the 2.0 M NaCl concentration is the five-fold increase of the two initially disfavored families $F_{(+,+)}$ and $F_{(-,-)}$ (from <2% to ~10% of the overall population of dimers). Also remarkable is the decrease in the

concentration of the family $F_{(+,-)}$ (from >45% to ~25%). Finally, less pronounced changes were observed for families $F_{(0,+)}$, $F_{(0,-)}$ and $F_{(0,0)}$: the salt increases the presence of $F_{(0,+)}$ and $F_{(0,-)}$ (from 16-18% each to ~21%) but decreases $F_{(0,0)}$ (from ~15% to ~10%). The representation of the HPLC percent area of the 6 families of members in front of the salt concentration (Figure 4.7) illustrates the described trends and points out a similar analogous paired behavior for families $F_{(+,+)}$ with $F_{(-,-)}$, and $F_{(0,+)}$ with $F_{(0,-)}$.

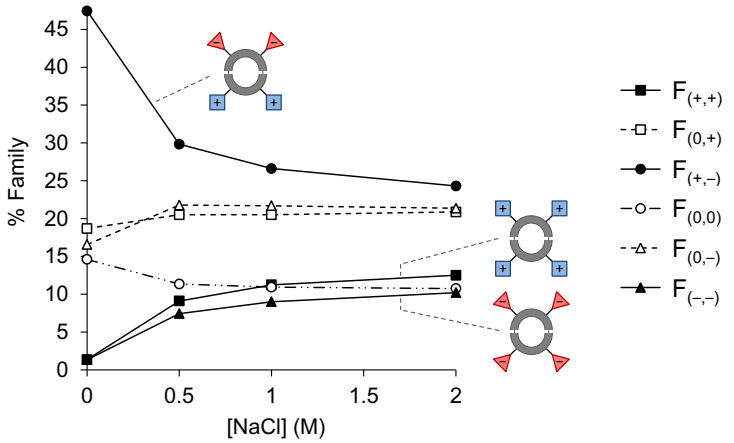


Figure 4.7. Plot of the percentage concentration of each family of members against the concentration of NaCl (M).

The statistical composition of the library, taking into account the 1:2 homo/heterodimer ratios,[54] can be calculated as 11.1% for each of the families $F_{(+,+)}$, $F_{(0,0)}$ and $F_{(-,-)}$, whereas it is 22.2% for the families $F_{(0,+)}$, $F_{(+,-)}$ and $F_{(0,-)}$. The plot in Figure 4.7 clearly shows that in the absence of salt the composition of the library is markedly different from the statistical distribution: the library is initially far from the statistically favored situation. However, the increase of the salt concentration appears to asymptotically approach the percentage concentration of each family to the corresponding values of 11.1% and 22.2%. This adaptive trend suggests the salt being disruptive to the stabilizing and destabilizing effects responsible for the initial distribution.

4.3.4. Unravelling the rules governing the co-adaptive process

Aiming to parameterize and eventually understand the salt-induced response of the complex DCL generated from **1b**+**1d**+**1h**+**1i**+**1j**+**1k**, we envisioned a set of simple adaptation rules that would describe the whole co-adaptive process. Similar to what happens in the evolution of natural ecosystems, there must be some inner information in

each member of the system that is expressed in the dynamic process for the adaptation to the induced stimulus.

The adaptation rules should have a general scope. Thus, for a given initial and final situations of an adaptive process, the rules should be applicable to any possible combination of the 6 BBs used herein. Additionally, the adaptation rules should be as simple as possible.

4.3.4.1. Dynamic deconvolution[55-56] into binary mixtures

The simplest dynamic system able to experience a population change as a response to an external stimulus must still be composed of more than one interrelated member. For the case of a dimer-based DCL, the smallest network able to show adaptive changes is the one formed by the mixture of two BBs (**A** and **B**). As previously commented in Chapter 2 (sections 2.3.4), this binary mixture leads to the formation of two homodimers (**AA** and **BB**) and one heterodimer (**AB**). We envisioned that the adaptation of the large DCL of 21 dimers could be understood as the combination of the contributions of all the corresponding binary mixtures. These small systems would contain the irreducible information of all the co-adaptive relationships. Thus, these informational portions would be an expression of the adaptation rules. Consequently, we studied the 15 possible binary mixtures of the 6 BBs in the absence of salt and in the presence NaCl (1.0 M).

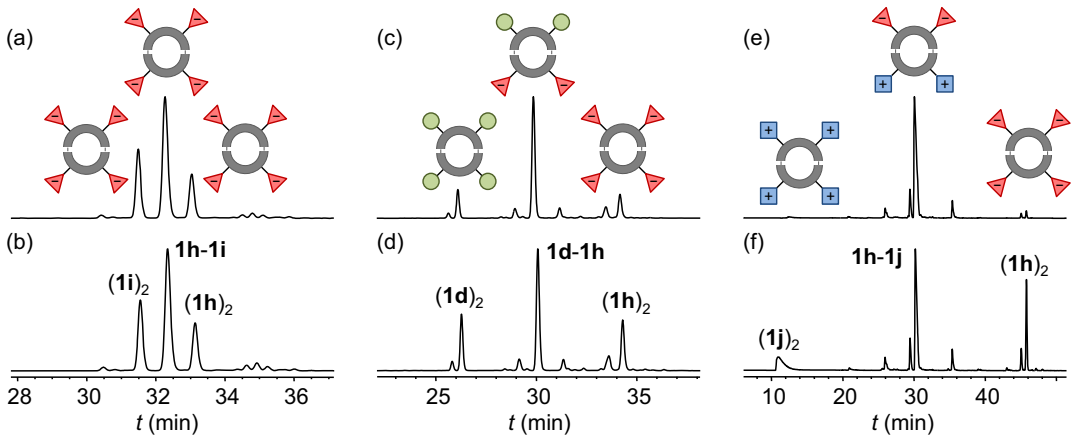


Figure 4.8. HPLC-UV traces (254 nm) of representative binary mixtures performed at 1 mM of each BB in 40 mM bis-Tris buffer (pH 6.5) with 25% (v/v) DMSO. Mixture of **1h**+**1i** in the absence (a) and presence of 1.0 M NaCl (b); **1d**+**1h** in the absence (c) and presence of 1.0 M NaCl (d); and **1h**+**1j** in the absence (e) and presence of 1.0 M NaCl (f).

In Figure 4.8, selected HPLC traces of the binary mixtures are shown. These representative examples allow a qualitative preliminary evaluation of the different general behaviors observed. The binary mixtures of neutral or equally charged BBs (**1h**+**1i**, **1b**+**1d** and **1j**+**1k**) do not experience appreciable population changes when increasing the salt concentration (Figure 4.8a-b). The binary mixtures of a neutral and a charged BB (**1b**+**1h-k** and **1d**+**1h-k**) suffer a disruptive selection, increasing the ratio of homodimers in the presence of salt (Figure 4.8c-d). Finally, the binary mixtures of differently charged BBs (**1h**+**1j-k** and **1i**+**1j-k**) also experience a disruptive selection, now with a drastic increase in the concentration of the homodimers: the mixtures rearrange their distribution from a starting composition in which these are barely detected, to a situation in which they have a considerable presence (Figure 4.8e-f).

4.3.4.2. Quantitative evaluation of the binary mixtures

For the quantification of the concentration changes exhibited by a DCL upon the introduction of an external stimulus, quite commonly the amplification factor (AF),[57-60] *e.g.* Chapter 3, and recently the normalized amplification factor (AF_n),[61] have been used. Although the amplifications depend at last on stability changes, they are also strongly affected by the overall mass balance of the library. Thus, the AF of a certain member can dramatically change depending on its initial concentration and the availability of its constituent parts within the library (amplifications are intrinsically affected by the changes in stability of the other members of the library). In order to find selection rules of general character, the quantitative evaluation of the binary mixtures cannot be based on a parameter that depends on the composition of the system. Alternatively, aiming to avoid the mass-balance influence, we decided to quantify the salt-induced changes by using thermodynamic equilibrium constants. These do not provide us directly with information about the adaptive characteristics but with information about the initial and final situations of the adaptive process.

In a disulfide-based DCL, because the starting free thiols are not present at the fully oxidized final composition, the equilibrium constants involving these virtual species (formation constants) are not experimentally practical. However, the equilibrium constants between disulfides (exchange constants) are readily accessible. For the quantitative description of the binary mixtures we used the exchange constant $K_{[\mathbf{A},\mathbf{B}]}$ previously defined in Chapter 2 (Equation 2.1). As already commented, this constant

presents the main advantage of being dimensionless: its value does not depend on the actual concentration of any of the two BBs and this provides experimental robustness. In the case of statistical distribution of dimers **AA**, **AB** and **BB** (1:2:1 for equimolar concentration of **A** and **B**),[54] the constant $K_{[\mathbf{A},\mathbf{B}]}$ has the value of 0.25.

The exchange constant $K_{[\mathbf{A},\mathbf{B}]}$ was independently determined for the 15 binary mixtures, both in the absence and presence of NaCl (1.0 M). Not surprisingly, the constants in the absence of salt have very different values depending mainly on the charge of the involved members. Thus, the binary mixtures with the three dimers of the same family, formed from the mixture of neutral or equally charged BBs, have constants with values close to 0.25 (Table 4.1, entries 1-3), indicating that the heterodimers are not favored in front of the homodimers and *vice versa*. The binary mixtures formed from the mixture of a neutral and a charged BB have constants with values between 0.04 and 0.2 (Table 4.1, entries 4-11), indicating that the heterodimers are slightly favored in front of the homodimers. Finally, the binary mixtures formed by a positively and a negatively charged BB have constants with values ≤0.002 (Table 4.1, entries 12-15), indicating that the heterodimers are much more stable than the homodimers (≥15 kJ·mol^{-1}).

Table 4.1. Values of the exchange constant in the absence ($K^{\text{No salt}}_{[\mathbf{A},\mathbf{B}]}$) and presence of 1.0 M NaCl ($K^{\text{1M NaCl}}_{[\mathbf{A},\mathbf{B}]}$), the ratio $K^{\text{1M NaCl}}_{[\mathbf{A},\mathbf{B}]}/K^{\text{No salt}}_{[\mathbf{A},\mathbf{B}]}$ and the $-\Delta\Delta G°$ energy, for the 15 binary mixtures.

entry	BB_1	BB_2	$K^{\text{No salt}[a]}_{[\mathbf{A},\mathbf{B}]}$	$K^{\text{NaCl}[a]}_{[\mathbf{A},\mathbf{B}]}$	$K^{\text{NaCl}}_{[\mathbf{A},\mathbf{B}]}/K^{\text{No salt}[a]}_{[\mathbf{A},\mathbf{B}]}$	$-\Delta\Delta G°^{[b]}$
1	**1h**	**1i**	0.18	0.19	1.08	0.19
2	**1d**	**1b**	0.24	0.24	1.01	0.01
3	**1j**	**1k**	0.26	0.29	1.13	0.30
4	**1h**	**1d**	0.044	0.17	3.92	3.39
5	**1h**	**1b**	0.043	0.18	4.16	3.53
6	**1i**	**1d**	0.069	0.20	2.88	2.63
7	**1i**	**1b**	0.093	0.27	2.96	2.69
8	**1j**	**1d**	0.081	0.22	2.69	2.46
9	**1j**	**1b**	0.13	0.36	2.68	2.44
10	**1k**	**1d**	0.093	0.24	2.58	2.35
11	**1k**	**1b**	0.19	0.46	2.39	2.16
12	**1h**	**1j**	<0.001	0.17	>170	>13
13	**1h**	**1k**	0.002	0.22	110	12
14	**1i**	**1j**	<0.001	0.11	>110	>12
15	**1i**	**1k**	0.002	0.14	70	11

[a] Values with a <3% estimated error. [b] Values shown in kJ·mol^{-1}, calculated for T = 298 K, and with a <3% estimated error for entries 4-15.

Regarding the constant values of the mixtures with 1.0 M NaCl ($K^{NaCl}_{[\mathbf{A},\mathbf{B}]}$), they present much smaller differences and, in general, the corresponding values are closer to the statistical value. This indicates a generally much less energetic difference between the homodimers and the heterodimers at higher ionic strengths.

The relative stability of a dimer in a DCL depends on how favorable it is to join the two constituent parts compared to the analogous process for the other members of the library. In the binary mixtures, when the two combined BBs are neutral or equally charged, the repulsive or attractive interactions are similar for the three dimers and the statistical distribution is roughly obtained. When a neutral and a charged BB are mixed, charges of the same sign have to approach to form only one of the homodimers. The corresponding electrostatic repulsions make this homodimer to have a lower relative stability, leading to a decrease in its concentration and, to fulfil the mass balance, also a decrease in the concentration of the other homodimer. Finally, when the two combined BBs have charges of different sign, all the electrostatic interactions are pushing in the same direction: the two homodimers are disfavored by repulsive interactions, and the heterodimer is favored by attractive interactions.[62] The higher ionic strength increases the relative permittivity (ε_r) of the solvent, shielding both repulsive and attractive forces. With this attenuation, the composition of the mixtures approaches that of the statistical distribution.

So far, only the initial and final situations of the overall process have been characterized. To get information about the full process, we calculated the ratio between the corresponding exchange constants ($K^{NaCl}_{[\mathbf{A},\mathbf{B}]}/K^{No\ salt}_{[\mathbf{A},\mathbf{B}]}$ values in Table 4.1). The result of this simple calculation is a measure of the magnitude of the responsive changes. This ratio and its associated $\Delta\Delta G°$ energy show the different influences of the salt on the mixtures and quantifies how much, in the absence of salt, the binary mixtures are shifted away from the final situation in 1.0 M NaCl. The lower the $\Delta\Delta G°$ is (*i.e.*, the larger the $-\Delta\Delta G°$ values represented in Table 4.1), the larger the stability change caused by the salt. This energetic parameter is very useful for the exclusive comparison of the effect of the salt. The $-\Delta\Delta G°$ energies are roughly zero ($\leq$0.3 kJ·mol^{-1}) for the mixtures formed from neutral or equally charged BBs (Table 4.1, entries 1-3), 2-4 kJ·mol^{-1} for the binary mixtures generated from a neutral and a charged BB (Table 4.1, entries 4-11),

and >10 kJ·mol^{-1} for the binary mixtures generated from a positively and a negatively charged BB (Table 4.1, entries 12-15).

Apart from these main trends, additional interesting information can be obtained from Table 4.1. In the binary mixtures of entries 4-11, only one of the three involved dimers contains two charged constituents. Therefore, the changes in concentration caused by the salt mainly depend on the stability changes of this particular dimer, the only one suffering intramolecular electrostatic interactions. When comparing the mixtures of a negative and a positive BB separately with the same neutral BB, the $-\Delta\Delta G°$ energy is larger for the negative than for the positive (Table 4.1, entries 4-7 *vs.* 8-11). This indicates a more efficient adaptation of the anionic macrocycles to the saline medium. This effect is due to a lower relative energy of the members of family $F_{(-,-)}$ compared to the ones of $F_{(+,+)}$. If we go back to the large DCL generated from the 6 BBs, Figure 4.6 shows a larger presence of $F_{(+,+)}$ than $F_{(-,-)}$ at all salt concentrations, in agreement with the stability order $F_{(+,+)} > F_{(-,-)}$. Finally, small differences between members of the same family can also be observed. The mixtures of **1h** or **1i** with the neutral **1b**,**d** show larger $-\Delta\Delta G°$ energies for **1h** than for **1i** (Table 4.1, entries 4 *vs.* 6 and 5 *vs.* 7), denoting a larger stability of $(\mathbf{1i})_2$ than $(\mathbf{1h})_2$. Similarly, the respective mixtures of **1j** or **1k** with the neutral **1b**,**d** show larger $-\Delta\Delta G°$ energies for **1j** than for **1k** (Table 4.1, entries 8 *vs.* 10 and 9 *vs.* 11), denoting a larger stability of $(\mathbf{1k})_2$ than $(\mathbf{1j})_2$. This trend correlates with the distance between the charges situated at the residues and the macrocyclic backbone. Thus, the longer that distance is, the more stable the species (section 4.3.5 for a detailed discussion on that topic).

4.3.4.3. Simulation with DCLSim software

We used the values in Table 4.1 to simulate the concentration of the 21 dimers by means of the DCLSim 1.1 software.[60] The simulated percentage amounts of each family, together with the corresponding experimental ones, were represented separately for the cases of absence of salt and presence of 1.0 M NaCl (Figure 4.9a and 4.9b, respectively). The simulated and experimental concentrations of each family showed an excellent agreement, which corroborates our hypothesis. The whole population distribution of our complex DCL generated from **1b**+**1d**+**1h**+**1i**+**1j**+**1k** is determined by the 15 corresponding $K_{[\mathbf{A},\mathbf{B}]}$ constants. Thereby, the complete information about the adaptive process is comprehended in the constant ratios $K^{\text{NaCl}}_{[\mathbf{A},\mathbf{B}]}/K^{\text{No salt}}_{[\mathbf{A},\mathbf{B}]}$ of Table 4.1.

Consequently, these values would eventually allow us to predict any equivalent adaptive process of different combinations of the same BBs.

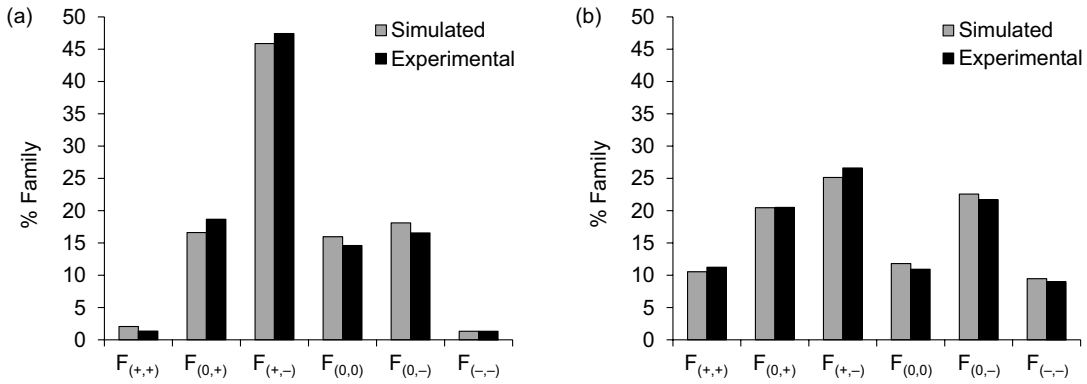


Figure 4.9. Representation of the simulated (gray) and experimental (black) percentage concentration of the six families of members for the case with no salt (a) and the case with 1.0 M NaCl (b).

4.3.5. Structural studies

The members of the dynamic library most affected by the ionic strength were the homodimers containing charged side chains. However, two intriguing trends were observed in the data contained in Table 4.1: i) macrocycles with positive charges are intrinsically more stable and, thus, less affected by the increase of polarity, than negative charges; ii) within each family, the shorter the length of the side chain, the less stable the homodimer and, therefore, more prone to amplify when increasing the salt content.

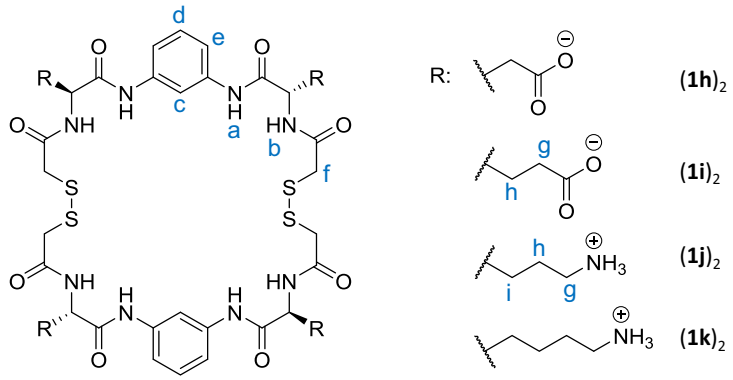


Figure 4.10. Chemical structure of the charged homodimers.

The effect of the length of the side chain can be rationalized by simple electrostatic forces. Because the electrostatic potential decreases with distance, the repulsion between charges of the same sign is stronger for shorter connections. Considering the

molecular structures of the homodimers (Figure 4.10), they share a common central macrocyclic scaffold, and they differ in the average distance between this scaffold and the extremes of the side chains (where the charged atoms are). This will make the electrostatic repulsion in the Asp derivative (two carbon atoms in the connections) bigger than that in the Glu counterpart (three carbon atoms). A parallel conclusion can be derived for Orn/Lys (three *vs*. four carbon atoms, respectively).

The differences between positive and negative charges are more difficult to explain with simple comparisons. This effect is perfectly illustrated by the behavior of the Glu (**1i**) and Orn (**1j**) derivatives, both bearing side chains with comparable lengths, *i.e.* the same number of carbon atoms, but different signs of the charges. Recently, it has been reported that the hydration energies of negative and positive charges are different due to the asymmetry of the water molecule.[63] However, these differences favor the solvation of negative charges, which would stabilize the anionic side chains, rendering exactly the reverse effect than the one observed. Therefore, we hypothesized that additional structural/conformational factors might be playing an important role. To unravel those effects, we performed a detailed NMR and molecular modeling study with the two homodimers derived from Glu ((**1i**)$_2$) and Orn ((**1j**)$_2$).

4.3.5.1. Nuclear magnetic resonance

First of all, the careful NMR titrations of both macrocycles were performed at representative pH values: pH 1.4-10.6 for (**1i**)$_2$ and pH 1.6-14.0 for (**1j**)$_2$ (see the spectra of the NMR titration in section 4.5.6). The analysis of the variations of the chemical shifts showed that (**1i**)$_2$ is fully deprotonated at pH ≥ 6.0 and (**1j**)$_2$ is fully protonated at pH ≤ 7.3. Therefore, at pH 6.5 both macrocycles have four (positive/negative) charges.

The ^{1}H NMR spectra of these dimers at pH ~6.5 are compared in Figure 4.11. The first striking observation is that, at this pH, the amide protons (signals a and b in Figure 4.11) are observable for (**1i**)$_2$ but not for (**1j**)$_2$. Considering the conditions used for the NMR experiments (section 4.5.6), we deduced that the amide protons of (**1j**)$_2$ exchange with the solvent much faster than those of (**1i**)$_2$. This observation would imply that the amido NH protons of the Orn derivative are exposed to the solvent, whereas those from the Glu macrocycle are somehow protected from the solvent. This suggests that the amide protons in (**1i**)$_2$ would be implicated in intramolecular H-bonds, which is in agreement with their observed chemical shifts. The formation of intramolecular H-

bonds could stabilize some sort of folded structures in solution, in which the carboxylates of the side chains are close to the macrocycle scaffold. The proximity of the anionic charges could also explain the upfield shift of the aromatic protons (c,d), especially for the signal c. Unfortunately, 2D NOESY, 2D ROESY and 1D ROESY experiments were inconclusive due to the high symmetry and flexibility of the macrocycle (see the corresponding spectra in section 4.5.6).

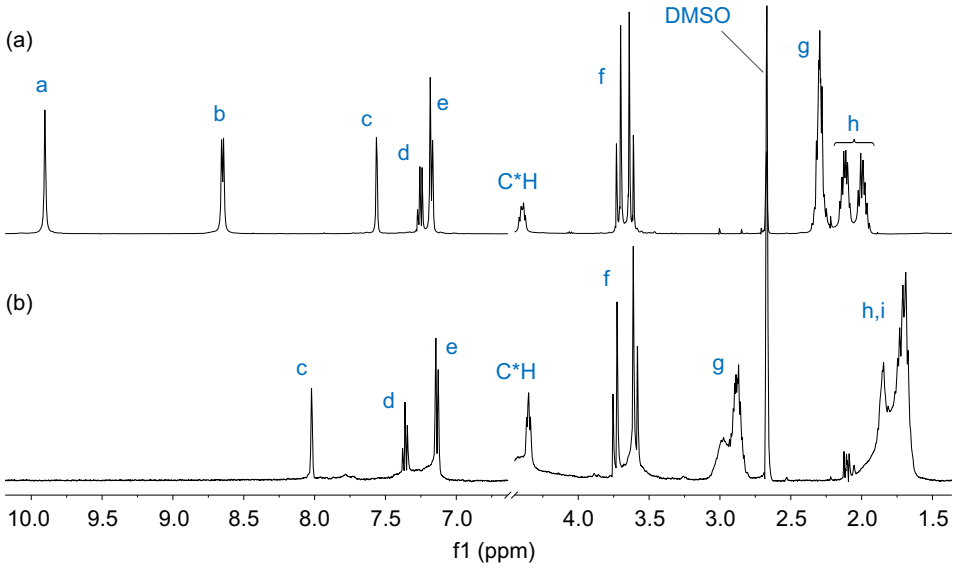


Figure 4.11. Partial ^{1}H NMR (500 MHz, "WATER_ES" water suppression pulse sequence, 298 K) spectra of (a) 2 mM (**1i**)$_2$ dimer in H_2O with 25% (v/v) DMSO-d_6 at pH ~6.5, and (b) 2 mM oxidized **1j** in H_2O with 25% (v/v) DMSO-d_6 at pH ~6.5. The assignment letters (a-h) correlate with Figure 4.10.

4.3.5.2. Molecular dynamics simulations

To further support this hypothesis, we decided to perform molecular modeling simulations on (**1i**)$_2$ and (**1j**)$_2$. In this case, the ambitious aim to observe potential conformational differences between these two homodimers sharing a common macrocyclic scaffold, prompted us to perform an exhaustive computational study by means of molecular dynamics simulations in explicit solvent molecules. These calculations were done in collaboration with Dr. Jordi Bujons (IQAC-CSIC).

Molecular dynamics (MD) simulations (100 ns each) on (**1i**)$_2$ and (**1j**)$_2$ were performed in triplicate, in a completely ionized state and in 25% (v/v) DMSO/water explicit solvent mixture (1:12 DMSO/water molar ratio, Figure 4.21 in section 4.5.7), and from different starting conformations to improve sampling (simulations 1-6, Figures

4.22 and 4.23 in section 4.5.7).[i] In good agreement with the NMR spectroscopic results, the simulations show that dimer $(\mathbf{1i})_2$ establishes multiple intramolecular H-bond interactions during most of the simulation time (2-3 on average, Figures 4.12b and 4.22), whereas $(\mathbf{1j})_2$ exhibits a much lower number (0-1 on average, Figures 4.12e and 4.23). The distances between pairs of carbon atoms of the carboxylates ($(\mathbf{1i})_2$) and pairs of carbons that hold the amino groups ($(\mathbf{1j})_2$), which are separated by the same number of bonds, were used to compare the distance between charged groups. These distances are, on average, about 1 Å shorter for $(\mathbf{1i})_2$ than for $(\mathbf{1j})_2$ (Figures 4.12a,d and 4.22 *vs.* 4.23), indicating that the carboxylates are generally closer than the ammonium groups. Furthermore, looking at the variation of the areas of the exposed polar and hydrophobic surfaces of the two compounds[64] (Figures 4.12c,f and 4.22 *vs.* 4.23), the area of the polar surface is similar but the area of the hydrophobic surface is substantially smaller for $(\mathbf{1i})_2$ (500-650 $Å^2$)[j] than for $(\mathbf{1j})_2$ (650-800 $Å^2$). Thus, these results indicate that under the simulation conditions compound $(\mathbf{1i})_2$ adopts more packed conformations, with less exposed hydrophobic moieties, than that of $(\mathbf{1j})_2$.

Additionally, the profiles shown in Figure 4.12c,f suggest that during the simulations both compounds switch between at least two states, namely one with a smaller total accessible area, *i.e.* a more packed/folded state, and another one with a larger total accessible area, *i.e.* a more unfolded/unpacked state. This is also observed in the radius of gyration profiles (Figures 4.22 and 4.23). These different packing/folding states are most evident when looking at the snapshots of the simulations.

[i] Videos of representative MD simulations performed for $(\mathbf{1i})_2$ and $(\mathbf{1j})_2$ can be found in the electronic Annex accompanying this thesis.

[j] These ASA values cannot be directly compared with the values previously calculated in Chapter 3 (section 3.3.7 and 3.5.7.1). First of all, the corresponding calculations were carried out by means of different computational methods, as specified in the corresponding experimental sections. Secondly, the values commented here were separately calculated for the polar and the hydrophobic surfaces, whereas in Chapter 3 the ASA values were calculated for the overall surface.

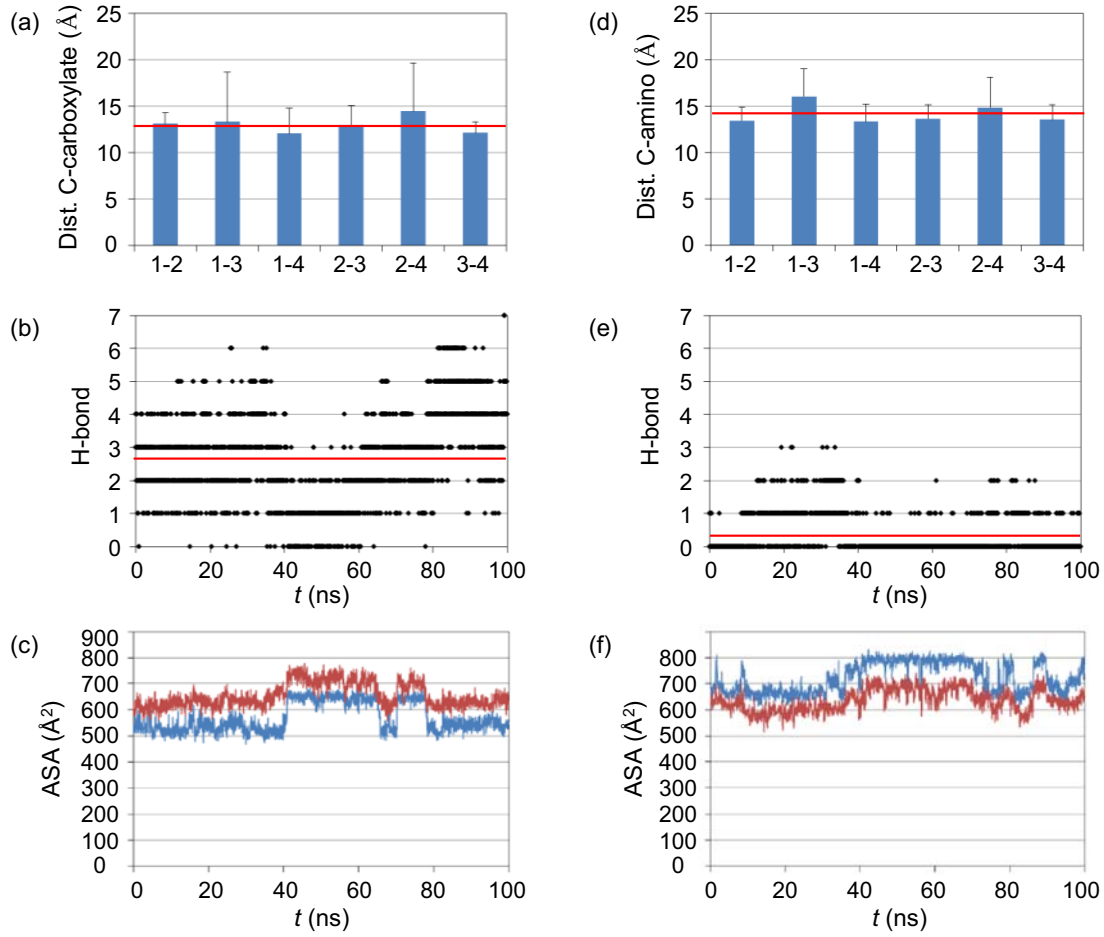


Figure 4.12. Summary of results obtained from two representative 100 ns MD simulations (simulations 2 and 5) carried out with homodimers (**1i**)$_2$ (a-c) and (**1j**)$_2$ (d-f). The graphics represent the distance between (a) the carbon atom of the four carboxylates or (d) those that hold the four protonated amino groups, averaged over the whole simulation; as well as the variation of (b,e) the number of intramolecular hydrogen bonds; and (c,f) the ASA of the polar (red) and hydrophobic (blue) surfaces against simulation time (ns). The horizontal red lines on the above four graphics represent the corresponding global averaged values. Similar results were obtained for the other two simulations performed for each compound (Figures 4.22 and 4.23).

Thus, 2500 snapshots from each simulation were saved and the conformations of (**1i**)$_2$ and (**1j**)$_2$ were clustered according to their heavy atoms root mean square deviation (RMSD). The most populated clusters from these simulations (Figure 4.13a,d) show conformations that are mostly packed, whereas the second relevant clusters (Figure 4.13b,e) show conformations that are mostly extended or unpacked, and the third ones (Figure 4.13c,f) are conformations that could be considered in between. As previously stated, compound (**1i**)$_2$ shows multiple H-bond interactions between the carboxylates and the amide groups of the macrocycle and also between oppositely placed amide groups (Figure 4.13a-c). This was also observed in the other two simulations ran for this compound (Figures A138-140 in the electronic Annex).

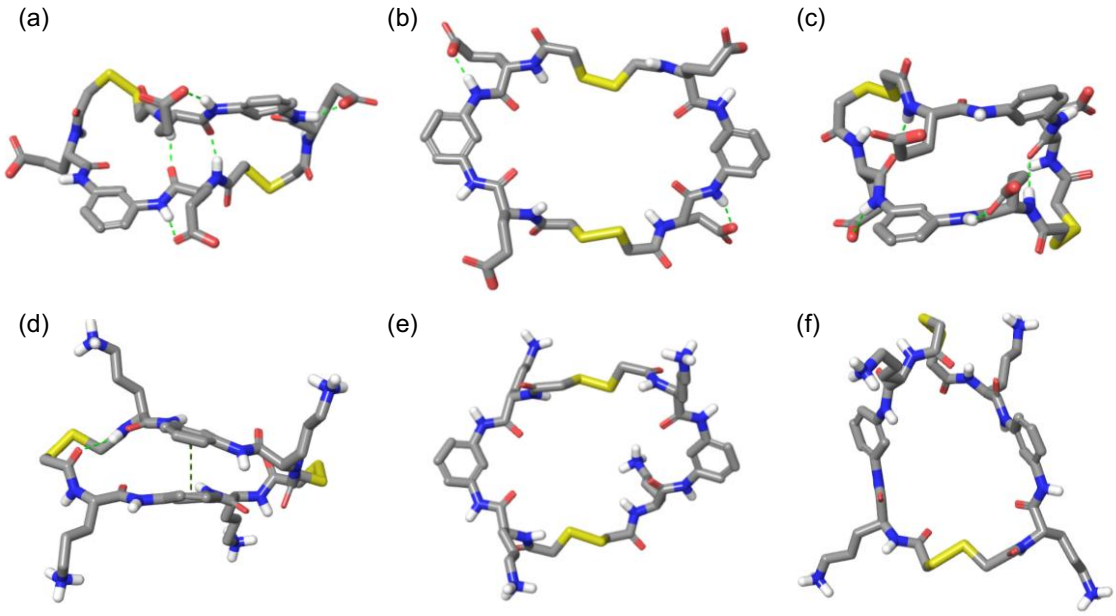


Figure 4.13. Most representative conformations of the most populated conformational clusters derived from simulations 2 and 5, performed with homodimers (**1i**)$_2$ (a-c) and (**1j**)$_2$ (d-f). The relative populations of each cluster are: (a) 55%, (b) 19%, (c) 15%, (d) 41%, (e) 21% and (f) 10%.

On the contrary, the conformations of (**1j**)$_2$ (Figure 4.13d-f) showed no intramolecular H-bonds even in the most packed state (Figure 4.13d). This packed conformation of (**1j**)$_2$ is the only one showing an intramolecular π-π stacking interaction between the two aromatic rings, being the side chains that hold the ammonium groups mostly extended towards the solvent. Only a few of the clusters determined from some of the simulations of (**1j**)$_2$ showed some intramolecular H-bond interactions (Figures A141-143 in the electronic Annex), as was already evident from Figure 4.12e, being the π-π stacking and hydrophobic interactions the main forces that stabilize the more packed conformations of this compound.

These results strongly support the previously formulated hypothesis that the amide protons of (**1i**)$_2$ are much more protected from exchange with the solvent than those of (**1j**)$_2$. Overall, the structural analysis showed the prevalence of more packed conformations for the negatively charged homodimers compared with the positively charged counterparts. These packed conformations set the charges closer and, therefore, they are more sensitive to the polarity of the medium in agreement with the values shown in Table 4.1.

4.3.6. The rational design of co-adaptive relationships

The establishment of the adaptation rules, apart from allowing us to give a reasonable explanation to the complex processes, should also be useful to rationally design controlled adaptive processes. Taking biological evolution in natural ecosystems as a source of inspiration, we envisioned a way to set up dynamic mixtures of pseudopeptidic macrocycles with controlled behaviors. Accordingly, we designed systems with competitive and cooperative relationships. These co-adaptive relationships are typically described for replicating molecular networks.[65-67]

The competitive relationship was exemplified by mixing a neutral (**1d**) and two negatively charged BBs (**1h-i**). The mixture was prepared in the presence of different salt concentrations (0.0-2.0 M NaCl) and the concentration for three of the six dimers ((**1h**)$_2$, **1h-1i** and (**1i**)$_2$) was represented as a function of the salt content (Figure 4.14a).[k] These three dimers are disfavored in the absence of salt for approaching charges of the same sign. As the ionic strength increases, the three are stabilized, slightly increasing their proportion in the mixture (from 9-11% to 11-12% for (**1h**)$_2$ and (**1i**)$_2$, and from 23% to 25% for **1h-1i**). This reorganization proceeds in a competitive manner, determined by the availability of **1h** and **1i**, because the three amplified species are competing for the same BBs. For all values of salt concentrations, dimer (**1h**)$_2$ has a lower presence in the library than the statistically equally favored (**1i**)$_2$, in agreement with the lower stability of the first.

The cooperative relationship was exemplified by mixing BBs with negative (**1h**), neutral (**1d**), and positive (**1k**) charges at the same conditions as before. The concentrations of the dimers (**1h**)$_2$, **1h-1k**, and (**1k**)$_2$ were plotted against the salt concentration (Figure 4.14b). For electrostatic reasons, the homodimers (**1h**)$_2$ and (**1k**)$_2$ are disfavored, whereas the heterodimer **1h-1k** is highly favored. As the ionic strength increases, dimers (**1h**)$_2$ and (**1k**)$_2$ are stabilized and **1h-1k** is destabilized. These changes in the library proceed cooperatively: the decrease in the concentration of **1h-1k** has a positive impact in the amplification of (**1h**)$_2$ and (**1k**)$_2$, now suffering much larger concentration changes than in the competitive environment (from 1-2% to 10-15% for (**1h**)$_2$ and (**1k**)$_2$; and from 44% to 19% for **1h-1k**). For all the salt concentrations,

[k] Although only three of the six dimers are represented in Figure 4.14, the percentage concentrations are relative to the overall population of dimers.

compound $(\mathbf{1h})_2$ has a lower presence in the library than the statistically equally favored $(\mathbf{1k})_2$, in agreement with the lower stability of the first.

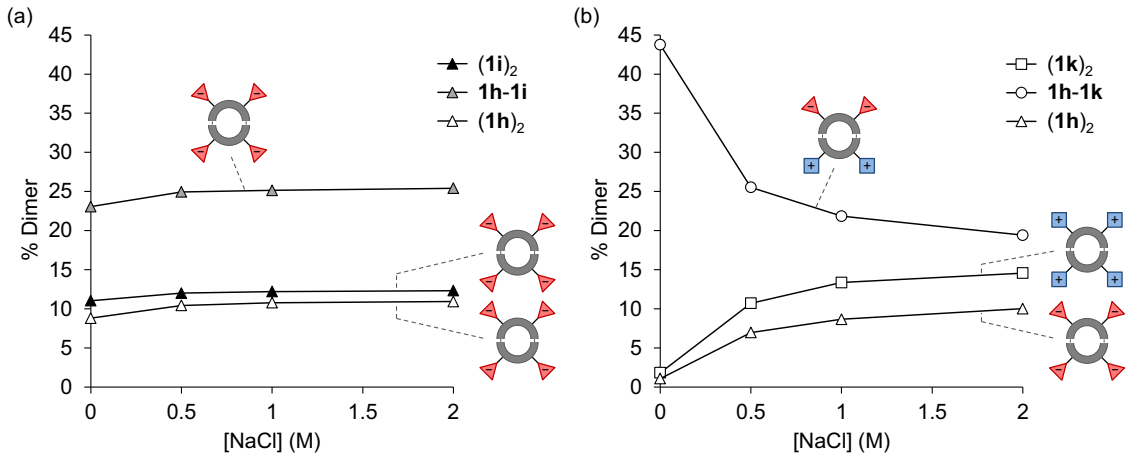


Figure 4.14. Plot of the concentration in percentage of the three dimers containing two charged BBs in front of the salt concentration, for the competitive case (a), and for the cooperative case (b). Both libraries were performed at 1 mM of each BB in 40 mM bis-Tris buffer (pH 6.5) with 25% (v/v) DMSO.

By the careful analysis of the results shown in Figure 4.14, we concluded that the full library composition (Scheme 4.1 and Figure 4.6) is the consequence of a complex network containing both competitive and cooperative relationships. Thus, the population of the members of the dynamic system is governed by the interconnections established between all the congeners. Hence, our dynamic chemical system behaves as a model of a minimalistic ecosystem, in which the behavior of a given species is determined by a combination of its information and the relationships with the members of the mixture (with different information). From the Systems Chemistry perspective, the understanding of the network comprises the characterization of the adaptive properties of every species in a dynamic communication with the other members.

4.4. Conclusions

The study of the salt-induced adaptation of a complex DCL consisting of 21 differently charged macrocyclic dimers has allowed reaching the following main conclusions:

i) In the absence of salt, the library optimizes the electrostatic interactions showing a composition markedly different from the statistical distribution. The increase of the ionic strength shields these interactions and the system approaches the statistically favored proportion. The salt-response of the species can be classified in different families attending to the charges. The families of members concentrating charged BBs are the ones most affected by the addition of salt.

ii) The whole adaptive process has been characterized in a top-down fashion by the dynamic deconvolution into the minimal components, *i.e.* the binary mixtures. These small systems have proved to contain the irreducible information of all the co-adaptive relationships present within the complex library.

iii) The careful analysis of the binary mixtures, together with structural studies performed for selected members of the library, has allowed unravelling interesting adaptive trends. Firstly, for the species bearing charges of the same sign, the salt-adaptation is inversely related to the length of the side chain. Secondly, negatively charged homodimers are more prone to being amplified by the salt than the positively charged counterparts. These two trends are ultimately related to the distance between the charges.

iv) The deep understanding of the behavior of the system as a network has brought us the opportunity to set up dynamic libraries displaying intended relationships. In this way, either competitive or cooperative co-adaptive trends have been reproduced as a response to the salt increase.

Overall, the full library is defined by a complex combination of relationships operating at once. These artificial molecular ecosystems demonstrate the potential of DCLs for mimicking important processes occurring in more elaborate (bio)molecular networks.

4.5. Experimental section

4.5.1 General methods

Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka and Merck) and were used without further purification. pH measurements were performed at room temperature on a Crison GLP21 pH-meter with the electrodes Crison 50 14T (≥10 mL samples) and PHR-146 Micro (NMR samples). NMR spectroscopic experiments were carried out on a Varian INOVA 500 spectrometer (500 MHz for ^{1}H) and a Varian Mercury 400 instrument (400 MHz for ^{1}H).

4.5.2. HPLC and MS analyses

The HPLC and MS analyses were performed as specified in Chapter 2 (section 2.5.2). The HRMS analyses of the DCLs are shown in the electronic Annex.

4.5.3. Preparation and evaluation of the DCLs

A 53.3 mM bis-Tris buffer was prepared by dissolving 558 mg of the free amine (bis(2-hydroxyethyl)aminotris(hydroxymethyl)methane) in 50 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq). Similarly, the same buffer was also prepared containing 0.67, 1.33 and 2.67 M NaCl. For each of them the pH of the solution was adjusted after the addition of the corresponding amount of NaCl.

Individual stocks (12 mM) of **1b**,**d**,**h**,**i**,**j**,**k** were prepared in DMSO by means of the Ellman's test in order to ensure exact concentration (Chapter 2, section 2.5.4.2.). For those experiments to be compared, *e.g.* mixtures of the same BBs in the presence of different salt concentrations, the reaction mixtures were prepared by dilution of a *stock mixture* of the corresponding BBs, ensuring no differences in concentration between the libraries of the same batch (similar process as the one represented in Chapter 3, Scheme 3.2). The DCLs of the complex system generated from the mixture of 6 BBs were prepared at 0.5 mM of each BB in a 40 mM bis-Tris buffer (pH 6.5) with 25% (v/v) DMSO in the presence of 0.0-2.0 M NaCl. The rest of the libraries (reversibility test, binary mixtures and co-evolutionary relationships) were performed at 1.0 mM of each BB in a 40 mM bis-Tris buffer (pH 6.5) with 25% (v/v) DMSO in the presence of 0.0-2.0 M NaCl.

The exchange constant $K_{[\mathbf{A},\mathbf{B}]}$ was calculated as previously explained in Chapter 2 (section 2.5.5, Equation 2.3). The $\Delta\Delta G°$ energy was calculated from the $K^{\text{NaCl}}_{[\mathbf{A},\mathbf{B}]}/K^{\text{No salt}}_{[\mathbf{A},\mathbf{B}]}$ ratio as explained in Chapter 3 (section 3.5.5.2, Equation 3.9). The estimated error for $K^{\text{No salt}}_{[\mathbf{A},\mathbf{B}]}$ and $K^{\text{NaCl}}_{[\mathbf{A},\mathbf{B}]}$, the $K^{\text{NaCl}}_{[\mathbf{A},\mathbf{B}]}/K^{\text{No salt}}_{[\mathbf{A},\mathbf{B}]}$ ratio and the $-\Delta\Delta G°$ energy is based on a ±1% integration error in the chromatographic analysis.[1,68]

4.5.4. Reversibility tests

The reversibility tests were performed following a similar procedure to the one depicted in Chapter 2 (Scheme 2.6). A *pre-equilibrated reaction mixture* was prepared by adding 20 μL of the *individual stocks* (12 mM) of **1d** and **1j** to 180 μL of a 53.3 mM bis-Tris buffer (pH 6.5). After 24 hours, the *reaction mixture **A*** was prepared by adding 15 μL of the *individual stock* of **1h** to 165 μL of the *pre-equilibrated reaction mixture*. Simultaneously, the *control reaction mixture **B*** was prepared by mixing 15 μL of each *individual stock* with 135 μL of a 53.3 mM bis-Tris buffer (pH 6.5). After 24 hours, the *reaction mixture **A***, the *control reaction mixture **B*** and the *pre-equilibrated reaction mixture* were analyzed by HPLC (Figures 4.15a,c,e).

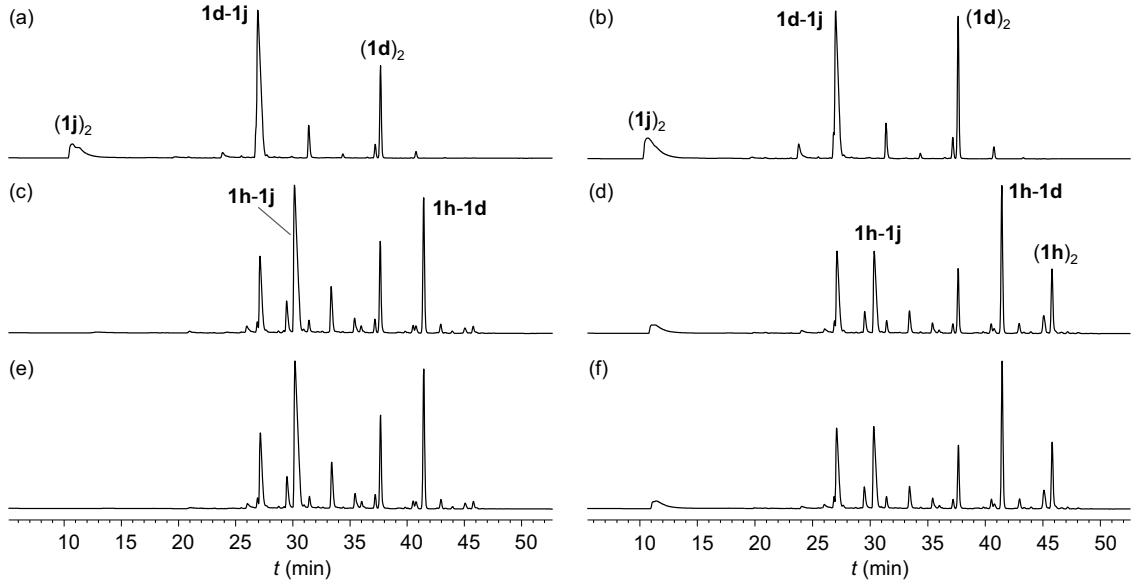


Figure 4.15. HPLC-UV traces (254 nm) of the *pre-equilibrated reaction mixtures* (a-b), the *reaction mixture **A*** (c-d) and the *control reaction mixture **B*** (e-f); performed in the absence of salt (a,c,e) and in the presence of 2.0 M NaCl (b,d,f).

[1] Notice that the real concentration of the two BBs does not affect the exchange constant value and, therefore, does not represent an additional source of error.

After the addition of **1h** to the *pre-equilibrated reaction mixture* containing **1d**+**1j**, the system evolved to the same final situation as when the three BBs were left to oxidize together (Figure 4.15c *vs.* 4.15e). The same reversibility test was also performed in the presence of 2.0 M NaCl (Figure 4.15b,d,f) and *reaction mixtures **A*** and ***B*** also showed identical composition (Figure 4.15d *vs.* 4.15f). Therefore, both in the absence and presence of salt, the library proved to reach the thermodynamic equilibrium.

4.5.5. Simulation with DCLSim software

The purpose of this simulation with DCLSim 1.1 software[60] was to check if the whole population distribution of the complex DCL generated by the mixture of the 6 BBs **1b**+**1d**+**1h**+**1i**+**1j**+**1k** is determined by the 15 corresponding exchange constants $K_{[\mathbf{A},\mathbf{B}]}$. With this purpose, first of all, the formation constants (K_f, Equations 3.3-3.5 in Chapter 3) of the heterodimers of the library were expressed as a function of the exchange constant $K_{[\mathbf{A},\mathbf{B}]}$ as shown in Equation 4.1. This equation was obtained by rearranging Equation 3.6.

$$K_f(\mathbf{AB}) = \sqrt{\frac{K_f(\mathbf{AA}) \cdot K_f(\mathbf{BB})}{K_{[\mathbf{A},\mathbf{B}]}}} \qquad \textbf{(Equation 4.1)}$$

The formation constants of the dimers of the library (input values of the simulation) were obtained as follow: for each of the homodimers, the formation constant was arbitrarily[m] set to the unit, whereas for the each of the heterodimer, the formation constant was calculated by means of Equation 4.1. The simulation was performed at the temperature of 298 K and the six BBs were set to be "virtual" and to have a $0.5 \cdot 10^{-3}$ M concentration each.

The concentration of the 21 dimers (output values of the simulation) was represented, by families, together with the corresponding experimental values (Figure 4.9 in section 4.3.4.3). This was done separately for the mixture prepared in the absence of salt and in the presence of 1.0 M NaCl. The excellent agreement between the experimental and the calculated values clearly corroborated that, both in the absence and in the presence of salt, the whole complex DCL is determined by the 15 corresponding $K_{[\mathbf{A},\mathbf{B}]}$ constants.

[m] It does not matter which are the attributed value to the formation constants of the homodimers: whatever their value is, the whole system will fit if the formation constant of the heterodimers is calculated by means of Equation 4.1.

4.5.6. Nuclear magnetic resonance

For the NMR titration of (**1i**)$_2$ the dimer was prepared on multi-milligram scale from dithiol **1i** as previously explained in Chapter 3 (section 3.5.7.3). A 2 mM (of dimer) NMR sample was prepared by dissolving 1.54 mg of the tetraacid (**1i**)$_2$ in 188 μL of DMSO-d_6 and 563 μL of H_2O. Finally, 10 μL of a 38 mM DSS solution in H_2O were added as an internal reference. Initially, some HCl (aq) was added in order to reach the lowest pH value of 1.4. Then, for the pH screening, NaOH (aq) was added to progressively increase its basicity up to pH 10.6 (Figure 4.16). For the NMR titration of (**1j**)$_2$ the dimer was prepared by *in situ* oxidation of **1j**. A 2 mM (of monomer) NMR sample was prepared by dissolving 1.07 mg of dithiol **1j**·2TFA in 188 μL of DMSO-d_6 and 563 μL of H_2O. The pH was then adjusted to 7.4 with NaOH (aq) and the sample was left to oxidize for 24 hours at room temperature. Finally, 5 μL of a 38 mM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) solution in H_2O were added as an internal reference. For the screening of the alkaline pHs, NaOH (aq) was added, progressively increasing the basicity up to pH 14.0 (Figure 4.17). For the screening of the acidic pHs, an identical sample was used and HCl (aq) was added, progressively increasing the acidity down to pH 1.6 (Figure 4.18). The ^{1}H NMR spectra of the two samples at different pH values were recorded on a Varian INOVA 500 spectrometer (500 MHz for ^{1}H, 298 K) using the Agilent Chempack pulse sequence WATER_ES (with excitation sculpting).[69] All spectra were acquired using the same optimized parameters for the solvent selective pulse and the gradient. The methyl group of the internal standard DSS was used as a reference at 0.00 ppm.

The 2D NOESY, 2D ROESY and 1D ROESY spectra of (**1i**)$_2$ (Figures 4.19 and 4.20) were recorded on a Varian Mercury 400 instrument (400 MHz for ^{1}H, 298 K). The NMR sample of the 2D NOESY and 2D ROESY was prepared by dissolving the tetraacid (**1i**)$_2$ (1 mM of dimer) in 20 mM deuterated phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6. For the 1D ROESY spectra, the NMR sample was prepared in the same conditions but with a non-deuterated buffer, and the Agilent Chempack pulse sequence (H)WET (Water suppression Enhanced through T_1 effects)[70] was used.

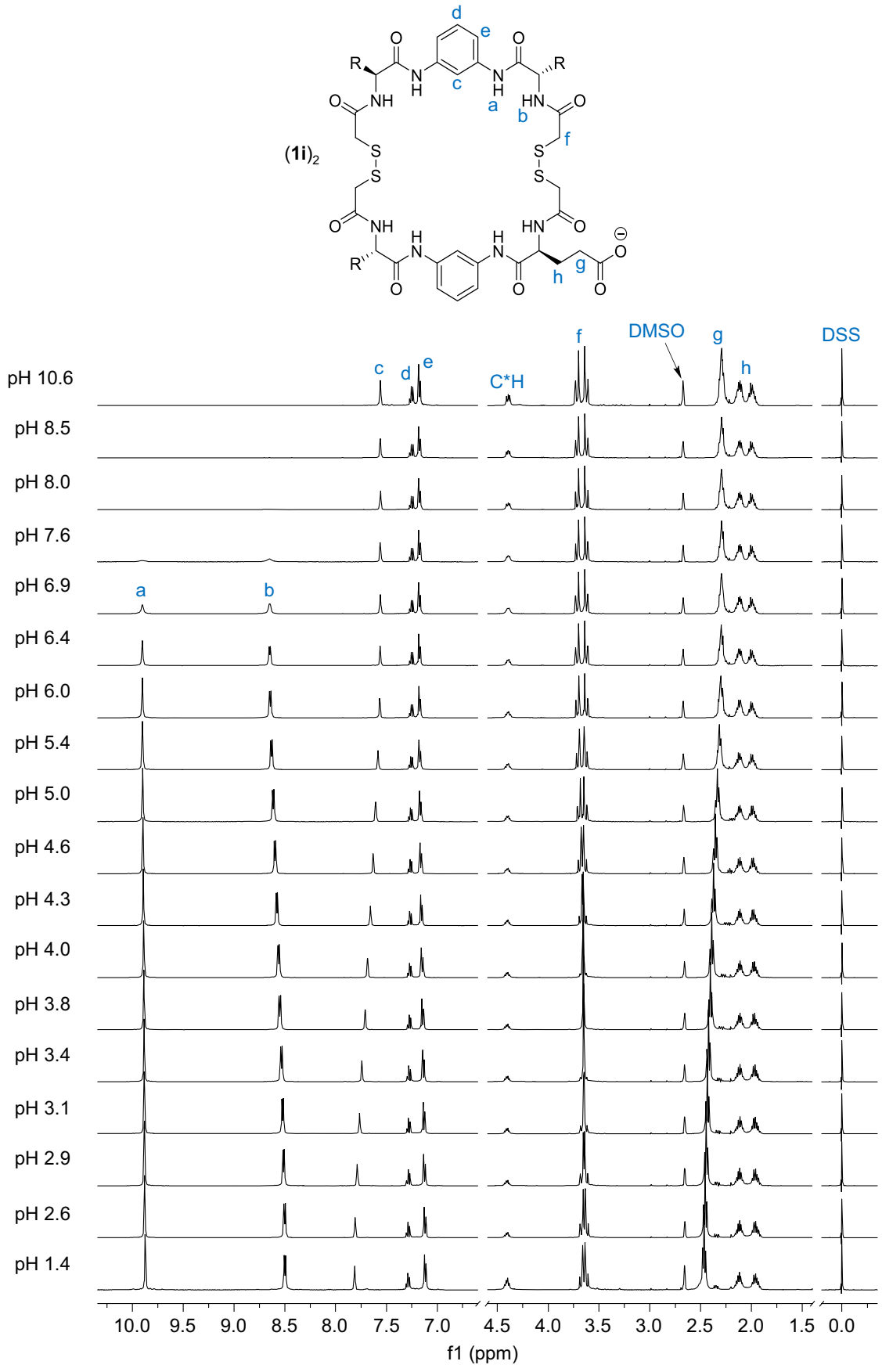


Figure 4.16. Partial ^{1}H NMR (500 MHz, "WATER_ES" water suppression pulse sequence, 298 K) of 2 mM (**1i**)$_2$ dimer in H_2O with 25% (v/v) DMSO-d_6 at pH from 1.4 to 10.6.

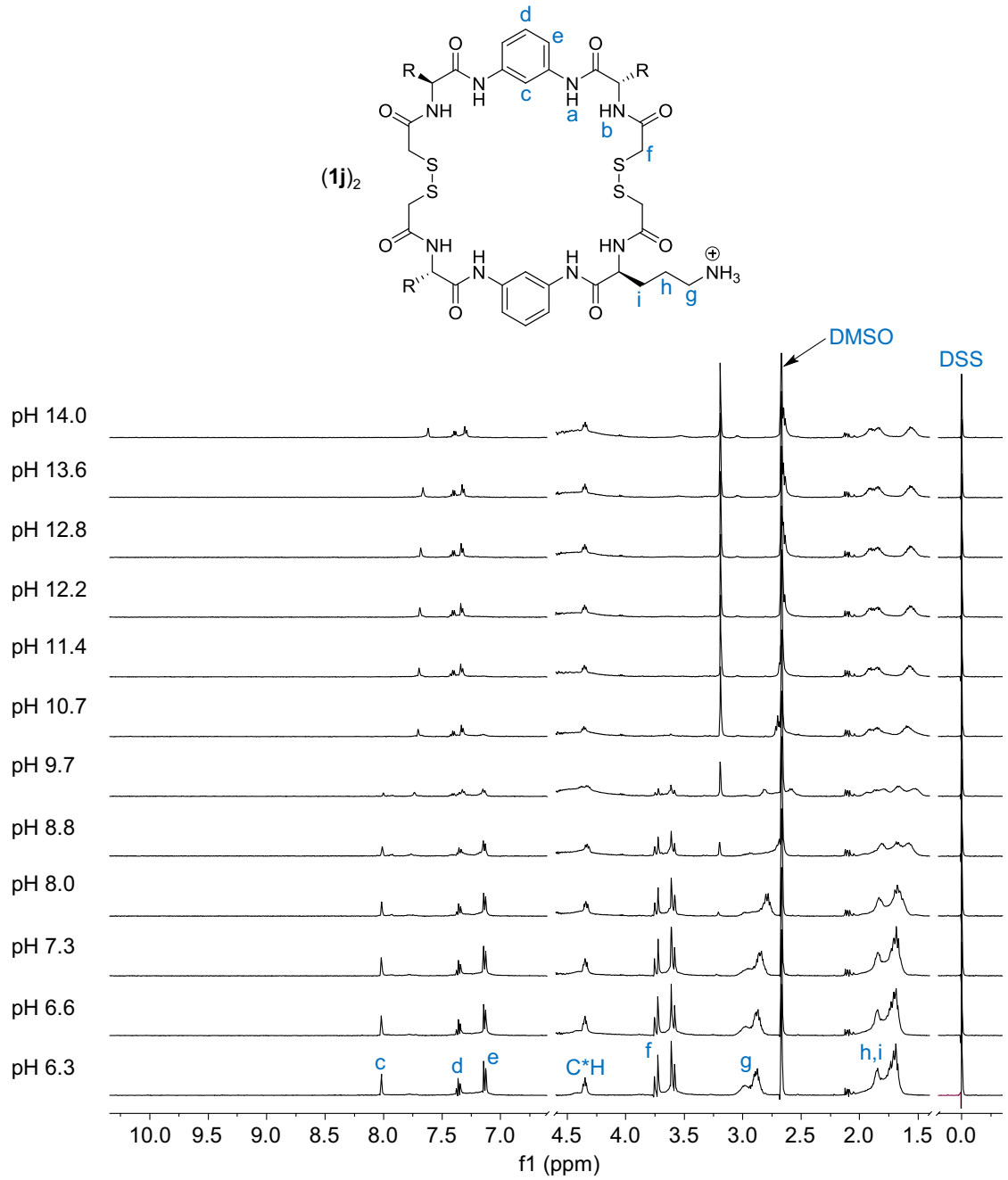


Figure 4.17. Partial ^{1}H NMR (500 MHz, "WATER_ES" water suppression pulse sequence, 298 K) of 2 mM oxidized **1j** in H_2O with 25% (v/v) DMSO-d_6 at pH from 6.3 to 14.0.

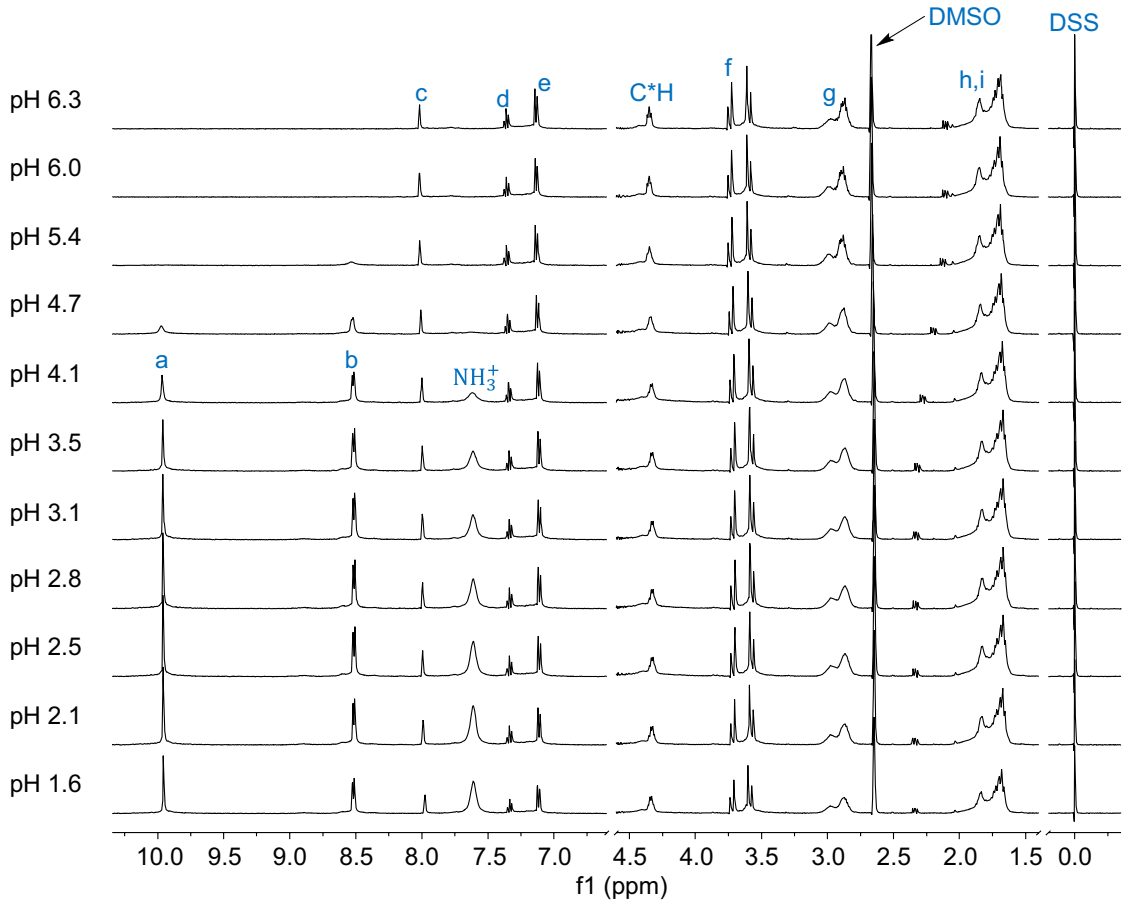


Figure 4.18. Partial ^{1}H NMR (500 MHz, "WATER_ES" water suppression pulse sequence, 298 K) spectra of 2 mM oxidized **1j** in H_2O with 25% (v/v) DMSO-d_6 at pH from 1.6 to 6.3. The assignment letters (a-h) correlate with Figure 4.10.

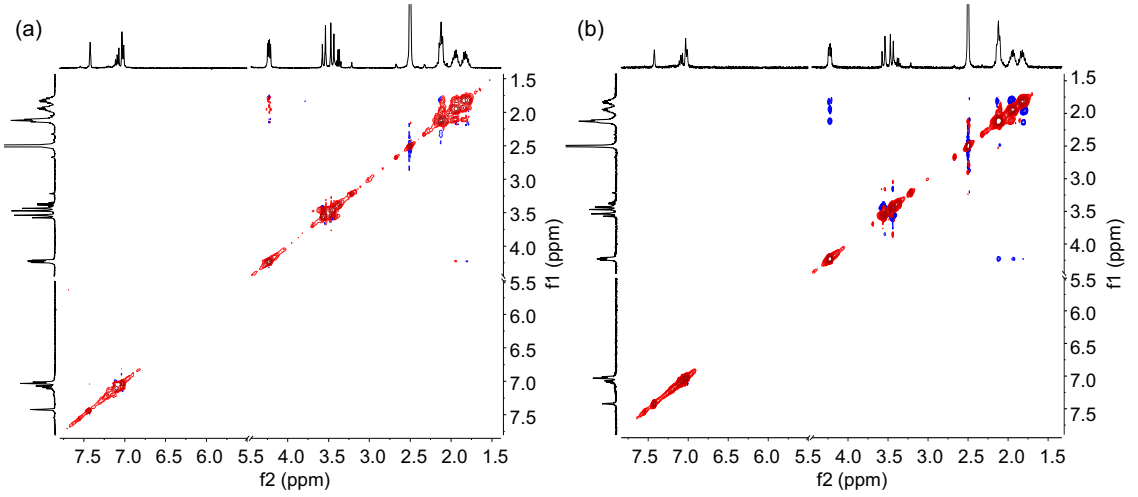


Figure 4.19. (a) 2D NOESY and (b) 2D ROESY (400 MHz, 298 K) spectra of 1 mM (**1i**)$_2$ in 20 mM deuterated phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6.

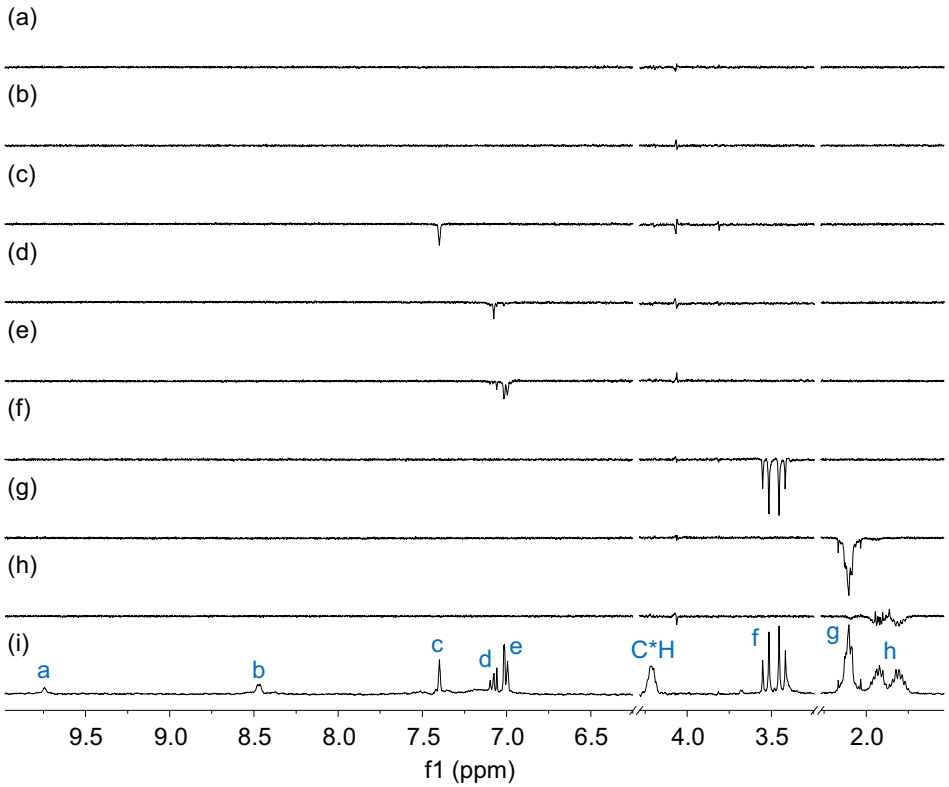


Figure 4.20. 1D ROESY (400 MHz, "WET" water suppression pulse sequence, 298 K) spectra of 1 mM $(\mathbf{1i})_2$ in 20 mM phosphate buffer (pH 7.5) with 25% (v/v) DMSO-d_6. Irradiation of signal a (a), b (b), c (c), d (d), e (e), f (f), g (g) and h (h). Partial ^{1}H NMR (400 MHz, 298 K) spectrum of the same sample (i). The assignment letters (a-h) correlate with Figure 4.10.

4.5.7. Molecular dynamics simulations

All molecular simulations were carried out with the package Schrödinger Suite 2014,[71] through its graphical interface Maestro.[72] The program Macromodel[73] with its default force field OPLS 2005, a modified version of the OPLS-AA force field,[74] and GB/SA water solvation conditions[75] were used for energy minimization of dimers $(\mathbf{1i})_2$ and $(\mathbf{1j})_2$. Molecular dynamics simulations were performed with the program Desmond 3.8[76-77] using the OPLS 2005 force field. Systems were set up using the System Builder of the Maestro-Desmond interface.[78]

Initial short (10 ns) simulations were carried out for each dimer in a completely ionized state, *i.e.* with a total charge of –4 ($(\mathbf{1i})_2$) or +4 ($(\mathbf{1j})_2$), plus 4 counter ions (Na^+ or Cl^-, respectively) in a 50 x 50 x 50 Å^{3}cubic box of TIP3P water. The systems were subjected to steepest descent minimization under periodic boundary conditions (PBC) and equilibrated for 500 ps at 300 K and 1 bar, under the NPT ensemble. Production MD simulations (10 ns, 2 fs timestep) were performed under the same conditions (PBC,

NPT ensemble, 300 K and 1.0 bar) using the Nose-Hoover thermostat method[79-80] with a relaxation time of 1.0 ps and the Martyna-Tobias-Klein barostat method[81] with isotropic coupling and a relaxation time of 2 ps. Integration was carried out with the RESPA integrator[82] using time steps of 2.0, 2.0, and 6.0 fs for the bonded van der Waals and short range and long range electrostatic interactions, respectively. A cut-off of 15 Å was applied to van der Waals and short-range electrostatic interactions, whereas long-range electrostatic interactions were computed using the smooth particle mesh Ewald method with an Ewald tolerance of 10^{-9}.[83-84] Bond lengths to hydrogen atoms were constrained using the Shake algorithm.[85]

Among the snapshots from these two preliminary MD simulations, three conformations were manually selected for each compound which covered different folding states, *i.e.* from a more folded/packed conformer to an essentially extended and an intermediate ones. These conformers were then used as starting point to run longer simulations (100 ns) in 25% (v/v) DMSO/water solvent mixture, which reflects the experimental conditions. Molecular systems contained one molecule of ionized dimer with the corresponding counter ions, immersed in a 50 x 50 x 50 $Å^3$ cubic box of a DMSO/water solvent mixture with an approximately 1:12 molar ratio. Systems were set up by first immersing the dimer molecule plus the four counter ions in a 50 x 50 x 50 $Å^3$ cubic box of DMSO, then a random removal of DMSO molecules was performed with the help of an external software until only 185 were left, and finally the whole system was resolvated with TIP3P waters (~ 2300 waters) keeping the dimensions of the cubic box. The full systems (~9000 atoms each) were minimized as above, first with the solute restrained and then without restrains until a gradient threshold of 0.1 kcal/mol/Å was reached, heated stepwise until 300 K with short MD runs (25 ps at 0.1, 10, 50, 100, 200 and 300 K), and then equilibrated for 2 ns at the same temperature and 1.0 bar in the NPT ensemble. Production MD simulations (100 ns, 2 fs timestep) were performed as already described. Coordinates were saved every 40 ps, hence 2500 snapshots (frames) were obtained from each MD run. The Simulation Event Analysis application included in the Desmond-Maestro interface was used to analyze the simulations results.

Thus, the number of intramolecular H-bonds, the radius of gyration and the distances between each pair of charged groups were calculated. Additionally, after removing the solvent molecules and ions from each frame, the 2500 conformations of the dimeric compounds $(\mathbf{1i})_2$ and $(\mathbf{1j})_2$ derived from each MD run were saved and the

water accessible areas (ASA) of the hydrophobic and polar surfaces (*i.e.* water accessible surface area of all hydrophobic ($|q_i|<0.2$) and polar ($|q_i|\geq0.2$) atoms) were calculated for each one with the program MOE.[86] Finally, the 2500 conformations of each MD simulation were clustered in 10 groups based on the atomic RMSD of the heavy atoms, by using the hierarchical clustering with average linkage method as implemented in the Conformer Cluster script included in Maestro.

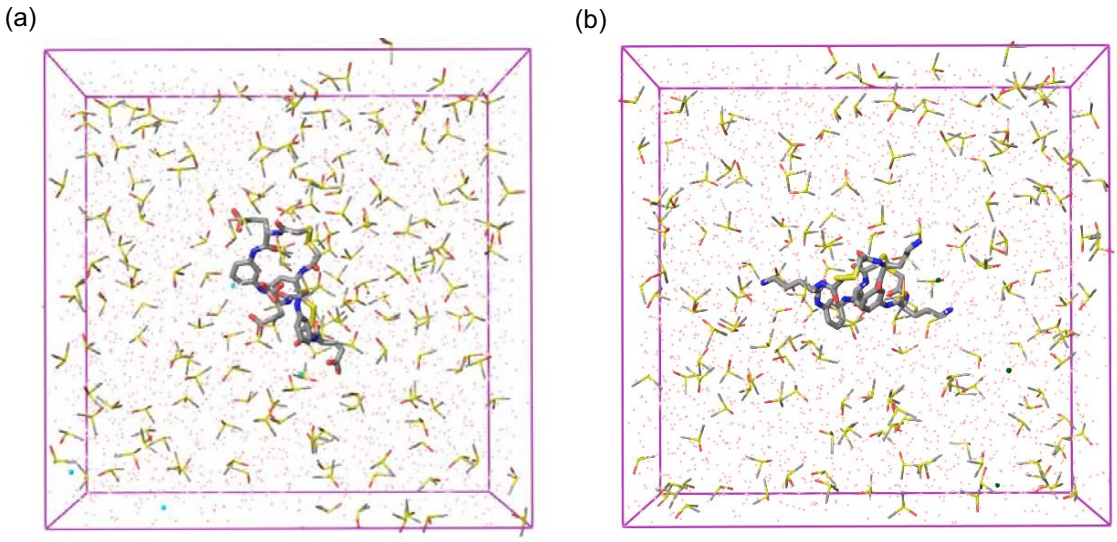


Figure 4.21. Molecular dynamics simulation systems for homodimers (**1i**)$_2$ (a) and (**1j**)$_2$ (b). Both were constituted by a 50 x 50 x 50 Å^3 cubic box containing one molecule of each homodimer, 4 counterions (Na^+, cyan or Cl^-, green), 185 DMSO molecules and ~2300 water molecules (red dots).

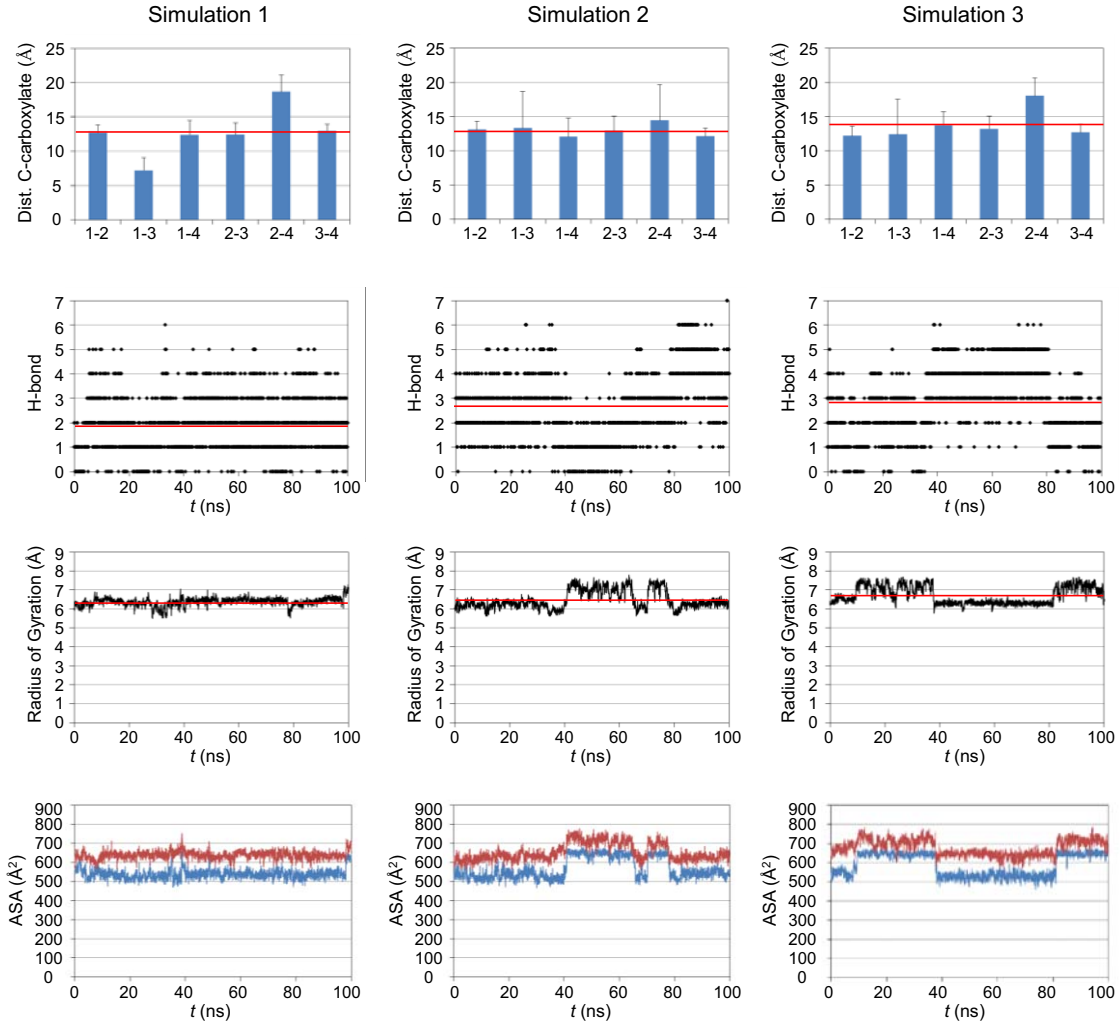


Figure 4.22. Summary of results obtained from the three 100 ns MD simulations carried out with homodimer (**1i**)$_2$. The graphics represent the distances between the carbon atoms of the four carboxylates averaged over the whole simulation, as well as the variation of the number of intramolecular hydrogen bonds, the radius of gyration and the water accessible areas (ASA) of the polar (red) and hydrophobic (blue) surfaces against simulation time (ns). The horizontal red lines on the above three graphs represent the corresponding global averaged values. Average C-carboxylate distances (Å): 12.7 (Sim. 1), 13.0 (Sim. 2) and 13.7 (Sim. 3). Average number of H-Bonds: 1.8 (Sim. 1), 2.6 (Sim. 2) and 2.8 (Sim. 3). Average radius of gyration (Å): 6.4 (Sim. 1), 6.5 (Sim. 2) and 6.7 (Sim. 3).

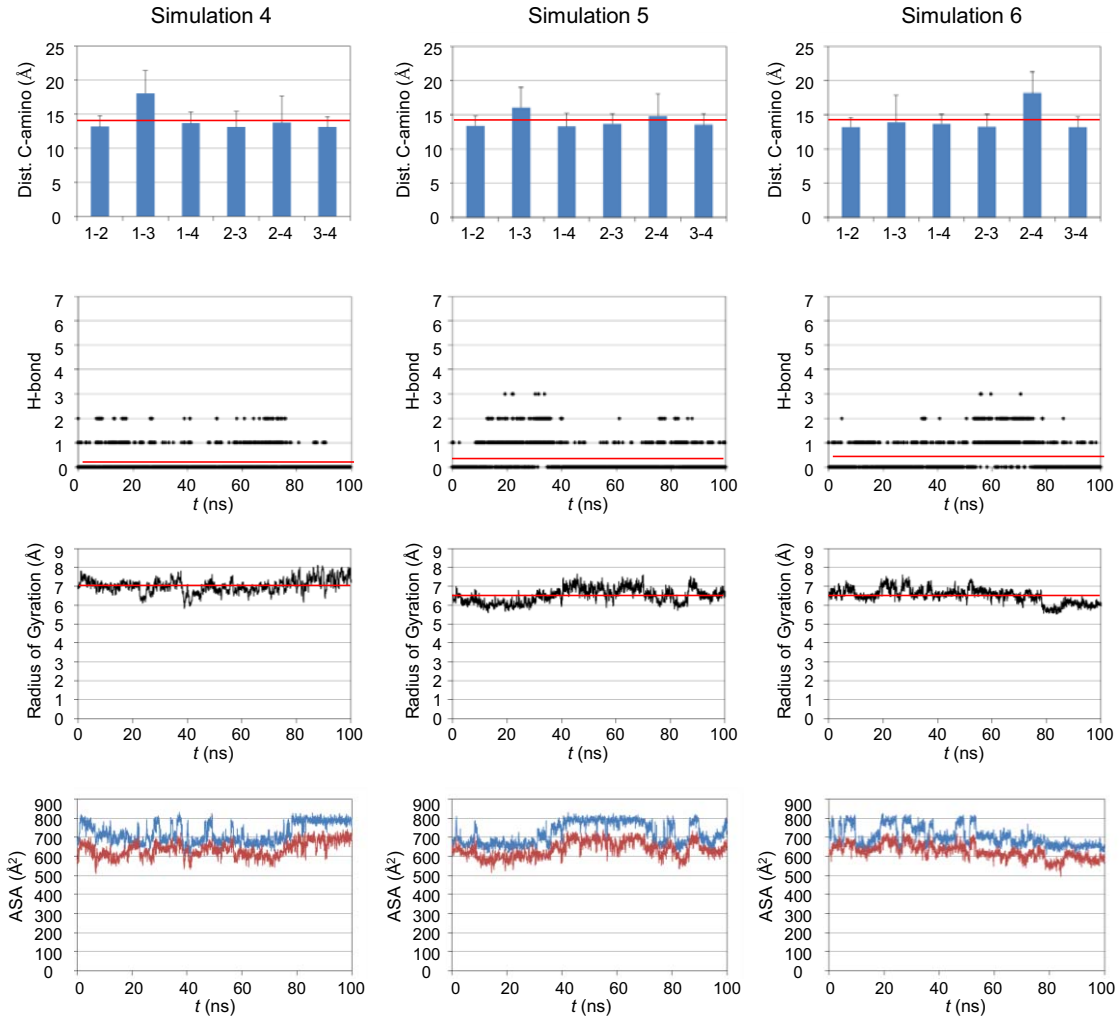


Figure 4.23. Summary of results obtained from the three 100 ns MD simulations carried out with homodimer $(\mathbf{1j})_2$. The graphics represent the distances between the carbon atoms that hold the four protonated amino groups averaged over the whole simulation, as well as the variation of the number of intramolecular hydrogen bonds, the radius of gyration and the water accessible areas (ASA) of the polar (red) and hydrophobic (blue) surfaces against simulation time (ns). The horizontal red lines on the above three graphs represent the corresponding global averaged values. Average C-amino distances (Å): 14.1 (Sim. 4), 14.1 (Sim. 5) and 14.2 (Sim. 6). Average number of H-Bonds: 0.19 (Sim. 4), 0.35 (Sim. 5) and 0.44 (Sim. 6). Average radius of gyration (Å): 7.0 (Sim. 4), 6.5 (Sim. 5) and 6.5 (Sim. 6).

4.6. References

(1) Liang, E.; Zhou, H.; Ding, X.; Zheng, Z.; Peng, Y. *Chem. Commun.* **2013**, *49*, 5384.
(2) Giuseppone, N.; Lehn, J.-M. *Chem. Eur. J.* **2006**, *12*, 1715.
(3) Giuseppone, N.; Fuks, G.; Lehn, J.-M. *Chem. Eur. J.* **2006**, *12*, 1723.
(4) Giuseppone, N.; Lehn, J.-M. *J. Am. Chem. Soc.* **2004**, *126*, 11448.
(5) Polyakov, V. A.; Nelen, M. I.; Nazarpack-Kandlousy, N.; Ryabov, A. D.; Eliseev, A. V. *J. Phys. Org. Chem.* **1999**, *12*, 357.
(6) Kovaříček, P.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 9446.
(7) Hafezi, N.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 12861.
(8) Au-Yeung, H. Y.; Pantoş, G. D.; Sanders, J. K. M. *Angew. Chem. Int. Ed.* **2010**, *49*, 5331.
(9) Reutenauer, P.; Boul, P. J.; Lehn, J.-M. *Eur. J. Org. Chem.* **2009**, *2009*, 1691.
(10) Vantomme, G.; Jiang, S.; Lehn, J.-M. *J. Am. Chem. Soc.* **2014**, *136*, 9509.
(11) Vantomme, G.; Lehn, J.-M. *Angew. Chem. Int. Ed.* **2013**, *52*, 3940.
(12) Di Stefano, S.; Mazzonna, M.; Bodo, E.; Mandolini, L.; Lanzalunga, O. *Org. Lett.* **2011**, *13*, 142.
(13) Chaur, M. N.; Collado, D.; Lehn, J.-M. *Chem. Eur. J.* **2011**, *17*, 248.
(14) Maeda, Y.; Javid, N.; Duncan, K.; Birchall, L.; Gibson, K. F.; Cannon, D.; Kanetsuki, Y.; Knapp, C.; Tuttle, T.; Ulijn, R. V.; Matsui, H. *J. Am. Chem. Soc.* **2014**, *136*, 15893.
(15) Nalluri, S. K. M.; Ulijn, R. V. *Chem. Sci.* **2013**, *4*, 3699.
(16) Wang, G.-T.; Lin, J.-B.; Jiang, X.-K.; Li, Z.-T. *Langmuir* **2009**, *25*, 8414.
(17) Buhler, E.; Sreenivasachary, N.; Candau, S.-J.; Lehn, J.-M. *J. Am. Chem. Soc.* **2007**, *129*, 10058.
(18) Sreenivasachary, N.; Lehn, J.-M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5938.
(19) Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S. *Science* **2010**, *327*, 1502.
(20) Herrmann, A.; Giuseppone, N.; Lehn, J.-M. *Chem. Eur. J.* **2009**, *15*, 117.
(21) Giuseppone, N.; Lehn, J.-M. *Angew. Chem. Int. Ed.* **2006**, *45*, 4619.
(22) Jadhav , K. B.; Lichtenecker , R. J.; Bullach, A.; Mandal, B.; Arndt, H.-D. *Chem. Eur. J.* **2015**, *21*, 5898.
(23) Ponnuswamy, N.; Cougnon, F. B. L.; Pantoş, G. D.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2014**, *136*, 8243.
(24) Ponnuswamy, N.; Cougnon, F. B. L.; Clough, J. M.; Pantoş, G. D.; Sanders, J. K. M. *Science* **2012**, *338*, 783.
(25) Cougnon, F. B. L.; Ponnuswamy, N.; Jenkins, N. A.; Pantoş, G. D.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2012**, *134*, 19129.
(26) Cougnon, F. B. L.; Jenkins, N. A.; Pantoş, G. D.; Sanders, J. K. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 1443.
(27) Cougnon, F. B. L.; Au-Yeung, H. Y.; Pantoş, G. D.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2011**, *133*, 3198.
(28) Au-Yeung, H. Y.; Pantoş, G. D.; Sanders, J. K. M. *J. Org. Chem.* **2011**, *76*, 1257.
(29) Au-Yeung, H. Y.; Cougnon, F. B. L.; Otto, S.; Pantos, G. D.; Sanders, J. K. M. *Chem. Sci.* **2010**, *1*, 567.
(30) Au-Yeung, H. Y.; Pantoş, G. D.; Sanders, J. K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10466.
(31) Au-Yeung, H. Y.; Dan Pantoş, G.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2009**, *131*, 16030.

(32) Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. *Science* **2005**, *308*, 667.
(33) Zangi, R.; Hagen, M.; Berne, B. J. *J. Am. Chem. Soc.* **2007**, *129*, 4678.
(34) Ludlow, R. F.; Otto, S. *J. Am. Chem. Soc.* **2008**, *130*, 12218.
(35) McNaughton, B. R.; Gareiss, P. C.; Miller, B. L. *J. Am. Chem. Soc.* **2007**, *129*, 11306.
(36) Otto, S.; Kubik, S. *J. Am. Chem. Soc.* **2003**, *125*, 7804.
(37) Hochgürtel, M.; Kroth, H.; Piecha, D.; Hofmann, M. W.; Nicolau, C.; Krause, S.; Schaaf, O.; Sonnenmoser, G.; Eliseev, A. V. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3382.
(38) McNaughton, B. R.; Miller, B. L. *Org. Lett.* **2006**, *8*, 1803.
(39) Solà, J.; Lafuente, M.; Atcher, J.; Alfonso, I. *Chem. Commun.* **2014**, *50*, 4564.
(40) Stefankiewicz, A. R.; Sanders, J. K. M. *Chem. Commun.* **2013**, *49*, 5820.
(41) Stefankiewicz, A. R.; Sambrook, M. R.; Sanders, J. K. M. *Chem. Sci.* **2012**, *3*, 2326.
(42) West, K. R.; Bake, K. D.; Otto, S. *Org. Lett.* **2005**, *7*, 2615.
(43) Ulatowski, F.; Sadowska-Kuziola, A.; Jurczak, J. *J. Org. Chem.* **2014**, *79*, 9762.
(44) Ludlow, R. F.; Otto, S. *J. Am. Chem. Soc.* **2010**, *132*, 5984.
(45) Belowich, M. E.; Valente, C.; Smaldone, R. A.; Friedman, D. C.; Thiel, J.; Cronin, L.; Stoddart, J. F. *J. Am. Chem. Soc.* **2012**, *134*, 5243.
(46) Belowich, M. E.; Valente, C.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2010**, *49*, 7208.
(47) Chichak, K. S.; Cantrill, S. J.; Pease, A. R.; Chiu, S.-H.; Cave, G. W.; Atwood, J. L.; Stoddart, J. F. *Science* **2004**, *304*, 1308.
(48) Pentecost, C. D.; Chichak, K. S.; Peters, A. J.; Cave, G. W. V.; Cantrill, S. J.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2007**, *46*, 218.
(49) Ayme, J.-F.; Beves, J. E.; Leigh, D. A.; McBurney, R. T.; Rissanen, K.; Schultz, D. *Nat. Chem.* **2012**, *4*, 15.
(50) Du, G.; Moulin, E.; Jouault, N.; Buhler, E.; Giuseppone, N. *Angew. Chem. Int. Ed.* **2012**, *51*, 12504.
(51) Bozdemir, O. A.; Barin, G.; Belowich, M. E.; Basuray, A. N.; Beuerle, F.; Stoddart, J. F. *Chem. Commun.* **2012**, *48*, 10401.
(52) Lehn, J.-M. *Angew. Chem. Int. Ed.* **2013**, *52*, 2836.
(53) Mukerjee, P.; Ostrow, J. D. *Tetrahedron Lett.* **1998**, *39*, 423.
(54) Bailey, W. F.; Monahan, A. S. *J. Chem. Educ.* **1978**, *55*, 489.
(55) Ramström, O.; Lohmann, S.; Bunyapaiboonsri, T.; Lehn, J.-M. *Chem. Eur. J.* **2004**, *10*, 1711.
(56) Bunyapaiboonsri, T.; Ramström, O.; Lohmann, S.; Lehn, J.-M.; Peng, L.; Goeldner, M. *ChemBioChem* **2001**, *2*, 438.
(57) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Chem. Eur. J.* **2008**, *14*, 2153.
(58) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Angew. Chem. Int. Ed.* **2007**, *46*, 8858.
(59) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *J. Am. Chem. Soc.* **2005**, *127*, 9390.
(60) Corbett, P. T.; Otto, S.; Sanders, J. K. M. *Chem. Eur. J.* **2004**, *10*, 3139.
(61) Hamieh, S.; Saggiomo, V.; Nowak, P.; Mattia, E.; Ludlow, R. F.; Otto, S. *Angew. Chem. Int. Ed.* **2013**, *52*, 12368.
(62) Sukenik, S.; Boyarski, Y.; Harries, D. *Chem. Commun.* **2014**, *50*, 8193.
(63) Scheu, R.; Rankin, B. M.; Chen, Y.; Jena, K. C.; Ben-Amotz, D.; Roke, S. *Angew. Chem. Int. Ed.* **2014**, *53*, 9560.
(64) Li, J.; Nowak, P.; Fanlo-Virgos, H.; Otto, S. *Chem. Sci.* **2014**, *5*, 4968.

(65) Dadon, Z.; Wagner, N.; Alasibi, S.; Samiappan, M.; Mukherjee, R.; Ashkenasy, G. *Chem. Eur. J.* **2014**, n/a.
(66) Ashkenasy, G.; Jagasia, R.; Yadav, M.; Ghadiri, M. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10872.
(67) Lee, D. H.; Severin, K.; Yokobayashi, Y.; Ghadiri, M. R. *Nature* **1997**, *390*, 591.
(68) Bicking, M. K. *LG GC N. Am.* **2006**, *24*, 402.
(69) Hwang, T. L.; Shaka, A. J. *J. Magn. Reson., Ser A* **1995**, *112*, 275.
(70) Ogg, R. J.; Kingsley, R. B.; Taylor, J. S. *J. Magn. Reson., Ser B* **1994**, *104*, 1.
(71) Schrödinger *Suite 2014 Update 3*; Schrödinger, LLC: New York, NY, 2014.
(72) Schrödinger *Maestro, version 9.9*; Schrödinger, LLC: New York, NY, 2014.
(73) Schrödinger *MacroModel, version 10.5*; Schrödinger, LLC: New York, NY, 2014.
(74) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
(75) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(76) Research, D. E. S. *Desmond Molecular Dynamics System, version 3.8*, New York, NY, 2014.
(77) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossváry, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. In *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)* Tampa, Florida, 2006.
(78) Schrödinger *Maestro-Desmond Interoperability Tools, version 3.8*; Schrödinger, LLC: New York, NY, 2014.
(79) Evans, D. J.; Holian, B. L. *J. Chem. Phys.* **1985**, *83*, 4069.
(80) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992**, *97*, 2635.
(81) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.
(82) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990.
(83) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(84) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(85) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Comput. Chem.* **2001**, *22*, 501.
(86) MOE; v. 2013.08 ed.; Chemical Computing Group Inc.: Montreal, QC, Canada, 2013.

CHAPTER 5

Homochiral self-sorting of dynamic macrocyclic pseudopeptides

5.1. Introduction

5.1.1. Biological homochirality

Chiral molecules in living organisms exist almost exclusively as single enantiomers, a property that is crucial for molecular recognition and replication processes.[1] Thus, biomolecules such as proteins, RNA, DNA and polysaccharides are made of monomers of only one handedness: L-amino acids and D-sugars. This chiral homogeneity of biomolecules, called homochirality, is a property of all living organisms and is critical for many bio-molecular processes. As an example, the activity of a protein is dependent on its spatial structure, and the functional specificity of its superstructure is determined by the homochirality of the monomeric constituent amino acids. Hence, homochirality is essential to life as we know it.

Models for the origin of biological homochirality have been proposed and reviewed,[2-7] but the appearance of natural biopolymers wholly consisting of homochiral subunits from the unanimated racemic prebiotic world is still unsolved. There are two hypotheses concerting the sequence of their emergence. The first one maintains that long homochiral biopolymers must have been formed after the appearance of the first living systems. The so-called *biotic theories* are by their very nature difficult to prove because of the uncertainties surrounding the origin of life and the early earth environments that fostered it.[8-11] The second hypothesis presumes that biopolymers preceded the primeval forms of life. The so-called *abiotic theories* have either to relay on the exogenous delivery of enantiomerically enriched meteorites to the early earth[12-15] or to postulate means for both symmetry breaking and the amplification of the initial imbalance. Theoretical investigations have been concerned with how an initial imbalance between left- and right-handed molecules might have been formed initially and then amplified.[16-18] Experimental investigations have yielded several plausible mechanisms for amplification of enantiomeric excess.[19-24]

5.1.2. Precedents of stereoselectivity in DCC

5.1.2.1. Template-induced stereoselectivity

Templating a racemic DCL with an enantiopure guest would amplify the best host, and if that receptor also bound the analyte enantioselectively, then the best matched enantiomer would be amplified over the mismatched antipode. Gagné and co-workers

explored this concept with the racemic pseudopeptide depicted in Figure 5.1. The addition of TFA to a solution of this compound deprotects the acetal and initiates reversible hydrazone exchange to form the corresponding cyclic oligomers. After the introduction of (–)-adenosine as a template, they observed the amplification of the dimer and, more importantly, this had one of its homochiral diasteromers selectively enhanced over the others.[a,25] Enantioselective responses of this DCL were also reported for (–)-cytidine (example depicted in Figure 5.1) and (–)-2-thiocytidine.[26]

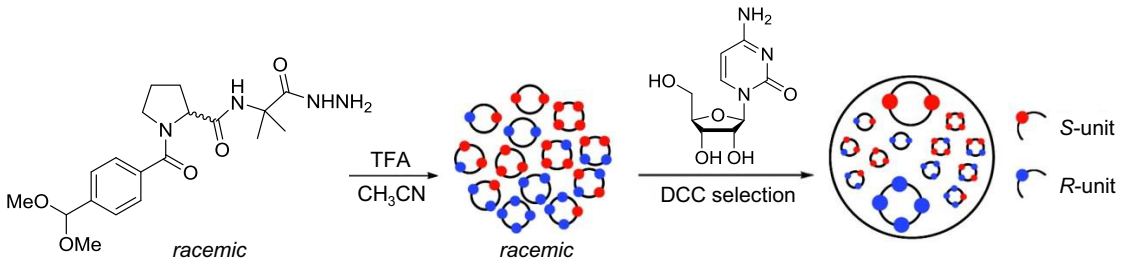


Figure 5.1. Deracemization of a DCL generated from a racemic BB by the introduction of (–)-cytidine (Figure modified from reference[26]).

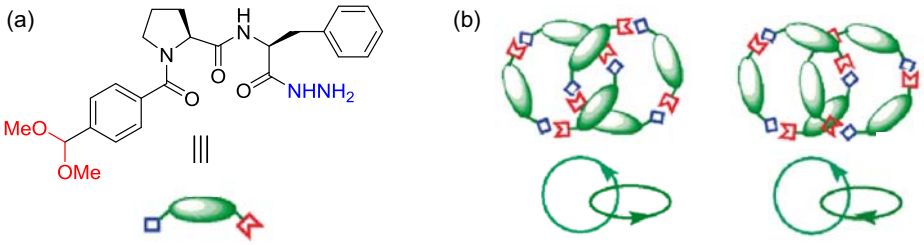


Figure 5.2. (a) Building block used by the group of Sanders for the preparation of DCLs through hydrazone reversible linkage. (b) Cartoon representation of the two diastereoisomeric [2]-catenanes (Figure modified from reference[27]).

Sanders and co-workers reported a DCL generated from an enantiomerically pure building block with a very similar chemical structure (Figure 5.2a). The exposure of the resulting library of hydrazones to diastereomeric cinchona alkaloids led to the amplification of two different macrocyclic receptors.[28] Very remarkably, the same authors also used this molecular network to show that diastereoselective responses can be induced by the addition of achiral templates.[27] Thus, the addition of acetylcoline induced the amplification of a [2]-catenane oligomer consisting of six building blocks assembled into two interlocked 42-membered rings. Because each ring is chiral, there are two possible diasteromers of the [2]-catenane (Figure 5.2b). The relative simplicity

[a] For the identification and quantification of the generated stereoisomers, the authors developed an *in situ* method consisting in the selective isotopic labeling of one of the two enantiomeric forms of the building block represented in Figure 5.1. The resulting unique mass signature of the diastereomeric library components enabled MS analysis for diastereo- and enantiocomposition assessment.

of the NMR spectrum of the catenane-acetylcoline complex (1:1) indicated that only one of the diasteromers was present.

Otto and co-workers also investigated the use of achiral guests to induce a diastereoselective templating effect.[29] They generated a disulfide-based DCL of oligomeric macrocycles in water by air-oxidation of the racemic dithiol depicted in Figure 5.3a. The addition of tetramethylammonium iodide as a template induced the efficient amplification of a tetrameric receptor. Since the starting dithiol is racemic, various stereoisomeric tetramers could be formed. The authors achieved baseline separation of all four diastereomeric products by HPLC, observing that the amplification was highly diastereoselective. The structure of the diasteromers was assigned using NMR and structural correlation through re-equilibration, and the amplified species was found to be the alternated *meso*-tetramer shown in figure 5.3a. The selectivity was attributed to a four-stave barrel folding of the *meso*-tetramer, enclosing a cavity ideally sized to accommodate the tetramethylammonium cation (Figure 5.3b).

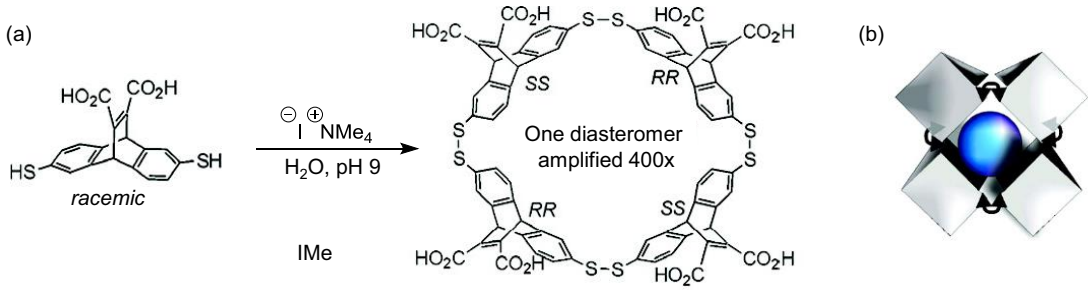


Figure 5.3. (a) Representation of the racemic dithiol building block, and the *meso*-tetramer disulfide amplified by the addition of tetramethylammonium cation as a template. (b) Cartoon representation of the four-stave barrel folding rationalizing the observed diastereoselectivity (Figure modified from reference[29]).

Metal cations have also been explored as achiral templates to induce stereoselective responses within imine-bases DCLs. Thus, Gotor and co-workers reported a remarkable diastereoselective amplification of a single member from a DCL of stereoisomeric macrocyclic imines.[30] The addition of Cd^{2+} ion induced the exclusive formation of a racemic heterochiral macrocyclic trimer, whereas the corresponding homo- counterparts were not detected. The Cd^{2+} amplification was completely diastereoselective as only one diasteromer was formed in the presence of the template. Very recently, Ziach and Jurczak reported the addition of Li^{+} to an imine-bases DCL of macrocyclic achiral oligomers.[31] The addition of the template induced the formation of crystalline

complexes, and these were found to be composed of a single enantiomeric form of the complex.

5.1.2.2. Chiral self-sorting

Self-sorting can be defined as high fidelity recognition between chemical species within complex mixtures.[32] Accordingly, chiral self-sorting refers to the same concept but applied to mixtures of stereoisomers. The chiral recognition between enantiomeric pairs can lead to self-recognition or self-discrimination,[33] depending on whether an enantiomer preferentially recognizes itself or its mirror image, to generate homo- or heterochiral species respectively.

Chiral self-sorting of a racemic mixture is a known phenomenon most commonly observed in the solid state,[34-37] being the spontaneous resolution of enantiomers a well-documented example since more than 150 years ago.[38] In recent years, several successful efforts to translate this self-sorting concept to mixtures in solution have been reported. However, most of these are based on non-covalent interactions such as metal coordination,[33,39,40] H-bonding,[33,41-43] π-π stacking[33] and van der Waals.[41] In this regard, DCvC has been proposed to enable chiral self-sorting *via* covalent bond formation. The following lines briefly comment on some of the most remarkable examples.

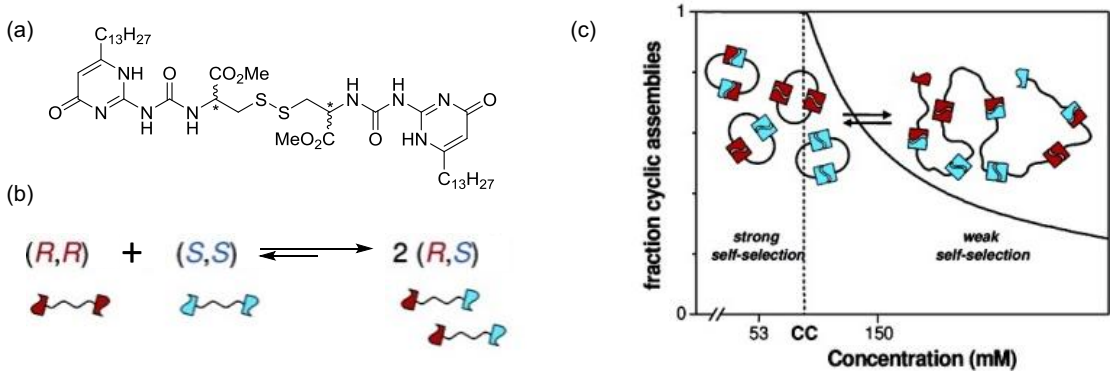


Figure 5.4. (a) Chemical structure of the 2-ureido-4[1*H*]-pyrimidinone racemic derivative. (b) Cartoon representation of the equilibrium between diastereomeric dimers in disulfide exchange reaction. (c) Correlation between self-selection and concentration (Figure modified from reference[44]).

Meijer and co-workers reported an interesting example of stereoselective self-selection by double dynamic chemistry.[44,45] Mixtures of the racemic disulfide depicted in Figure 5.4a were produced by simultaneous disulfide exchange and dimerization of the 2-ureido-4[1*H*]-pyrimidinone group by quadrupole hydrogen bonding. Equilibration

of the two different types of reversible bonds (disulfides and H-bonds) resulted in enrichment of the *meso*-isomer ((*R*,*S*) configuration, Figure 5.4b) because the formation of cyclic assemblies is stereoselective and the thermodynamic stability of these assemblies is different. Additionally, the diastereoselectivity was found to depend on the concentration of the building blocks (Figure 5.4c).

The use of DCvC has also been explored for the stereoselective preparation of large cavitands and nanocapsules. Thus, Warmuth *et al.* described the dynamic covalent synthesis of polyimide-based chiral nanocapsules.[46,47] More recently, Szumna and co-workers reported the impressive synthesis of self-assembled capsules based on short peptides of various chiralities. The approach of the authors combines dynamic covalent synthesis (reversible imine bond formation) with self-assembly. They found that it is possible to induce self-organization of complex nanosized chiral molecular containers through the use of very short peptidic sequences and a macrocyclic skeleton that provides additional preorganization.[48]

There are only two reported examples of macrocyclization *via* DCvC with inherent selectivity for homochiral products in solution. In a first example, Tilley and co-workers reported an efficient and high-yield synthetic route to multigram quantities of chiral BINOL-containing macrocycles using zirconocene coupling.[49] However, no rationalization was provided to explain the observed homochiral selectivity. A more in-depth energy analysis was later reported by Sisco and Moore.[50] In this second example, BINOL-based building blocks were also used to generate macrocyclic dimers by means of the alkyne metathesis reversible chemistry (Figure 5.5a). The authors observed that the inherent symmetry differences between the homo- and heterochiral dimers (Figure 5.5b) can be a useful attribute to achieve homochiral self-sorting. Thus, the macrocyclization process was found to be entropy-driven due to increased symmetry of the homochiral products. Consequently, the increase in temperature favored the homochiral prevalence in the final product distribution.

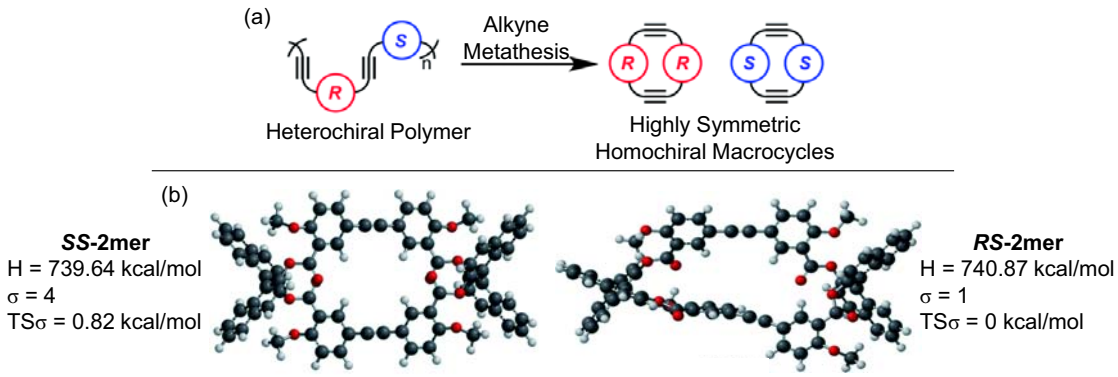


Figure 5.5. Symmetry-driven macrocyclization *via* DCvC. (b) Energy minimized structures of *SS*-2mer and *RS*-2mer (Figure modified from reference[50]).

The two reported examples of macrocyclization *via* DCvC with inherent selectivity for homochiral products are all based on rigid BINOL-containing molecules dissolved in purely organic solvents. To the best of our knowledge, there are no precedents of such a homochiral self-sorting phenomenon with highly flexible and/or bio-inspired macrocycles in aqueous media.

5.2. Objectives and hypothesis

The main objective of the present Chapter is to investigate how the chiral information of the building blocks is transmitted in a simple DCL of flexible pseudopeptidic compounds. In this regard, we hypothesized that the inherent structural differences between diastereomeric library members would lead to the emergence of chiral self-sorting processes.

This main objective can be divided into the following three specific aims:

i) To design and prepare a minimalistic DCL of stereoisomeric pseudopeptides as a tool to investigate the potential emergence of chiral self-sorting phenomena.

ii) To evaluate the decrease in the polarity of the medium as an environmental factor to induce changes in the composition of the dynamic library by favoring the polar intramolecular interactions.

iii) To study the effect of three additional variables (temperature, pH and concentration) in order to better characterize and understand the behavior of the whole system.

5.3. Results and discussion

5.3.1. Motivation for the present study

The actual origin of the homochirality is still a matter of debate, but its observed preservation necessarily implies the accurate transmission of the stereochemical information. This process requires an efficient homochiral self-selection of the implicated species. Fascinated by this idea, we decided to investigate how the chiral information could be transmitted in simple DCLs under thermodynamic equilibrium.

In the previous Chapter of this thesis we observed that, in the frame of a DCL, compounds with different structures (based on different amino acid moieties and having different charges) can show notably different responses to the increase of the ionic strength. In the present study we wondered whether practically identical stereoisomeric compounds could also show different responses to an alike external stimulus. In this regard, the aim for chiral self-sorting within a DCL of flexible and bio-inspired pseudopeptidic BBs is an extraordinarily ambitious challenge.

5.3.2. The building blocks

Once again, for the simplicity of analysis and quantitative evaluation, we envisioned a minimalistic system generated by the combination of only two building blocks, *i.e.* a binary mixture. For this small dynamic library to be used as a model system to study the transmission of chiral information within a DCL, the two building blocks were designed to be mirror images. Thus, the combination of only two enantiomeric building blocks was conceived as the simplest way to generate a molecular network with the potential ability to show chiral self-sorting phenomena.

The study was focalized on the use of negatively charged species. Once dissolved in an oxidizing medium, their subsequent covalent union would presumably lead to the formation of dimers classified in Chapter 4 as members of family $F_{(-,-)}$. These have interestingly shown to prominently change their relative stability depending on the distance between the charges and the polarity of the medium. Taking these precedents into account, an alike polarity-related factor was envisioned as a potential driving force behind the aimed chiral self-sorting. It is worth mentioning that also the members of family $F_{(+,+)}$ proved to show this behavior. However, for the members of this family the stability changes are limited by the longer average distance between charges.

Additionally, one would expect that the prevalence of a more packed/folded conformation for the members of family $F_{(-,-)}$ would make them more prone to show structural and energetic differences between stereoisomers. Their remarkable ability to form intramolecular H-bonds is foreseen to depend on the orientation of the side chains and, ultimately, on the relative configuration of the chiral centers.

In order to be composed by members of family $F_{(-,-)}$, and considering the BBs synthesized in Chapter 1, the binary mixture could be generated by the combination of either **1h** or **1i** with its corresponding enantiomers. The Glu-based dithiol **1i** was preferred mainly because, unlike the Asp-based dithiol **1h**, it does not undergo the undesired imide formation when exposed to acidic media for prolonged periods of time (section 2.3.6).[51-54] The mirror image of **1i** was synthesized as explained in Chapter 2 with its four $C\underline{H}_{Ar}$ replaced by deuterium atoms (Figure 5.6). The resulting unique mass signature of the library components was expected to enable MS analysis for diastereo- and enantiocomposition assessment. Even in the case of not being able to chromatographically separate the compounds of the library, the isotopic labeling should allow a direct MS quantitative evaluation. The isotopic labeling strategy for MS identification of the stereoisomeric library members has previously been used by Gagné and co-workers (section 5.1.2.1).[25]

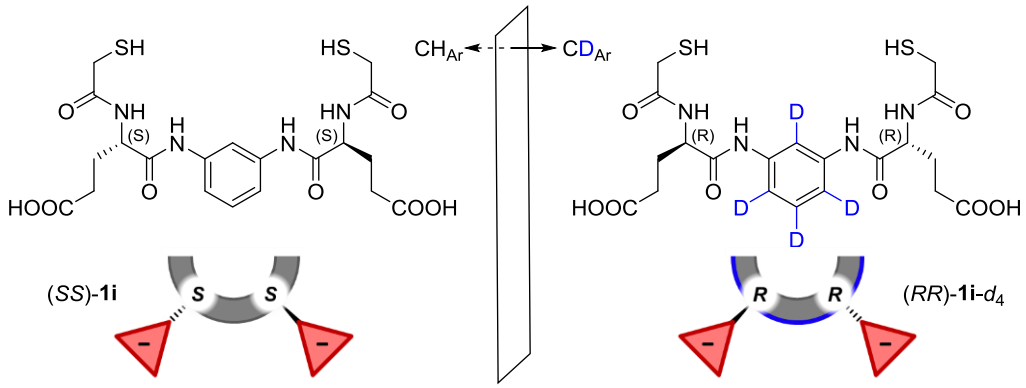


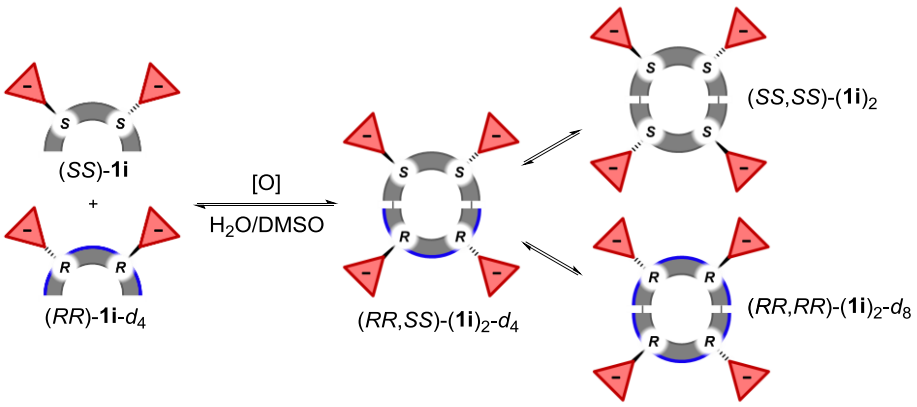
Figure 5.6. Representation of the two pseudo-enantiomeric BBs (*SS*)-**1i** and (*RR*)-**1i**-d_4.

Rigorously speaking, because of the isotopic labeling, the two compounds represented in Figure 5.6 are no longer exact mirror images but pseudo-enantiomers. It is important to notice that the approximation of using pseudo-enantiomeric BBs is only valid if the CH_{Ar}/CD_{Ar} replacement has no impact on the stability of the library members. In this case, the slightly stronger acidity[55] of CH_{Ar} compared to CD_{Ar} definitely does not translate into appreciable structural or energetic differences. From

now on, the two pseudo-enantiomers of **1i** are specified in this Chapter by explicit mention of the configuration and deuterium atoms content in the following way: (*SS*)-**1i** and (*RR*)-**1i**-d_4.

5.3.3. The molecular network

To first identify and characterize the disulfides produced when mixing the two pseudo-enantiomeric forms of BB **1i**, the experimental conditions used in Chapter 3 were implemented. Thus, dithiols (*SS*)-**1i** and (*RR*)-**1i**-d_4 were mixed at 2.0 mM each in 20 mM aqueous phosphate buffer (pH 7.5)[b] with 25% (v/v) DMSO (Scheme 5.1). Simultaneously, the two BBs were separately oxidized in exactly the same experimental conditions. After 24 hours at room temperature, the libraries were analyzed by achiral-HPLC-UV (Figure 5.7) and achiral-UPLC-MS.



Scheme 5.1. Representation of the major members of the DCL generated from the mixture of (*SS*)-**1i** and (*RR*)-**1i**-d_4.

The HPLC trace of the equimolar binary mixture (Figure 5.7a) shows the prevalence of two major peaks. The first one, at the retention time of 30.8 min, corresponds to the dimer joining two BBs with opposed chirality: the pseudo-C_{2h}-symmetric heterochiral dimer (*RR*,*SS*)-(**1i**)$_2$-d_4. The second one, at the retention time of 31.5 min, corresponds to the simultaneous elution of the two possible pseudo-enantiomeric dimers joining identical BBs: the D_2-symmetrical homochiral dimers (*SS*,*SS*)-(**1i**)$_2$ and (*RR*,*RR*)-(**1i**)$_2$-d_8.[c] Finally, the minor broad peak at the retention time of 34.7-35.4 min corresponds to the overlap of the four possible trimers (*SS*,*SS*,*SS*)-(**1i**)$_3$, (*RR*,*SS*,*SS*)-(**1i**)$_3$-d_4,

[b] At this pH the carboxylic acid groups of the side chains are fully deprotonated.
[c] The "homo- and heterochiral" designation alludes to the relative configuration of the two joined BBs. The two homochiral dimers are chiral compounds, whereas the heterochiral dimer is a non-optically active pseudo-*meso* compound.

(*RR*,*RR*,*SS*)-(**1i**)$_3$-d_8 and (*RR*,*RR*,*RR*)-(**1i**)$_3$-d_{12}. In summary, as expected, the library mainly consists of the three possible macrocyclic dimers, together with a small amount of the four possible macrocyclic trimers. We decided to focus our study on the predominant dimeric species (*SS*,*SS*)-(**1i**)$_2$, (*RR*,*SS*)-(**1i**)$_2$-d_4 and (*RR*,*RR*)-(**1i**)$_2$-d_8, thus only considering the three dimers represented in Scheme 5.1.

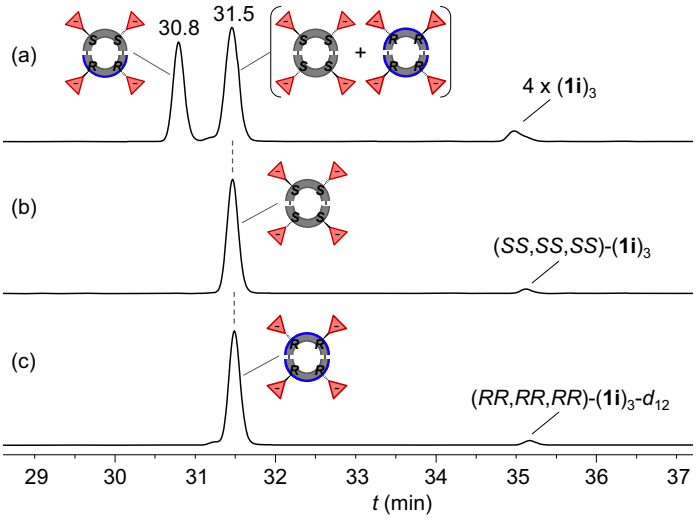


Figure 5.7. Achiral-HPLC-UV traces (254 nm) of the DCLs generated from (a) 2.0 mM (*SS*)-**1i** + 2.0 mM (*RR*)-**1i**-d_4, (b) 2.0 mM (*SS*)-**1i**, and (c) 2.0 mM (*RR*)-**1i**-d_4; in 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO.

Regarding the experimental evidences supporting the assignment of the principal members of the library, the MS analysis provided unambiguous proof. The separately oxidized BBs (*SS*)-**1i** and (*RR*)-**1i**-d_4 rendered the isotopic pattern of (*SS*,*SS*)-(**1i**)$_2$ and (*RR*,*RR*)-(**1i**)$_2$-d_8 respectively, being the latter eight m/z units heavier because of the isotopic labeling (Figures 5.8a and 5.8b).[d] By comparison with the isotopic pattern of these two homochiral species, the identification of the members of the binary mixture was quite straightforward: the MS analysis of the heterochiral dimer (*RR*,*SS*)-(**1i**)$_2$-d_4 formally consists in the "averaged" isotopic profiles of (*SS*,*SS*)-(**1i**)$_2$ and (*RR*,*RR*)-(**1i**)$_2$-d_8 (Figure 5.8c), whereas the MS analysis of the peak containing the two simultaneously eluted homochiral dimers ((*SS*,*SS*)-(**1i**)$_2$ + (*RR*,*RR*)-(**1i**)$_2$-d_8) consists in their superposed isotopic profiles (Figure 5.8d). Additionally, the two homochiral dimers in Figures 5.7b and 5.7c show the same retention time as the second peak in the HPLC trace of the binary mixture (Figure 5.7a), in agreement with their assignment.

[d] Due to partial D/H exchange during the synthesis of *m*-phenylenediamine-d_4 (Chapter 1, section 1.3.2.6), the most intense peak in the isotopic pattern of (*RR*,*RR*)-**1i$_2$**-d_8 (Figure 5.8b) does not correspond to the octadeuterated form, and the intensities distribution of the (*SS*,*SS*)-**1i$_2$** and (*RR*,*RR*)-**1i$_2$**-d_8 isotopic profiles are not comparable.

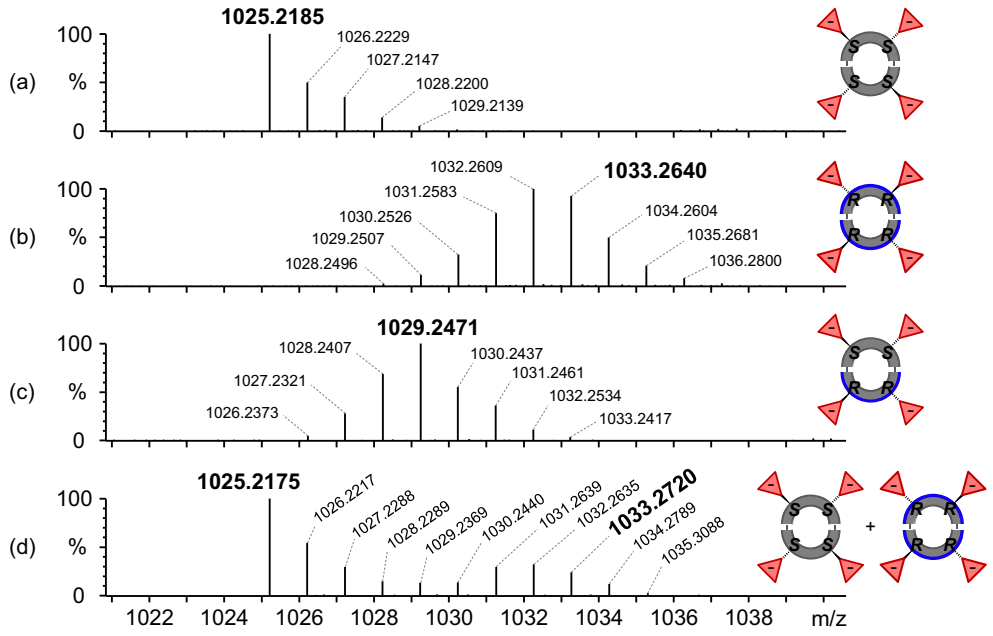


Figure 5.8. ESI(+)-TOF isotopic patterns (experimental mass peaks) of (a) (*SS*,*SS*)-(**1i**)$_2$ (peak at 31.5 min in Figure 5.7b), (b) (*RR*,*RR*)-(**1i**)$_2$-d_8 (peak at 31.5 min in Figure 5.7c), (c) (*RR*,*SS*)-(**1i**)$_2$-d_4 (peak at 30.8 min in Figure 5.7a), and (d) (*SS*,*SS*)-(**1i**)$_2$ + (*RR*,*RR*)-(**1i**)$_2$-d_8 (peak at 31.5 min in Figure 5.7a).

5.3.4. Starting situation and quantitative evaluation

An improved binary mixture was designed with the following three adjustments: i) a lower concentration of the BBs to favor the formation of dimers over trimers; ii) a lower %DMSO to be able to incorporate higher amounts of a second organic co-solvent (section 5.3.5), as well as to prevent overoxidation when using high temperatures (section 5.3.6); and iii) a slight increase in the pH to accelerate the oxidation process.[e] Dithiols (*SS*)-**1i** and (*RR*)-**1i**-d_4 were mixed at 0.5 mM each in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO. The resulting library was analyzed by chiral-HPLC and, after several trials at different column temperatures, polarity gradients and flow rates; the separation of the three dimers was achieved as shown in Figure 5.9. The identity of the three dimers was uncovered by deconvolution experiments (section 5.5.5).

[e] When using small amounts of DMSO (1-2%), although these proved to already accelerate the formation of disulfides by a fourfold decrease of the oxidation half-life time (Chapter 2), a slightly basic pH further accelerates the process by allowing in parallel the air-oxidation mechanism.

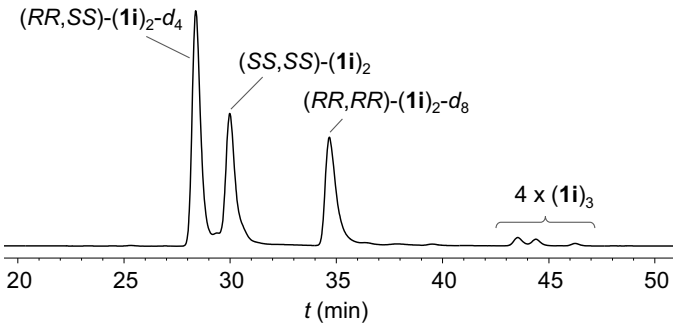
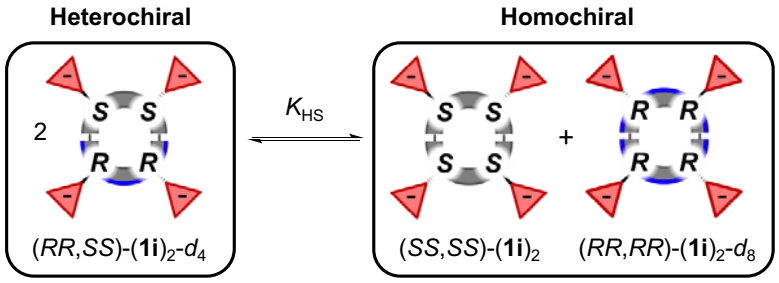


Figure 5.9. Chiral-HPLC-UV trace (254 nm) of the DCL generated from (*SS*)-**1i** + (*RR*)-**1i**-d_4 at 0.5 mM each in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO.

The library distribution was converted into valuable quantitative physicochemical information by calculating the exchange constant $K_{[\mathbf{A},\mathbf{B}]}$ (Equation 2.1) in its adapted version K_{HS} (homochiral selection exchange constant) defined in Equation 5.1. This constant is especially useful in the framework of this study since it corresponds to the reaction of converting 2 equivalents of the heterochiral dimer into 1 equivalent of each of the homochiral dimers (Scheme 5.2). Thus, the dimensionless K_{HS} constant is actually a ratio between the amount of homo- and heterochiral dimers, turning out to be a simple quantitative parameter to "detect" the prevalence of any of the two relative configurations. As previously discussed in Chapter 4, for an equimolar mixture of the two BBs, the statistical distribution of the dimers (*SS*,*SS*)-(**1i**)$_2$, (*RR*,*SS*)-(**1i**)$_2$-d$_4$ and (*RR*,*RR*)-(**1i**)$_2$-d_8 is 1:2:1 and K_{HS} = 0.25.[56]



Scheme 5.2. Exchange reaction between homo- and heterochiral dimers.

$$K_{HS} = \frac{\left[(SS,SS)\text{-}(\mathbf{1i})_2\right]\cdot\left[(RR,RR)\text{-}(\mathbf{1i})_2\text{-}d_8\right]}{\left[(RR,SS)\text{-}(\mathbf{1i})_2\text{-}d_4\right]^2} \qquad \textbf{(Equation 5.1)}$$

The exchange constant was calculated for the binary mixture shown in Figure 5.9 and, very interestingly, its value was found to be 0.36 ± 0.01 (slightly larger than the statistical 0.25 value). Hence, the library distribution is slightly shifted from the

statistical distribution in favor of the homochiral configuration. Very remarkably, this observation implies a certain degree of homochiral self-sorting.

5.3.5. The effect of the polarity of the solvent

We know from experience that polar interactions can have a huge impact on the relative stability of the members of those dynamic libraries generated from charged BBs (Chapter 4). As a consequence, the composition of these libraries is likely to be sensitive to the "polarity" of the medium. Taking these precedents into account, we wondered whether a polarity change could favor the homochiral self-selection, further shifting the equilibrium represented in Scheme 5.2 toward the formation of the homochiral dimers.

5.3.5.1. The polarity as an environmental factor

By changing the polarity of the medium no external chiral agents are added to the system. Therefore, a polarity change would never differentiate enantiomers but only diastereomeric species. Thus, a polarity change would always maintain the two pseudo-enantiomeric homochiral dimers (*SS*,*SS*)-(**1i**)$_2$ and (*RR*,*RR*)-(**1i**)$_2$-d_8 at equal stability levels, while it might cause an energetic differentiation between these two and the heterochiral dimer (*RR*,*RR*)-(**1i**)$_2$-d_8. In this respect, a polarity variation could have an effect on the equilibrium represented in Scheme 5.2, and consequently also on the K_{HS} exchange constant.

5.3.5.2. Polarity screening

A decrease in the polarity was envisioned to favor the polar intramolecular interactions of the library members. With this aim, we decided to incorporate acetonitrile (AN) as an organic co-solvent. In 1981 Balakrishnan and Easteal[57] reasoned that, because AN is less polar than pure water, the polarity of the mixture should decrease upon the addition of AN and would be dictated by the fact that the water molecules are more polar in isolation than as a part of a cluster. Since then, many other studies on the physicochemical properties of this mixture have been reported and reviewed.[58] Water/AN mixtures are recurring reaction media in physical organic chemistry,[59] are known to play an important role in atmospheric chemistry,[60] and have gained practical significance through use in RP-LC.[61] In some examples water/AN

mixtures[62-66] or even pure AN[f,67] have been used for the preparation of disulfide based DCLs.

The polarity of water/AN mixtures can be macroscopically measured by the relative static permittivity (formerly called static dielectric constant).[68] At 25 °C the permittivity of water/AN mixtures progressively decreases from 78.3[69] (pure water) to 35.9[69] (pure AN). From a microscopic point of view, the variations are not so gradual. Douhéret *et al*. inferred the presence of three structurally different regions over the water/AN mixture composition range.[70] The bounds for such regions, in mole fraction of AN (χ_{AN}), are $0 \leq \chi_{AN} \leq 0.2$, $0.2 < \chi_{AN} < 0.75$ and $0.75 \leq \chi_{AN} \leq 1$.

When using AN as a co-solvent, the use of phosphate buffer was avoided for solubility reasons and, alternatively, the pH of the mixtures was adjusted by the addition of tetrabutylammonium hydroxide (TBAOH). Apart from its solubility in AN, the use of this base has some additional important advantages. Firstly, the tetrabutylammonium (TBA^+) cataions have poor coordination abilities and, therefore, should not efficiently compete with the formation of intramolecular H-bonds. Secondly, the replacement of the phosphate buffer system by the TBAOH base contributes to the decrease in polarity by decreasing the ionic strength. TBAOH has already been used in the framework of DCC, proving to be a suitable base catalyst to allow disulfide equilibration even in pure organic solvents.[71]

A set of binary mixtures of (*SS*)-**1i** and (*RR*)-**1i**-d_4 was prepared at 0.5 mM each in TBAOH basified water (pH ~8) with 1.2% (v/v) DMSO and incorporating increasing amounts of AN: 0, 10, 30, 50, 70, 86 and 92% (v/v). No higher proportions of AN were investigated because a minimum water content is needed for the oxidation of the dithiols to take place in a reasonable rate.[71,72] All %AN were chosen arbitrarily except the 86% that corresponds to the azeotropic proportion.[73] Additionally, one last sample was conceived to investigate the opposite stimulus: a polar and chaotropic medium. This was prepared at the same experimental conditions as the others, but the TBAOH basified water was replaced by an aqueous phosphate buffer (pH 8.0) containing 1.0 M guanidinium chloride (GndHCl).

[f] When pure AN is used for generating DCLs, the BBs are added in already oxidized forms (as disulfides). In the absence of water the oxidation process does not take place in a reasonable rate.

The samples were left to oxidize at 22 °C and different reaction times were needed depending on the solvent composition (Table 5.3 in section 5.5.8). While the presence of GndHCl is known to assist the thiols oxidation,[74] the presence of AN was found to slow the process down. Once the mixtures reached the thermodynamic equilibrium, they were subsequently analyzed by chiral-HPLC.

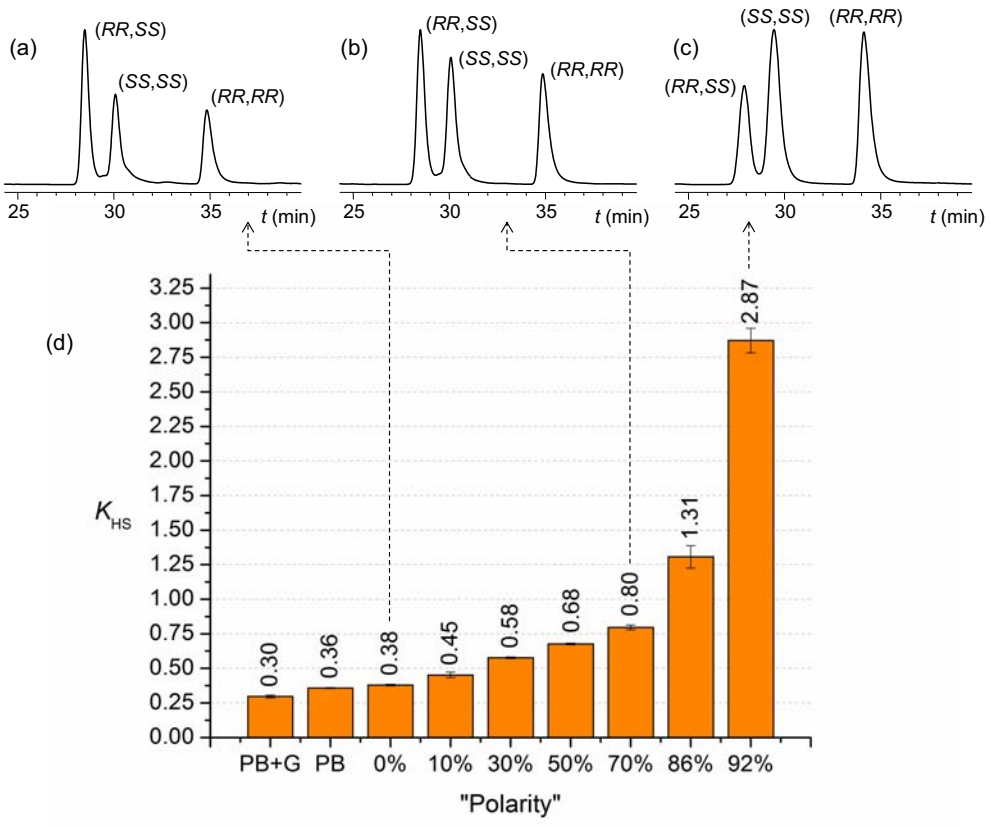


Figure 5.10. (a) Chiral-HPLC-UV traces (254 nm) of the binary mixture of (*SS*)-**1i** + (*RR*)-**1i**-d_4 (0.5 mM each) performed at 22 °C in (a) basified water + 1.2% (v/v) DMSO; (b) basified water + 70% (v/v) AN + 1.2% (v/v) DMSO; and (c) basified water + 92% (v/v) AN + 1.2% (v/v) DMSO. (d) Representation of the K_{HS} constant as a function of a qualitative "polarity" scale. (PB+G = aqueous phosphate buffer (pH 8.0) with 1.0 M Gnd·HCl; PB = aqueous phosphate buffer (pH 8.0); the 0, 10, 30, 50, 70, 86 and 92% values indicate the mixtures prepared with this volume % of AN).

For a preliminary qualitative analysis, Figures 5.10a-c show representative HPLC traces corresponding to the mixtures containing 0, 70 and 92% (v/v) AN respectively. Very interestingly, the library radically changed its composition upon the incorporation of AN in the medium. Initially, in the absence of AN, each of the statistically disfavored homochiral dimers (*SS*,*SS*)-(**1i**)$_2$ and (*RR*,*RR*)-(**1i**)$_2$-d_8 has a notably lower concentration than the heterochiral dimer (*RR*,*SS*)-(**1i**)$_2$-d_4 (Figure 5.10a). As the proportion of AN increases, the two homochiral members are equally amplified at the expenses of the heterochiral dimer, reaching a final situation at 92% (v/v) AN in which each of the

homochiral dimers has a quite larger concentration than the heterochiral counterpart (Figure 5.10c). Thus, very remarkably, these observations imply that a decrease in the polarity of the solvent promotes the homochiral self-selection.

The K_{HS} constant was calculated for all the libraries and represented in front of a qualitative polarity scale (Figure 5.10d). The three first columns clearly show the effect of the buffer and the GndHCl salt. Taking the sample prepared in phosphate buffer as the reference ("PB" in Figure 5.10d), the incorporation of 1.0 M GndHCl ("PB+G" in Figure 5.10d) causes a decrease in the K_{HS} constant from 0.36 ± 0.01 to 0.30 ± 0.01, in agreement with the corresponding increase in the ionic strength and the denaturant nature of the guanidinium salt.[75] On the contrary, when the phosphate buffer is replaced by the TBAOH basified medium ("0%" in Figure 5.10d), because of the previously commented advantages of the latter, the constant slightly increases up to 0.38 ± 0.01. Regarding the effect of the AN, the homochiral self-selection prominently increases as the polarity of the medium decreases, largely pushing up above the azeotropic composition (86% AN in water). Remarkably, at 92% AN the K_{HS} value is tenfold the theoretical one for a statistical distribution (absence of chiral selection). These results suggest that the homochiral self-selection is driven by polar intramolecular interaction that are favored in less competitive media.

The effect of the polarity was studied in more detail by calculating the difference between the standard Gibbs free energy of reaction of each of the experimental conditions "i" ($\Delta G^{o}_{i} = -RT\ln(K_{HS})$) and the statistical situation ($\Delta G^{o}_{stat} = -RT\ln(0.25)$). The obtained energetic values ($\Delta\Delta G^{o}_{i}$, Equation 5.2) quantify how much the libraries are shifted away from the statistical distribution. Hence, the $\Delta\Delta G^{o}_{i}$ energy is a very useful indicator, as it is a direct measure of the homochiral self-selection magnitude.

$$\Delta\Delta G^{o}_{i} = \Delta G^{o}_{i} - \Delta G^{o}_{stat} = -RT\ln\left(\frac{K_{HS}}{0.25}\right) \qquad \textbf{(Equation 5.2)}$$

For the samples prepared in phosphate buffer in the absence and presence of 1.0 M GndHCl, the calculated $-\Delta\Delta G^{o}_{i}$ energy was 0.88 and 0.42 kJ·mol^{-1} respectively. These small values are in agreement with the small preference for the homochiral relative configuration at these highly competitive media. Regarding the samples containing 0-92% (v/v) AN, their $-\Delta\Delta G^{o}_{i}$ energy was represented as a function of the AN mole fraction (Figure 5.11). This plot unambiguously shows that the selectivity for the

homochiral configuration is higher as the polarity decrease (the $-\Delta\Delta G_i^o$ energy rising up to 5.99 kJ·mol^{-1} in the presence of 92% AN). Besides, this trend defines a "cubic-like" shape with three different regions. Thus, the variation of the molar fraction of AN exerts a higher impact at low and high concentrations of the organic co-solvent, being less sensitive within the range $0.15 < \chi_{AN} < 0.45$.

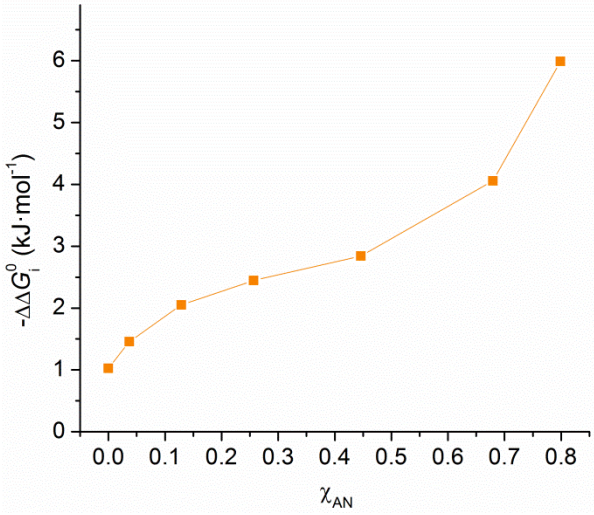


Figure 5.11. Representation of the $-\Delta\Delta G_i^o$ energy as a function of the AN mole fraction for the binary mixtures oxidized at 22 °C.

5.3.6. The effect of the temperature

Before investigating the effect of the temperature on our system, we wondered about the stability of the TBA^+ cation. As a quaternary ammonium salt, TBA^+ can undergo a Hofmann-type elimination when exposed to high temperatures and basic media.[76] In this side reaction 1-butene is produced by elimination (E_2) of tributylamine and an adjoining hydrogen atom (eventually releasing tributylammonium). Such transformation would not alter the pH of the mixture but would suppose an undesired change on the experimental conditions. The TBA^+/Bu_3NH^+ counterion replacement could have an influence on the relative energy of the library members and mask the effect of the temperature. Fortunately, ^{1}H NMR and MS analyses confirmed the TBA^+ stability both in the presence and absence of AN, at least when exposed to 45 °C for up to 3 days (section 5.5.7).

Two batches of nine binary mixtures were prepared with exactly the same compositions as before (PB+G, PB, 0-92% AN). One batch was left to oxidize in a fridge at 2 °C, and the other in an oven at 45 °C. Samples at 2 °C required 9-11 days to

fully oxidize, while samples at 45 °C only needed 46-64 hours (section 5.5.8). Once the mixtures reached the equilibrium composition (reversibility tests in section 5.5.9) all libraries were subsequently analyzed by chiral-HPLC. Table 5.1 contains the values of the K_{HS} constant and $-\Delta\Delta G_i^o$ energy calculated for the nine different "polarities" at the three different temperatures (2 and 45 °C, together with the previous data set performed at 22 °C). For the sample prepared in phosphate buffer in the presence of 1.0 M GndHCl (entry 1 in Table 5.1), no reliable quantitative data was obtained at 45 °C because of the fast overoxidation of the library members at these experimental conditions.

Table 5.1. Calculated values for the K_{HS} constant and the $-\Delta\Delta G_i^o$ energy ($kJ \cdot mol^{-1}$) of the binary mixtures prepared at different solvent compositions and temperatures.

			2 °C		22 °C		45 °C	
entry	sample	χ_{AN}	K_{HS}	$-\Delta\Delta G_i^o$	K_{HS}	$-\Delta\Delta G_i^o$	K_{HS}	$-\Delta\Delta G_i^o$
1	PB+G	0.00	0.27 ± 0.01	0.16	0.30 ± 0.01	0.42	— [a]	— [a]
2	PB	0.00	0.30 ± 0.01	0.44	0.36 ± 0.01	0.88	0.39 ± 0.01	1.17
3	0%	0.00	0.33 ± 0.01	0.65	0.38 ± 0.01	1.03	0.42 ± 0.01	1.34
4	10%	0.04	0.39 ± 0.01	1.01	0.45 ± 0.02	1.46	0.47 ± 0.01	1.65
5	30%	0.13	0.54 ± 0.01	1.77	0.58 ± 0.01	2.05	0.58 ± 0.01	2.23
6	50%	0.26	0.62 ± 0.01	2.09	0.68 ± 0.01	2.44	0.67 ± 0.01	2.61
7	70%	0.45	0.70 ± 0.02	2.34	0.80 ± 0.02	2.84	0.80 ± 0.01	3.09
8	86%	0.68	1.10 ± 0.04	3.39	1.30 ± 0.08	4.05	1.40 ± 0.05	4.55
9	92%	0.80	2.68 ± 0.06	5.42	2.87 ± 0.09	5.99	3.0 ± 0.2	6.6

[a] No reliable quantitative data was obtained at these experimental conditions because of the fast overoxidation of the library members.

At a first glance, values in Table 5.1 show two evident tendencies. First of all, for the three different temperatures the decrease in the polarity of the medium causes an increase in the $-\Delta\Delta G_i^o$ energy. Secondly, and very interestingly, the values in Table 5.1 are ordered not only within each column but also within each row. As the temperature increases the $-\Delta\Delta G_i^o$ values also increase. Hence, the homochiral self-sorting increases with the temperature.

In Figure 5.12 the $-\Delta\Delta G_i^o$ energies corresponding to entries 3-9 in Table 5.1 are represented as a function of the χ_{AN}. In order to evaluate the effect of the temperature, the values are gathered in three series, one for each temperature. These series respond equivalently to the increase in the AN content. However, the position of each series over the vertical axis clearly depends on the temperature. For all %AN the $-\Delta\Delta G_i^o$ energy clearly increases with the temperature. Very interestingly, this trend suggests a positive entropic contribution ($\Delta\Delta S_i^o > 0$) to the selection process. Back to our initial source of

inspiration, some authors have proposed that the origin of biological homochirality would be an entropically driven event.[16,77-80]

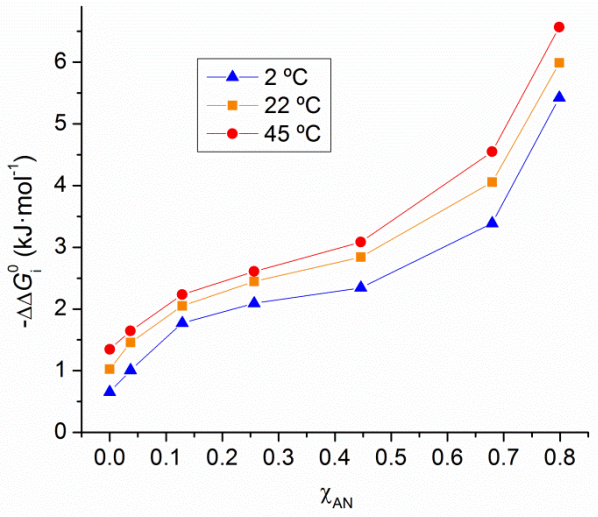


Figure 5.12. Representation of the $-\Delta\Delta G_i^o$ energy as a function of the AN mole fraction, for the batches oxidized at 2 °C (—▲—), 22 °C (—■—) and 45 °C (—●—).

In a recent paper, Moore and co-workers reported that the increase of temperature can stabilize the diastereomers with higher symmetry (section 5.1.2.2).[50] A difference in symmetry has an impact on the rotational entropy, since structures with higher symmetry have a larger number of indistinguishable orientations reached by rotation.[g,81,82] However, since in our case both homo- and heterochiral dimers have an identical symmetry order (σ = 4, determined from point group),[83] this entropic factor cannot explain the observed behavior.

5.3.7. Additional experiments

5.3.7.1. The effect of the pH

The effect of the protonation degree of the library members on the composition of the molecular network was studied by performing the binary mixture at different pH values. The (*SS*)-**1i** and (*RR*)-**1i**-d_4 BBs were mixed at 0.5 mM each in different aqueous buffers (pH 2.5-7.5)[h] with 56% (v/v) AN and 1.2% (v/v) DMSO. Once the mixtures reached the thermodynamic equilibrium, they were subsequently analyzed by chiral-

[g] Entropy differences are typically associated with higher disorder. Alternatively, entropy can be interpreted in terms of changes in information. Specifically, an increase in entropy is associated with loss of information (indistinguishability) or increased symmetry.

[h] The McIlvaine phosphate-citrate buffer system was used to cover the 2.5-7.5 pH range and different amounts of sodium chloride were added as an inert salt in order to fix the ionic strength.

HPLC and both K_{HS} and $-\Delta\Delta G_i^o$ parameters were calculated. Also the percent amount of trimers was measured for each of the tested pH values.

In Figure 5.13 the $-\Delta\Delta G_i^o$ energy and the % Trimers are represented as a function of the pH. Interestingly, these two monitored parameters show opposed behaviors. On one hand, the relative amount of trimers decreases from 11.8% at pH 2.5 to 3.7% at pH 7.5. This tendency is in agreement with the progressive deprotonation of the library members as the basicity increases. The system accommodates to the increase in the number of charged carboxylate groups by disfavoring those species concentrating negative charges. On the other hand, and more importantly, the homochiral self-selection ($-\Delta\Delta G_i^o$) gradually increases as the pH goes from 2.5 to 7.5, reaching a plateau at pH 6.5-7.5. At this pH the members of the library are already fully deprotonated and, therefore, the tetraanionic form is the one showing the largest energetic difference between homo- and heterochiral dimers. This observation is in agreement with the polar interactions being implicated in the homochiral self-selection: a larger number of deprotonated carboxylates implies a larger number of possible electrostatic and H-bonding intramolecular interactions.

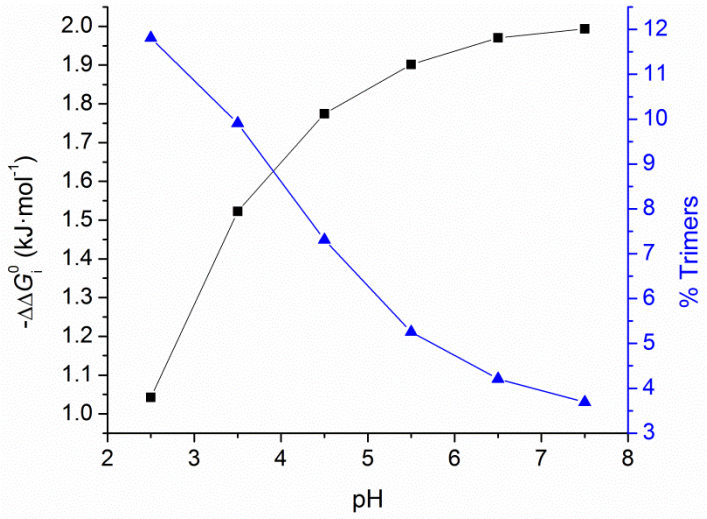


Figure 5.13. Representation of the $-\Delta\Delta G_i^o$ energy (left axis, —■—) and the %Trimer (right axis, —▲—) as a function of the pH.

5.3.7.2. The effect of the concentration

Aggregation processes can influence the composition of a DCL by favoring those species that form the most stable aggregates.[84] Thus, in principle aggregation could be conceived as the driving force behind the observed homochiral self-sorting.[44,45] However, strong intermolecular interactions are hardly imaginable to take place for the

members of this molecular network, as they are tetraanionic species at pH ~8, and the corresponding repulsive electrostatic interactions should prevent aggregation processes to occur. Even so, a simple experiment was designed in order to exclude the possibility of aggregation. Two binary mixtures were prepared with exactly the same experimental conditions as samples "PB" and "70%" in section 5.3.5.2, but four times more concentrated. They were left to oxidize at 22 °C for 7 and 8 days respectively, until they reached the thermodynamic equilibrium. Finally, the mixtures were analyzed by chiral-HPLC and the corresponding K_{HS} and $-\Delta\Delta G_i^o$ values were calculated in order to be compared with those of the mixtures performed at lower concentration.

Table 5.2. Calculated values for the K_{HS} constant and the $-\Delta\Delta G_i^o$ energy ($kJ \cdot mol^{-1}$) of the binary mixtures performed at 0.5 and 2.0 mM concentration of each BB.

			0.5 mM		2.0 mM	
entry	sample	χ_{AN}	K_{HS}	$-\Delta\Delta G_i^o$	K_{HS}	$-\Delta\Delta G_i^o$
1	PB	0.00	0.36 ± 0.01	0.88	0.37 ± 0.01	0.93
2	70%	0.45	0.80 ± 0.02	2.84	0.82 ± 0.01	2.92

As shown in Table 5.2, both in the presence and absence of AN, the K_{HS} constant has the same value regardless the concentration of the BBs. Since intermolecular interactions intrinsically dependent on the concentration, the observed equal outcome for the experiments performed at different concentration corroborates that no aggregation processes are responsible for the homochiral self-selection.

5.3.8. Preliminary structural studies

In order to understand why in water/AN mixtures the homochiral configuration has an inherent larger stability, we decided to perform the detailed NMR characterization of both homo- and heterochiral dimers. With this purpose, two samples were prepared, the first one containing (*SS*)-**1i** (2.0 mM) in TBAOH basified water (pH 7.6) with 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6, and the second one containing (*SS*)-**1i** + (*RR*)-**1i**-d_4 (1 mM each) in exactly the same solvent. After 5 days at room temperature, the *in situ* oxidation of the starting dithiol building blocks led to the formation of the corresponding homochiral dimer in the first case, and a mixture of homo- and

heterochiral dimers in the second case.[i] The ^{1}H NMR spectra of both samples (acquired at low temperature)[j] are compared in Figure 5.14.

Considering the stereoisomeric relationship between the members of the mixture, the two homochiral dimers must show perfectly overlapped NMR signals, as they are enantiomers, whereas the heterochiral dimer could present a different NMR signature. At a first glance, the spectrum of the diastereoisomeric mixture (Figure 5.14c) seems to be more complex than that of the pure homochiral dimer (Figure 5.14b). Very importantly, this observation implies distinct NMR profiles for the homo- and heterochiral dimers. This is a remarkable result, given that the two diastereomeric dimers have practically identical structures.

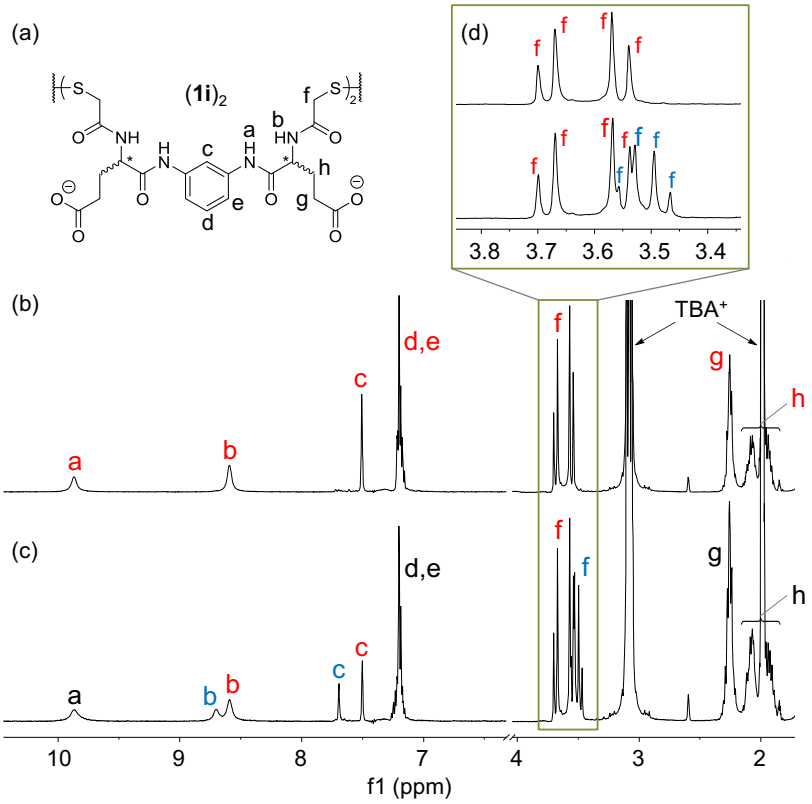


Figure 5.14. (a) Chemical structure of (**1i**)$_2$. Partial ^{1}H NMR (500 MHz, "WATER_ES" water suppression pulse sequence, 278 K) spectra of (b) (*SS*)-**1i** (2.0 mM) oxidized in TBAOH basified water (pH 7.6) containing 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6, and (c) (*SS*)-**1i** + (*RR*)-**1i**-d_4 (1 mM each) oxidized in TBAOH basified water (pH 7.6) containing 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6. (d) Zoom at the 3.8-3.4 region of (b) and (c) spectra. Signals of the homo- and heterochiral dimers are identified in red and blue respectively. Overlapped signals of both diastereoisomeric species are identified in black.

[i] Regarding the mixture of homo- and heterochiral dimers, the composition of an equivalent sample was previously commented in section 5.3.5.2, Figure 5.10b).

[j] Considering the conditions used for the experiment (water suppression pulse sequence, section 5.5.11), at room temperature the amide protons were not observable due to their fast exchange with water. By lowering the temperature to 5 °C the exchange slows down and the amide signals become observable.

The signals of the binary mixture were assigned by direct comparison with the spectrum of the pure homochiral dimer (Figure 5.14b *vs.* 5.14c) and by ^{1}H-^{1}H correlation experiments (gCOSY, Figure 5.21a in section 5.5.11). Very interestingly, the protons showing different signals for the two diastereomeric species are precisely those described in Chapter 3 as the ones exhibiting larger changes depending on the folding degree of dimer (**1i**)$_2$. Thus, the aliphatic amido NH (b) of the heterochiral dimer is deshielded with respect to the one of the homochiral counterpart, whereas the aromatic c proton is substantially shielded. Additionally, the AB quartet of the methylene in α to the disulfide bond (f) presents a markedly larger anisochrony for the homochiral species (Figure 5.14d). Surprisingly, these three observations seem somehow contradictory. On one hand, the deshielded amido NH of the aliphatic amide would suggest the establishment of more H-bonding interactions for the heterochiral dimer, a feature associated with a more folded/packed conformation. On the other hand, the other two signals would suggest the opposite trend: the shielded aromatic proton c of the heterochiral dimer is associated with a more unfolded/unpacked conformation (as previously observed in section 3.3.6.2), and the smaller anisochrony of the proton f of the heterochiral dimer would also suggest a more flexible conformation for this species. In an attempt to shine some light on this topic, the 2D ROESY spectrum was recorded for the sample of the mixture of stereoisomers. Unfortunately, this experiment did not provide any relevant information regarding the conformation of the two diastereomeric dimers (Figure 5.21b in section 5.5.11).

Overall, the differences observed at the NMR spectra suggest significantly different conformations for the homo- and heterochiral dimers in water/AN mixtures. Unfortunately, these differences are inconclusive for determining the conformational preferences of the two diastereomeric dimers. Molecular modeling simulations are foreseen to provide valuable information and will be carried out in due course. These simulations are out of the scope of this thesis.

5.4. Conclusions

A simple DCL made from the combination of an enantiomeric pair of pseudopeptidic building blocks has proved to be a suitable molecular network in which chiral self-sorting phenomena can take place. This is evidenced in the following conclusions:

i) The mixture of BBs (*SS*)-**1i** and (*RR*)-**1i**-d_4 leads to the formation of a disulfide-based DCL consisting of two homochiral dimers and one heterochiral dimer. The isotopic labeling of the building blocks has enabled MS analysis for diastereo- and enantiocomposition assessment.

ii) The composition of the library generated in buffered water has been quantitatively evaluated, showing to be slightly biased in favor of the homochiral dimers. The decrease in the polarity resulting from the incorporation of acetonitrile as an organic co-solvent favors the homochiral self-selection. This result suggests that the selection process is driven by polar intramolecular interactions that are favored in less polar media.

iii) The homochiral selectivity also increases with the temperature. This trend suggests a positive entropic contribution to the selection process, in close resemblance to the origin of the biological homochirality.

iv) The homochiral self-selection increases with the pH, indicating that the fully deprotonated form of the library members is the one showing the largest selectivity. On the contrary, the library distribution is not dependent on the concentration of the BBs, indicating that no aggregation processes are responsible for the homochiral self-sorting.

v) Preliminary NMR experiments suggest significantly different conformations for the homo- and heterochiral dimers. Further structural studies are foreseen in order to better understand the self-sorting process at the molecular level.

Overall, these findings illustrate how an accurate transmission of the chiral information from the molecular to the macromolecular level can lead to homochiral self-sorting processes. In this regard, DCLs have proved to be suitable tools for the experimental modeling and study of such phenomena.

5.5. Experimental section

5.5.1. General methods

Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka and Merck) and were used without further purification. pH measurements were performed at room temperature on a Crison GLP21 pH-meter with the electrodes Crison 50 14T (≥10 mL samples) and PHR-146 Micro (<10 mL samples). NMR spectroscopic experiments were carried out on a Varian Mercury 400 instrument (400 MHz for ^{1}H and 101 MHz for ^{13}C). Absorbance measurements were performed on a Molecular Devices SpectraMax M5 (all spectra were recorded at room temperature).

5.5.2. HPLC and MS analyses

The achiral-HPLC and MS analyses were performed as specified in Chapter 2 (section 2.5.2). The chiral-HPLC analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using a reversed-phase CHIRALPAK® IA (25 x 0.46 cm, 5 μm) chiral column and (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures were used as mobile phase. The eluent used was: 2 min at 20% CH_3CN in H_2O, then linear gradient from 20% to 30% CH_3CN over 48 min (constant flow set at 0.5 $mL \cdot min^{-1}$). The monitoring wavelength was set at 254 nm and the temperature of the column was set at 25 °C. The HPLC samples were prepared by dilution with an acidic solution of 89% H_2O, 10% CH_3CN and 1% TFA.

5.5.3. Identification of the library members

Dithiols (*SS*)-**1i** and (*RR*)-**1i**-d_4 were mixed at 2.0 mM each in 20 mM aqueous phosphate buffer (pH 7.5) with 25% DMSO. After 24 hours at room temperature the libraries were analyzed by achiral-UPLC-ESI-TOF (Figure 5.15). The identification of the dimers (peaks at 13.0 and 13.3 min in Figure 5.15) was already discussed in the results and discussion section. Regarding the peak at 14.9 min in Figure 5.15, the following data confirmed the overlap of the four possible trimers: HRMS (ESI+) calcd. for [(*SS*,*SS*,*SS*)-(**1i**)$_3$+H]$^+$ (m/z): 1537.3179, found: 1537.3337; calcd. for [(*RR*,*SS*,*SS*)-(**1i**)$_3$-d_4+H]$^+$ (m/z): 1541.3430, found: 1541.3616; calcd. for [(*RR*,*RR*,*SS*)-(**1i**)$_3$-d_8+H]$^+$ (m/z): 1545.3682, found: 1545.3593; calcd. for [(*RR*,*RR*,*RR*)-(**1i**)$_3$-d_{12}+H]$^+$ (m/z): 1549.3933, found: 1549.4169.

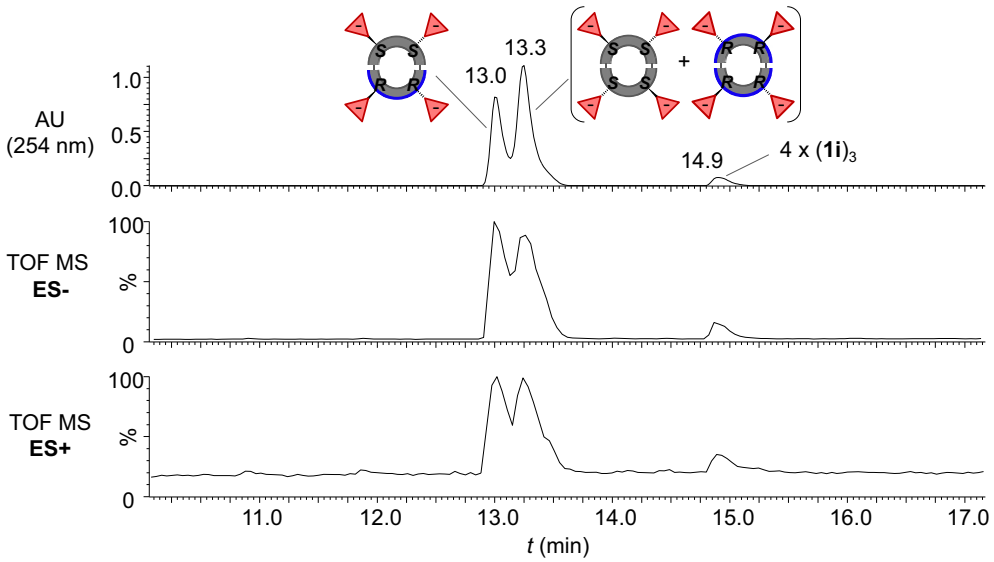


Figure 5.15. UPLC-UV(254 nm)-ESI-TOF traces of the DCL generated from (*SS*)-**1i** + (*RR*)-**1i**-d_4 (2.0 mM each) in 20 mM aqueous phosphate buffer (pH 7.5) with 25% (v/v) DMSO.

5.5.4. Preparation of the *individual stocks* and a the *stock mixture*

For each of dithiols (*SS*)-**1i** and (*RR*)-**1i**-d_4 a 14.7 mM *individual stock* was prepared in basified water. First of all, the corresponding amount of solid was added to the needed amount of Milli-Q water, obtaining an acidic suspension. Then, a small volume of a 1.0 M TBAOH aqueous solution was added in order to adjust the pH to ~8. As the TBAOH solution was added, the suspension became clearer and eventually all solid was completely dissolved. A *stock mixture* of the two BBs (7.35 mM of each) was prepared by mixing equal volumes of the two *individual stocks*. The pH of the resulting mixture was measured, confirming to be ~8. This stock mixture was stored at –80 °C and used for the preparation of all DCLs of this Chapter.

5.5.5. Assignment of the library members in the chiral-HPLC trace

The three previously identified dimers of the binary mixture were assigned in the chiral-HPLC trace by deconvolution experiments. Building blocks (*SS*)-**1i** and (*RR*)-**1i**-d_4 were mixed at 0.5 mM each in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO. Simultaneously, the two BBs were separately oxidized in exactly the same experimental conditions. After 5 days at room temperature, the libraries were analyzed by chiral-HPLC. The peaks of the HPLC trace of the binary mixture were assigned by comparison with the chromatograms of the two separately oxidized building blocks (Figure 5.16).

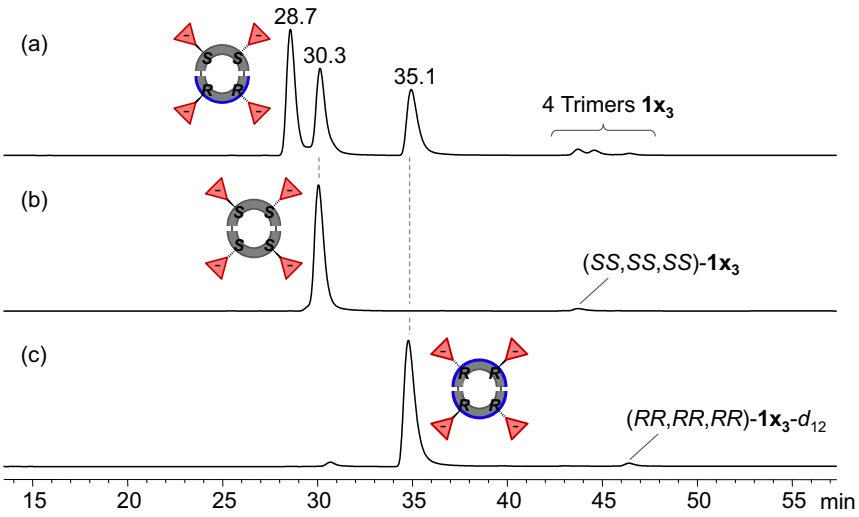


Figure 5.16. Chiral-HPLC-UV traces (254 nm) of the DCL generated from (a) (*SS*)-**1i** + (*RR*)-**1i**-d_4 (2.0 mM each), (b) (*SS*)-**1i** (2.0 mM), and (c) (*RR*)-**1i**-d_4 (2.0 mM); in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO.

5.5.6. Quantitative evaluation of the DCLs

The assumption of equal molar extinction coefficients at 254 nm (ε_{254}) for dithiols (*SS*)-**1i** and (*RR*)-**1i**-d_4 is not evident, since the deuterium labeling of (*RR*)-**1i**-d_4 is part of the *m*-diamidophenyl chromophore and could have an effect on its electronic properties. Thus, we decided to perform UV absorbance measurements at different concentrations for both (*SS*)-**1i** and (*RR*)-**1i**-d_4 in order to determine the corresponding ε_{254} values (Figure 5.17). The molar extinction coefficient was calculated as the slope of the least square regression line for the Abs_{254} *vs*. [BB] plot, and the obtained ε_{254} values were found to be the same within the experimental error (in methanol, $17.9 \cdot 10^3$ and $17.6 \cdot 10^3$ $M^{-1} \cdot cm^{-1}$ for (*SS*)-**1i** and (*RR*)-**1i**-d_4 respectively).

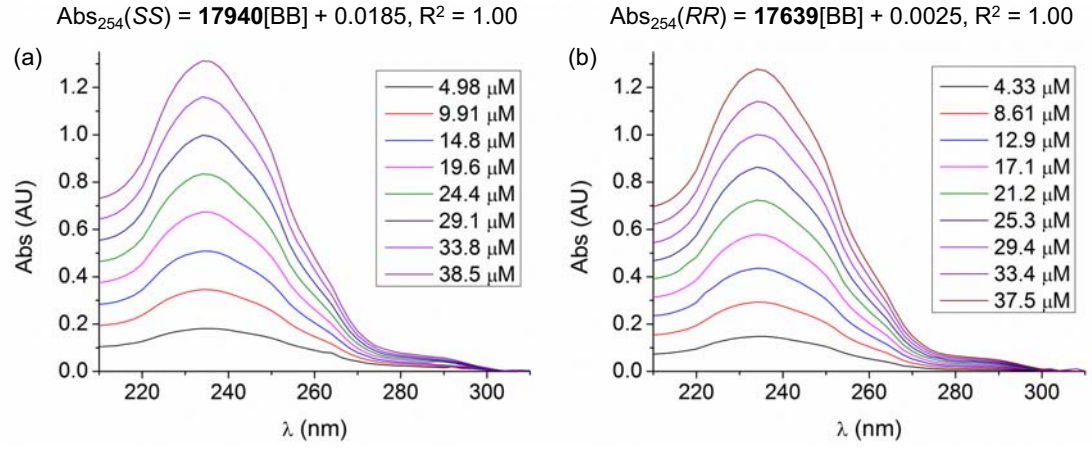


Figure 5.17. UV spectra of (*SS*)-**1i** (a) and (*RR*)-**1i**-d_4 (b) in methanol at different concentrations, and linear least square equations for λ = 254 nm.

The confirmed equal molar extinction coefficients for the two BBs and their known additive molar absorptivity when forming oligomers (Chapter 2, section 2.3.3), allowed directly using the chiral-HPLC peak areas for calculating the exchange constant. The presence of an unidentified impurity under the (*SS*,*SS*)-(**1i**)$_2$ peak (marked with an asterisk in Figure 5.18) prompted us to calculate the K_{HS} by means of Equation 5.3. Notice that, for an equimolar mixtures of (*SS*)-**1i** and (*RR*)-**1i**-d_4 the two generated homochiral dimers are necessarily also in equimolar amount.

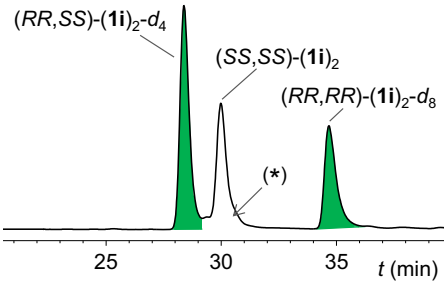


Figure 5.18. Chiral-HPLC-UV trace (254 nm) of the DCL generated from (*SS*)-**1i** + (*RR*)-**1i**-d_4 (0.5 mM each) in 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO. Green peaks are the ones used for calculating the K_{HS} constant and the asterisk corresponds to the unidentified impurity.

$$K_{HS} = \left(\frac{\left[(RR,RR)\text{-}(\mathbf{1i})_2\text{-}d_8 \right]}{\left[(RR,SS)\text{-}(\mathbf{1i})_2\text{-}d_4 \right]} \right)^2 \qquad \textbf{(Equation 5.3)}$$

All dynamic libraries were performed in triplicate. The given K_{HS} constants are the corresponding averaged values, and the associated confidence intervals were calculated for a 95% significance by means of the OriginPro 8.1 software.

5.5.7. Evaluation of the TBA$^+$ stability

Two NMR samples were prepared emulating the solvents composition and basicity of the binary mixtures but containing sodium acetate instead of the two Glu-based BBs. One sample was prepared by dissolving 10 mM CH_3COONa in D_2O with 1.2% (v/v) DMSO-d_6; and the other by dissolving 10 mM CH_3COONa in 50%[k] (v/v) CD_3CN with 48.8% (v/v) D_2O and 1.2% DMSO-d_6. TBAOH was used in both cases to adjust the pD to ~8. The ^{1}H NMR spectrum of the two samples was recorded before and after exposure to 45 °C for 3 days, showing no appreciable variations during this period of time (Figure 5.19). Additionally, after being exposed to 45 °C for 3 days, the two samples were analyzed by UPLC-ESI-TOF and no Et_3NH^+ was detected in any of them.

[k] No higher %AN were tested in order to avoid sodium acetate precipitation.

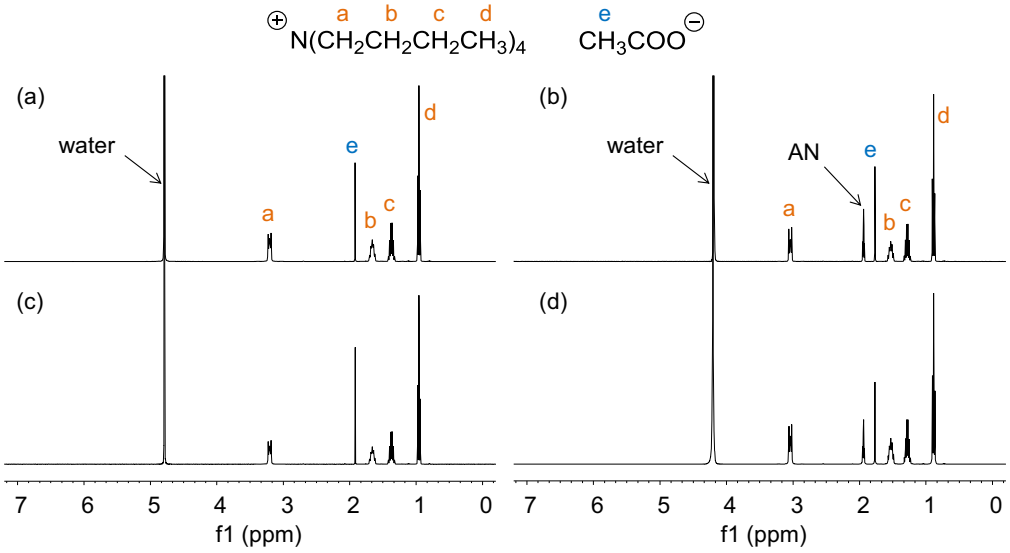


Figure 5.19. ^{1}H NMR (400 MHz, 298 K) spectra of (a) 10 mM CH_3COONa in TBAOH basified D_2O with 1.2% (v/v) DMSO-d_6, (b) 10 mM CH_3COONa in TBAOH basified D_2O with 50% (v/v) AN and 1.2% (v/v) DMSO-d_6, (c) sample (a) after 3 days at 45 °C, and (d) sample (b) after 3 days at 45 °C. The spectra in (a,c) are referenced with the water residual peak (4.79 ppm) and the spectra in (b,d) are referenced with the AN residual peak (1.94 ppm).

5.5.8. Preparation of the DCLs at different %AN and temperatures

The dynamic libraries were prepared at 0.5 mM of each BB by dilution of the *stock mixture* of (*SS*)-**1i** + (*RR*)-**1i**-d_4 (7.35 mM of each) with the corresponding amount of aqueous solution (phosphate buffer (pH 8.0) in the absence and presence of GndHCl for entries 1-2 in Table 5.3, and Milli-Q water for entries 3-9 in Table 5.3) and organic co-solvents (1.2% (v/v) DMSO for entries 1-9 in Table 5.3, and the corresponding volume % of AN for entries 4-9 in Table 5.3). Depending on the solvent composition, different reaction times were needed to fully oxidize the dithiol building blocks (Table 5.3).

Table 5.3. Reaction time of the oxidations carried out at 2, 22 and 45 °C.

entry	sample	χ_{AN}	2 °C	22 °C	45 °C
1	PB+G	0.00	9 d	5 d	— [a]
2	PB	0.00			46 h
3	0%	0.00	10 d	6 d	49 h
4	10%	0.04			53 h
5	30%	0.13			
6	50%	0.26			
7	70%	0.45	11 d	7 d	64 h
8	86%	0.68			
9	92%	0.80			

[a] No reliable quantitative data was obtained at these experimental conditions because of the fast overoxidation of the library members.

5.5.9. Reversibility tests

Two solutions of tris(2-carboxyethyl)phosphine (TCEP, 25 mM) were prepared by dissolving the corresponding amount of TCEP·HCl in: i) Milli-Q water with 1.2% (v/v) DMSO, and ii) Milli-Q water with 92% (v/v) AN and 1.2% (v/v) DMSO. A small volume of NaOH (aq) was added to adjust the pH of both solutions to ~8. The reversibility test was performed at the four extreme conditions of %AN and temperature: 0% AN at 2 °C, 92% (v/v) AN at 2 °C, 0% AN at 45 °C, and 92% (v/v) AN at 45 °C. To the binary mixtures oxidized at these four extreme conditions, a small volume of a TCEP solution (the one prepared with the same %AN) was added. After re-oxidation at the corresponding temperature, the libraries were subsequently analyzed by chiral-HPLC and the K_{HS}^{TCEP} constant (*i.e.* K_{HS} after the addition of TCEP and re-oxidation) was calculated (Table 5.4).

Table 5.4. Calculated exchange constants for the libraries before (K_{HS}) and after (K_{HS}^{TCEP}) the addition of 0.35 equivalents of TCEP and re-oxidation; at the extreme conditions of %AN (0 and 92% (v/v)) and temperature (2 and 45 °C).

		2 °C			45 °C		
entry	sample	K_{HS}	re-oxidation time	K_{HS}^{TCEP}	K_{HS}	re-oxidation time	K_{HS}^{TCEP}
1	0%	0.33 ± 0.01	3 d	0.31 ± 0.01	0.42 ± 0.01	14 h	0.40 ± 0.01
2	92%	2.68 ± 0.06		2.6 ± 0.1	3.0 ± 0.2		3.0 ± 0.1

After the addition of TCEP and re-oxidation, the exchange constant of the four libraries remained the same within the experimental error. Therefore, the binary mixtures were confirmed to reach the thermodynamic equilibrium at the four extreme conditions. By extension, all intermediate %AN and temperatures were also considered to be under thermodynamic control.

5.5.10. Preparation of the DCLs at different pHs and concentrations

The McIlvaine phosphate-citrate buffer system was used to prepare buffers at pH 2.5-7.5, as explained in Chapter 2 (section 2.5.3). In this case, however, the buffers were prepared at lower concentration (twofold dilution factor with respect to Chapter 2) in order to avoid water/AN immiscibility.[62] The dynamic libraries of the study of the effect of the pH were prepared at 0.5 mM of each BB by dilution of the *stock mixture* of (*SS*)-**1i** + (*RR*)-**1i**-d_4 (7.35 mM of each) with the corresponding amount of aqueous solution (buffers prepared at pH 2.5-7.5) and organic co-solvents (56% (v/v) AN and 1.2% (v/v)

DMSO). Different oxidation times were needed depending on the pH of the libraries (Table 5.5).

Table 5.5. Values of the K_{HS} constant, the $-\Delta\Delta G_i^o$ energy (kJ·mol^{-1}) and the % amount of Trimers (**1i**)$_3$, together with the oxidation times required to fully oxidize the binary mixtures prepared at different pH values.

entry	pH	K_{HS}	$-\Delta\Delta G_i^o$	% Trimers	reaction time
1	2.5	0.38 ± 0.01	1.04	11.8 ± 0.7	8 d
2	3.5	0.47 ± 0.01	1.52	9.9 ± 0.3	7 d
3	4.5	0.52 ± 0.01	1.77	7.3 ± 0.1	
4	5.5	0.34 ± 0.01	1.90	5.3 ± 0.2	6 d
5	6.5	0.56 ± 0.01	1.97	4.2 ± 0.1	
6	7.5	0.56 ± 0.01	1.99	3.7 ± 0.4	5 d

Whereas in Chapter 2 the oxidation rate was found not to significantly change within the pH range 2.5-6.5 for samples containing 10-25% DMSO, in this case the decrease in the pH clearly slows the oxidation process down. When relatively large amounts of DMSO are used the thiols are rapidly oxidized mainly by reaction with DMSO. However, in the presence of much smaller amounts of DMSO (1.2% (v/v) in this study) the reaction with DMSO is notably slower and the air-oxidation mechanism becomes important. The latter is strongly dependent on the pH, and this is why at low %DMSO the oxidation rate changes with the pH.

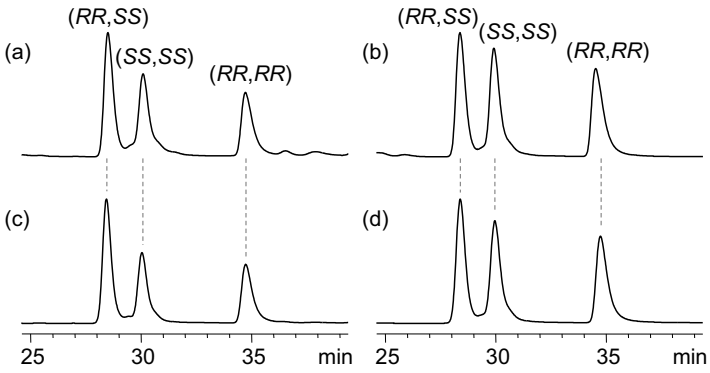


Figure 5.20. Chiral-HPLC-UV traces (254 nm) of the binary mixture of (*SS*)-**1i** + (*RR*)-**1i**-d_4 performed at: (a) 2.0 mM each BB in aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO; (b) 2.0 mM each BB in TBAOH basified water (pH ~8) with 70% (v/v) AN and 1.2% (v/v) DMSO; (c) 0.5 mM each BB in aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO; and (d) 0.5 mM each BB in TBAOH basified water (pH ~8) with 70% (v/v) AN and 1.2% (v/v) DMSO.

The dynamic libraries of the study of the effect of the concentration were prepared at 2.0 mM of each BB by dilution of the *stock mixture* with: i) 20 mM aqueous phosphate buffer (pH 8.0) with 1.2% (v/v) DMSO; and ii) TBAOH basified water (pH ~8) with 70% (v/v) AN and 1.2% (v/v) DMSO. The libraries were left to oxidize for 7 and 8 days

respectively. In Figure 5.20 the HPLC traces of these two concentrated samples are compared with the corresponding samples performed at 0.5 mM concentration.

5.5.11. Nuclear magnetic resonance

For each of dithiols (*SS*)-**1i** and (*RR*)-**1i**-d_4 a 7.00 mM *individual stock* was prepared in TBAOH basified water (pH 7.6) as explained in section 5.5.4. From these, two NMR samples were prepared: i) (*SS*)-**1i** (2.0 mM) in TBAOH basified water (pH 7.6) with 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6 (prepared by mixing 216 µL of the (*SS*)-**1i** *individual stock* + 525 µL of CD_3CN + 9 µL of DMSO-d_6); and ii) (*SS*)-**1i** + (*RR*)-**1i**-d_4 (1.0 mM each) in TBAOH basified water (pH 7.6) with 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6 (prepared by mixing 108 µL of the (*SS*)-**1i** *individual stock* + 108 µL of the (*RR*)-**1i**-d_4 *individual stock* + 525 µL of CD_3CN + 9 µL of DMSO-d_6). Both samples were left to oxidize for 5 days at room temperature.

The ^{1}H NMR spectrum of the two samples (Figure 5.14) was recorded on a Varian INOVA 500 spectrometer (500 MHz for ^{1}H, 278 K). Additionally, only for the sample containing both BB, gCOSY (Figure 5.21a) and 2D ROESY (Figure 5.21b) experiments were also performed on the same spectrometer. All spectra were acquired using the same optimized parameters for the Agilent Chempack pulse sequence WATER_ES (with excitation sculping).[85]

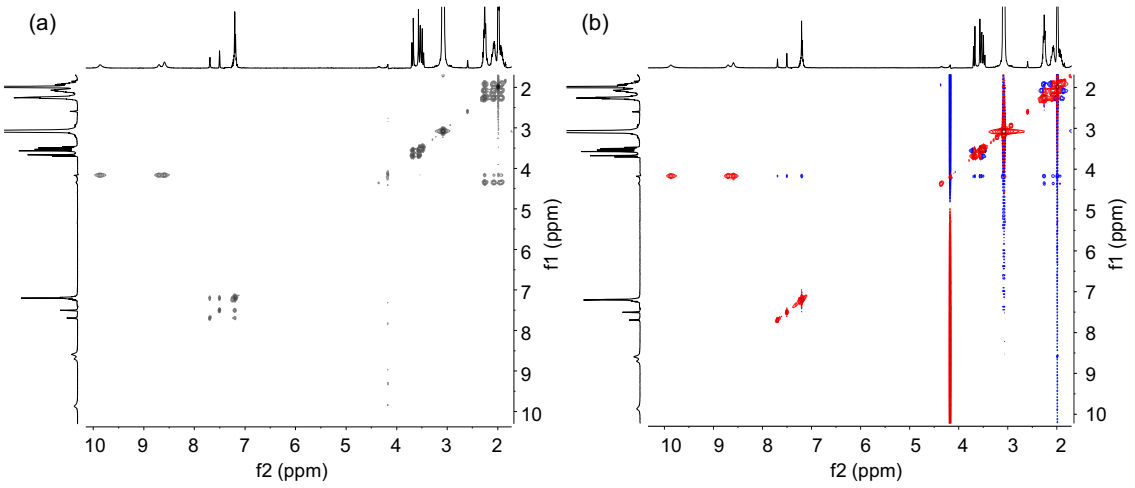


Figure **5.21.** (a) ^{1}H-^{1}H gCOSY and (b) 2D ROESY (500 MHz, "WATER_ES" water suppression pulse sequence, 278 K) spectra of (*SS*)-**1i** + (*RR*)-**1i**-d_4 (1 mM each) oxidized in TBAOH basified water (pH 7.6) containing 70% (v/v) CD_3CN and 1.2% (v/v) DMSO-d_6.

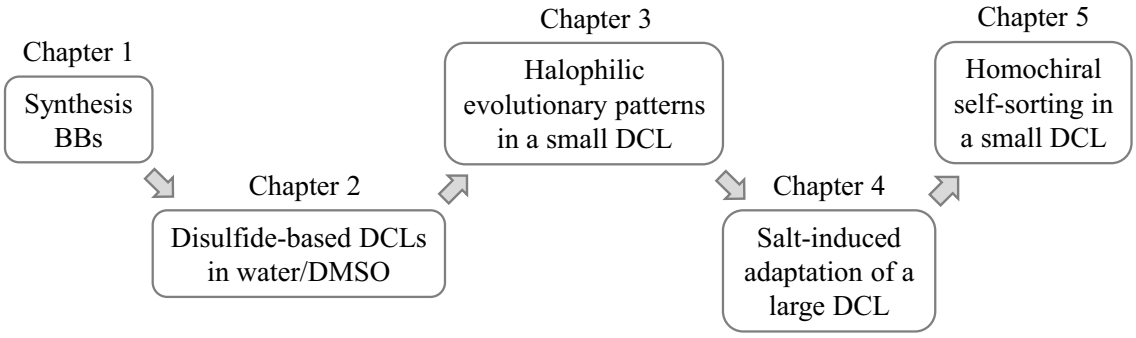
5.6. References

(1) Gal, J.; Cintas, P. In *Biochirality*; Cintas, P., Ed.; Springer Berlin Heidelberg: 2013; Vol. 333, p 1.
(2) Hein, J.; Gherase, D.; Blackmond, D. In *Biochirality*; Cintas, P., Ed.; Springer Berlin Heidelberg: 2013; Vol. 333, p 83.
(3) Hein, J. E.; Blackmond, D. G. *Acc. Chem. Res.* **2012**, *45*, 2045.
(4) Kojo, S. *Symmetry* **2010**, *2*, 1022.
(5) Blackmond, D. G. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a002147.
(6) Weissbuch, I.; Illos, R. A.; Bolbach, G.; Lahav, M. *Acc. Chem. Res.* **2009**, *42*, 1128.
(7) Kuhn, H. *Curr. Opin. Colloid Interface Sci.* **2008**, *13*, 3.
(8) Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. *Chem. Rev.* **2013**, *114*, 285.
(9) Bada, J. L. *Chem. Soc. Rev.* **2013**, *42*, 2186.
(10) Szostak, J. W. *Nature* **2009**, *459*, 171.
(11) Pross, A. *Origins Life Evol. B.* **2004**, *34*, 307.
(12) Wong, J. T.-F.; Lazcano, A. *Prebiotic evolution and astrobiology*; Landes Bioscience, 2009.
(13) Pizzarello, S. *Chem. Biodivers.* **2007**, *4*, 680.
(14) Pizzarello, S.; Cronin, J. R. *Geochim. Cosmochim. Acta* **2000**, *64*, 329.
(15) Cronin, J. R.; Pizzarello, S. *Adv. Space Res.* **1999**, *23*, 293.
(16) Brewer, A.; Davis, A. P. *Nat. Chem.* **2014**, *6*, 569.
(17) Calvin, M. *Chemical evolution: molecular evolution towards the origin of living systems on the earth and elsewhere*; Oxford University Press, 1969.
(18) Frank, F. C. *Biochim. Biophys. Acta* **1953**, *11*, 459.
(19) Steendam, R. R. E.; Verkade, J. M. M.; van Benthem, T. J. B.; Meekes, H.; van Enckevort, W. J. P.; Raap, J.; Rutjes, F. P. J. T.; Vlieg, E. *Nat. Commun.* **2014**, *5*.
(20) Dressel, C.; Reppe, T.; Prehm, M.; Brautzsch, M.; Tschierske, C. *Nat. Chem.* **2014**, *6*, 971.
(21) Klussmann, M.; Iwamura, H.; Mathew, S. P.; Wells, D. H.; Pandya, U.; Armstrong, A.; Blackmond, D. G. *Nature* **2006**, *441*, 621.
(22) Blackmond, D. G.; McMillan, C. R.; Ramdeehul, S.; Schorm, A.; Brown, J. M. *J. Am. Chem. Soc.* **2001**, *123*, 10103.
(23) Shibata, T.; Morioka, H.; Hayase, T.; Choji, K.; Soai, K. *J. Am. Chem. Soc.* **1996**, *118*, 471.
(24) Soai, K.; Shibata, T.; Morioka, H.; Choji, K. *Nature* **1995**, *378*, 767.
(25) Voshell, S. M.; Lee, S. J.; Gagné, M. R. *J. Am. Chem. Soc.* **2006**, *128*, 12422.
(26) Chung, M.-K.; Hebling, C. M.; Jorgenson, J. W.; Severin, K.; Lee, S. J.; Gagné, M. R. *J. Am. Chem. Soc.* **2008**, *130*, 11819.
(27) Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. *Science* **2005**, *308*, 667.
(28) Bulos, F.; Roberts, S. L.; Furlan, R. L. E.; Sanders, J. K. M. *Chem. Commun.* **2007**, 3092.
(29) Corbett, P. T.; Tong, L. H.; Sanders, J. K. M.; Otto, S. *J. Am. Chem. Soc.* **2005**, *127*, 8902.
(30) Gonzalez-Alvarez, A.; Alfonso, I.; Gotor, V. *Chem. Commun.* **2006**, 2224.
(31) Ziach, K.; Jurczak, J. *Chem. Commun.* **2015**, *51*, 4306.
(32) Wu, A.; Isaacs, L. *J. Am. Chem. Soc.* **2003**, *125*, 4831.
(33) Safont-Sempere, M. M.; Fernández, G.; Würthner, F. *Chem. Rev.* **2011**, *111*, 5784.

(34) Destoop, I.; Xu, H.; Oliveras-Gonzalez, C.; Ghijsens, E.; Amabilino, D. B.; De Feyter, S. *Chem. Commun.* **2013**, *49*, 7477.
(35) Viedma, C. *Phys. Rev. Lett.* **2005**, *94*, 065504.
(36) Brock, C. P.; Schweizer, W. B.; Dunitz, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 9811.
(37) Kondepudi, D. K.; Kaufman, R. J.; Singh, N. *Science* **1990**, *250*, 975.
(38) Pasteur, L. *Recherches sur les relations qui peuvent exister entre la forme cristalline, la composition chimique et les sens de la polarisation rotatoire*; Impr. Bachelier, 1848.
(39) Kramer, R.; Lehn, J. M.; Marquis-Rigault, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5394.
(40) Telfer, S. G.; Yang, X.-J.; Williams, A. F. *Dalton T.* **2004**, 699.
(41) Sato, K.; Itoh, Y.; Aida, T. *Chem. Sci.* **2014**, *5*, 136.
(42) Chung, D. M.; Nowick, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 3062.
(43) Murguly, E.; McDonald, R.; Branda, N. R. *Org. Lett.* **2000**, *2*, 3169.
(44) ten Cate, A. T.; Dankers, P. Y. W.; Sijbesma, R. P.; Meijer, E. W. *J. Org. Chem.* **2005**, *70*, 5799.
(45) ten Cate, A. T.; Dankers, P. Y. W.; Kooijman, H.; Spek, A. L.; Sijbesma, R. P.; Meijer, E. W. *J. Am. Chem. Soc.* **2003**, *125*, 6860.
(46) Sun, J.; Bennett, J. L.; Emge, T. J.; Warmuth, R. *J. Am. Chem. Soc.* **2011**, *133*, 3268.
(47) Xu, D.; Warmuth, R. *J. Am. Chem. Soc.* **2008**, *130*, 7520.
(48) Jędrzejewska, H.; Wierzbicki, M.; Cmoch, P.; Rissanen, K.; Szumna, A. *Angew. Chem. Int. Ed.* **2014**, *53*, 13760.
(49) Schafer, L. L.; Tilley, T. D. *J. Am. Chem. Soc.* **2001**, *123*, 2683.
(50) Sisco, S. W.; Moore, J. S. *Chem. Sci.* **2014**, *5*, 81.
(51) Capasso, S. *Thermochim. Acta* **1996**, *286*, 41.
(52) Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S. *J. Chem. Soc., Chem. Commun.* **1992**, 919.
(53) Wright, H. T. *Protein Eng.* **1991**, *4*, 283.
(54) Geiger, T.; Clarke, S. *J. Biol. Chem.* **1987**, *262*, 785.
(55) McConnell, R.; Godwin, W.; Stanley, B.; Green, M. S. *J. Ark. Acad. Sci.* **1997**, *51*, 1.
(56) Bailey, W. F.; Monahan, A. S. *J. Chem. Educ.* **1978**, *55*, 489.
(57) Balakrishnan, S.; Easteal, A. *Aust. J. Chem.* **1981**, *34*, 943.
(58) Marcus, Y. *J. Phys. Org. Chem.* **2012**, *25*, 1072.
(59) Blandamer, M. J.; Burgess, J.; Engberts, J. B. F. N.; Sanchez, F. *Faraday Discuss. Chem. Soc.* **1988**, *85*, 309.
(60) Deakyne, C. A.; Meot-Ner , M.; Campbell, C. L.; Hughes, M. G.; Murphy, S. P. *J. Chem. Phys.* **1986**, *84*, 4958.
(61) Guillaume, Y. C.; Guinchard, C. *Anal. Chem.* **1997**, *69*, 183.
(62) Hafezi, N.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 12861.
(63) Rodriguez-Docampo, Z.; Eugenieva-Ilieva, E.; Reyheller, C.; Belenguer, A. M.; Kubik, S.; Otto, S. *Chem. Commun.* **2011**, *47*, 9798.
(64) Saur, I.; Scopelliti, R.; Severin, K. *Chem. Eur. J.* **2006**, *12*, 1058.
(65) Vial, L.; Sanders, J. K. M.; Otto, S. *New J. Chem.* **2005**, *29*, 1001.
(66) Otto, S.; Kubik, S. *J. Am. Chem. Soc.* **2003**, *125*, 7804.
(67) Sarma, R. J.; Otto, S.; Nitschke, J. R. *Chem. Eur. J.* **2007**, *13*, 9542.
(68) Braslavsky, S. E. *Pure Appl. Chem.* **2007**, *79*, 293.
(69) Gagliardi, L. G.; Castells, C. B.; Ràfols, C.; Rosés, M.; Bosch, E. *J. Chem. Eng. Data* **2007**, *52*, 1103.

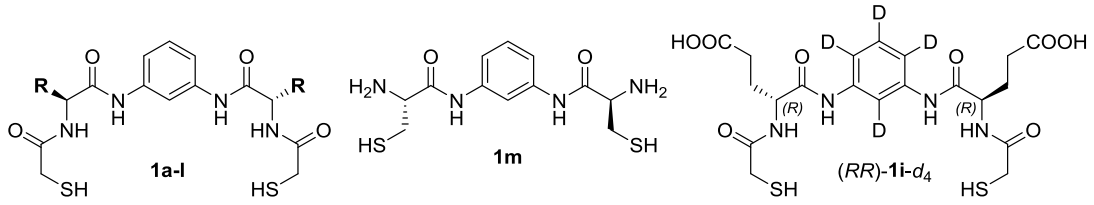
(70) Catalán, J.; Diaz, C.; Garcia-Blanco, F. *Org. Biomol. Chem.* **2003**, *1*, 575.
(71) Ulatowski, F.; Sadowska-Kuziola, A.; Jurczak, J. *J. Org. Chem.* **2014**, *79*, 9762.
(72) Atcher, J.; Alfonso, I. *RSC Adv.* **2013**, *3*, 25605.
(73) In *Azeotropic Data-III*; AMERICAN CHEMICAL SOCIETY: 1973; Vol. 116, p 1.
(74) Adamson, J.; Lajoie, G. In *PEPTIDES-AMERICAN SYMPOSIUM-*; ESCOM SCIENCE PUBLISHERS: 1994; Vol. 13, p 44.
(75) Watlaufer, D. B.; Malik, S. K.; Stoller, L.; Coffin, R. L. *J. Am. Chem. Soc.* **1964**, *86*, 508.
(76) Cope, A. C.; Mehta, A. S. *J. Am. Chem. Soc.* **1963**, *85*, 1949.
(77) Blanco, C.; Hochberg, D. *Phys. Chem. Chem. Phys.* **2011**, *13*, 839.
(78) Plasson, R.; Brandenburg, A. *Origins Life Evol. B.* **2010**, *40*, 93.
(79) Plasson, R.; Bersini, H. *J. Phys. Chem. B* **2009**, *113*, 3477.
(80) Julian, R. R.; Myung, S.; Clemmer, D. E. *J. Phys. Chem. B* **2005**, *109*, 440.
(81) Ben-Naim, A. *J. Chem. Educ.* **2011**, *88*, 594.
(82) Lin, S.-K. *J. Chem. Inform. Comput. Sci.* **1996**, *36*, 367.
(83) Estrada, E.; Avnir, D. *J. Am. Chem. Soc.* **2003**, *125*, 4368.
(84) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652.
(85) Hwang, T. L.; Shaka, A. J. *J. Magn. Reson., Ser A* **1995**, *112*, 275.

GENERAL CONCLUSIONS



The detailed study of the sequence of topics listed above has led to the following general conclusions:

1) Fourteen new pseudopeptidic dithiols (**1a-m** and (*RR*)-**1i**-d_4) have been designed, synthesized and fully characterized in order to be used as bipodal building blocks to generate all the disulfide-based DCLs studied in this thesis. The design is based on a C_2-symmetric scaffold consisting of a central *m*-phenylenediamine chromophore that rigidly joins two identical arms, each formed by an amino acid residue. Additionally, compounds **1a-l** and (*RR*)-**1i**-d_4 incorporate a mercaptoacetyl moiety attached at the *N*-termini of the two amino acids. The pseudopeptidic nature provides the building blocks with pertinent peptide-like information, differently charged functional groups and chiral information.



2) Suitable experimental conditions for the generation of useful disulfide-based DCLs have been successfully developed. The use of DMSO as co-solvent in aqueous mixtures has shown several beneficial effects. Firstly, it increases the solubility of the building blocks, allowing the use of water insoluble non-charged species. Secondly, it promotes thiol oxidation, highly reducing the reaction time for the disulfide formation. Finally, it accelerates the disulfide exchange, allowing the systems to fully equilibrate even at slightly acidic pH values.

3) A minimalistic DCL has been successfully used to reproduce adaptive trends described for the evolution of biological systems. Thus, the addition of salt to a dynamic library of macrocyclic pseudopeptides induces the amplification of those

species concentrating anionic amino acids, with the Asp derivatives showing a better salt-adaptation than the Glu counterparts. Structural studies, including molecular modeling, NMR and CD experiments, suggest a folded conformation for the amplified species and reveal the better salt-adaptation of those species showing a smaller accessible surface area. The adaptive process is driven by the increase of the ionic strength and has a remarkable resemblance with the natural evolution of the proteins of halophilic organisms for surviving in hypersaline media.

4) The effect of increasing the ionic strength has also been studied in a larger DCL consisting of 21 differently charged dimeric macrocycles. The salt-induced adaptation of this complex system has been characterized in a top-down fashion by the dynamic deconvolution into the minimal components, and by performing structural studies for selected dimers. In the absence of salt, the library optimizes the electrostatic interactions showing a composition markedly different from the statistical distribution. The increase of the ionic strength shields these interactions and the system approaches the statistically favored proportion. Consequently, the salt-response of the members of the library can be classified in different families attending to the charges. The behavior of each member is determined by a combination of its structural information and the co-adaptive relationships with the other members of the complex network.

5) A simple DCL consisting of homo- and heterochiral dimeric pseudopeptides has successfully been used to study the chiral self-sorting phenomenon. In buffered water, the composition of the library is slightly biased in favor of the homochiral species. A decrease in the polarity of the medium favors the homochiral self-sorting, suggesting that the selection process is driven by polar intramolecular interactions. Additionally, the homochiral selectivity also increases with the temperature, indicating a positive entropic contribution. Preliminary NMR experiments suggest significantly different conformations for the homo- and heterochiral species, and further structural studies are foreseen in order to fully understand the self-sorting process at the molecular level.

Overall, the results comprehended in this thesis represent a contribution to expand the use of the dynamic combinatorial libraries beyond the frontiers of chemistry. The adaptive nature of these dynamic systems, together with a suitable bio-inspired design,

have demonstrated to allow the minimalistic experimental modeling of different processes of biological interest like the natural evolution of the halophilic proteins, the co-adaptive relationships in a complex network and the homochiral self-sorting phenomenon. In this regard, dynamic combinatorial chemistry has a large scope for future development, with foreseen implication in the understanding of some relevant scientific problems such as the origin of homochirality and life.

RESUM EN CATALÀ

INTRODUCCIÓ GENERAL

El terme "química de sistemes", introduït l'any 2005 per von Kiedrowski,[1] es refereix a un camp emergent de la química centrat en l'estudi de sistemes moleculars complexos. En aquests sistemes, diversos compostos reaccionen i interaccionen entre ells donant lloc a una xarxa molecular capaç de presentar propietats emergents, és a dir, propietats que no són simplement el resultat de la suma dels atributs de cadascun dels components.

Un dels punts de partida del desenvolupament de la química de sistemes ha estat la química combinatòria dinàmica (DCC per les seves sigles en anglès).[2-4] Aquest tipus de química va ser introduït a meitat de la darrera dècada del segle XX amb la finalitat d'unificar dues idees fonamentals: i) la química combinatòria com a estratègia per generar diversitat estructural, i ii) l'evolució molecular, que planteja la possibilitat que els sistemes químics siguin capaços de seleccionar, amplificar i inclús autocorregir les seves estructures i propietats com a resposta a l'acció d'un estímul extern. La DCC inicialment va ser desenvolupada en el marc de la química supramolecular[5-6] com a metodologia per a la identificació i preparació de receptors moleculars sintètics.[7] Amb aquesta finalitat, la DCC proposa crear una quimioteca de compostos (DCL de "dynamic combinatorial library") interconnectats per mitjà de processos químics reversibles. El constant intercanvi de blocs constituents entre els membres de la quimioteca fa que el sistema arribi a una situació d'equilibri termodinàmic en la qual la concentració de cadascun dels compostos oligomèrics formats, *i.e.* els diferents membres de la quimioteca, està determinada per la seva estabilitat relativa.

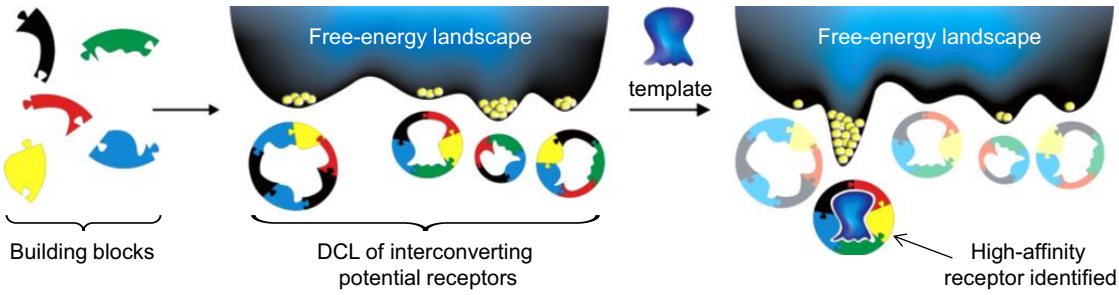


Figura R.1. Representació esquemàtica d'una DCL amb el seu perfil d'energia lliure, mostrant com la interacció d'una molècula diana pot conduir a la selecció i amplificació dels receptors que presenten una major afinitat per la plantilla (Figura extreta de la referència[7]).

En presència d'una plantilla molecular (molècula diana) que actua com a estímul extern, la interacció d'alguns dels membres de la mescla amb la plantilla fa que el sistema dinàmic reorganitzi la seva distribució augmentant la concentració dels compostos que interaccionen de manera més eficaç amb la plantilla (Figura R.1). D'aquesta manera s'aconsegueix la selecció i amplificació dels membres de la quimioteca que formen una estructura supramolecular més estable amb la molècula diana.

Si bé l'aplicació que inicialment va motivar el desenvolupament de la DCC va ser la identificació i preparació de receptors moleculars sintètics, actualment la química combinatòria dinàmica també s'utilitza amb moltes altres finalitats. Entre aquestes destaquen la síntesi de lligands per biomolècules,[8-10] de gàbies moleculars[11-12] i de sensors.[13-14] Recentment, la DCC també ha estat utilitzada en l'estudi de sistemes multifase,[15-16] en catàlisis,[17-18] en química de superfícies[19-20] i en la síntesi d'estructures enllaçades mecànicament.[21-24]

CAPÍTOL 1

Introducció i objectius

La vida es basa en sistemes químics molt complexos formats per un nombre limitat d'unitats estructurals.[25] Els aminoàcids constitueixen un d'aquests elements estructurals bàsics i són presents en un gran nombre de biomolècules. La combinació de tan sols vint aminoàcids diferents formant cadenes polimèriques permet generar un gran nombre de pèptids i proteïnes amb una gran varietat de funcions biològiques.[26] Fascinats i inspirats per les propietats d'aquestes biomolècules, els investigadors han desenvolupat pèptids abiòtics amb la finalitat d'aconseguir nous compostos amb propietats similars a les dels sistemes naturals.

Una estratègia habitual a l'hora de dissenyar molècules bioinspirades consisteix a combinar unitats estructurals naturals amb fragments abiòtics. Els pseudopèptids són molècules que combinen aminoàcids naturals amb altres elements estructurals sintètics, i han estat utilitzats per a un gran nombre d'aplicacions.[27] Exemples representatius són el desenvolupament de catalitzadors,[28-30] de compostos amb activitat biològica[31-33] i de receptors moleculars.[11,34-35]

L'objectiu principal d'aquest Capítol és dissenyar, sintetitzar i caracteritzar un seguit de noves molècules per tal d'utilitzar-les com a blocs constituents per preparar quimioteques dinàmiques. L'estructura química d'aquests nous compostos ha de reunir un seguit de requisits. Primerament, les molècules han d'incorporar grups funcionals que permetin la unió reversible dels blocs constituents. A més, la presència d'un cromòfor és desitjable per facilitar la corresponent anàlisi quantitativa. Per últim, tenint en compte les potencials aplicacions d'aquests compostos, el corresponent esquelet molecular hauria d'estar proveït de tres propietats o característiques fonamentals: i) informació estructural de tipus peptídic, ii) grups funcionals amb diferents càrregues, i iii) informació quiral.

Resultats i discussió

Es van dissenyar i sintetitzar tretze ditiols pseudopeptídics (**1a-m**, Figura R.2a). Els compostos **1a-l** tenen un mateix esquelet estructural, mentre que el compost **1m** presenta una estructura lleugerament més simple. Totes les molècules tenen simetria C_2 i consten d'una *m*-fenilendiamina central que uneix dos braços idèntics, cadascun d'ells

format per un aminoàcid. Unit al nitrogen d'aquest aminoàcid, els compostos **1a-l** incorporen el substituent mercaptoacetil.

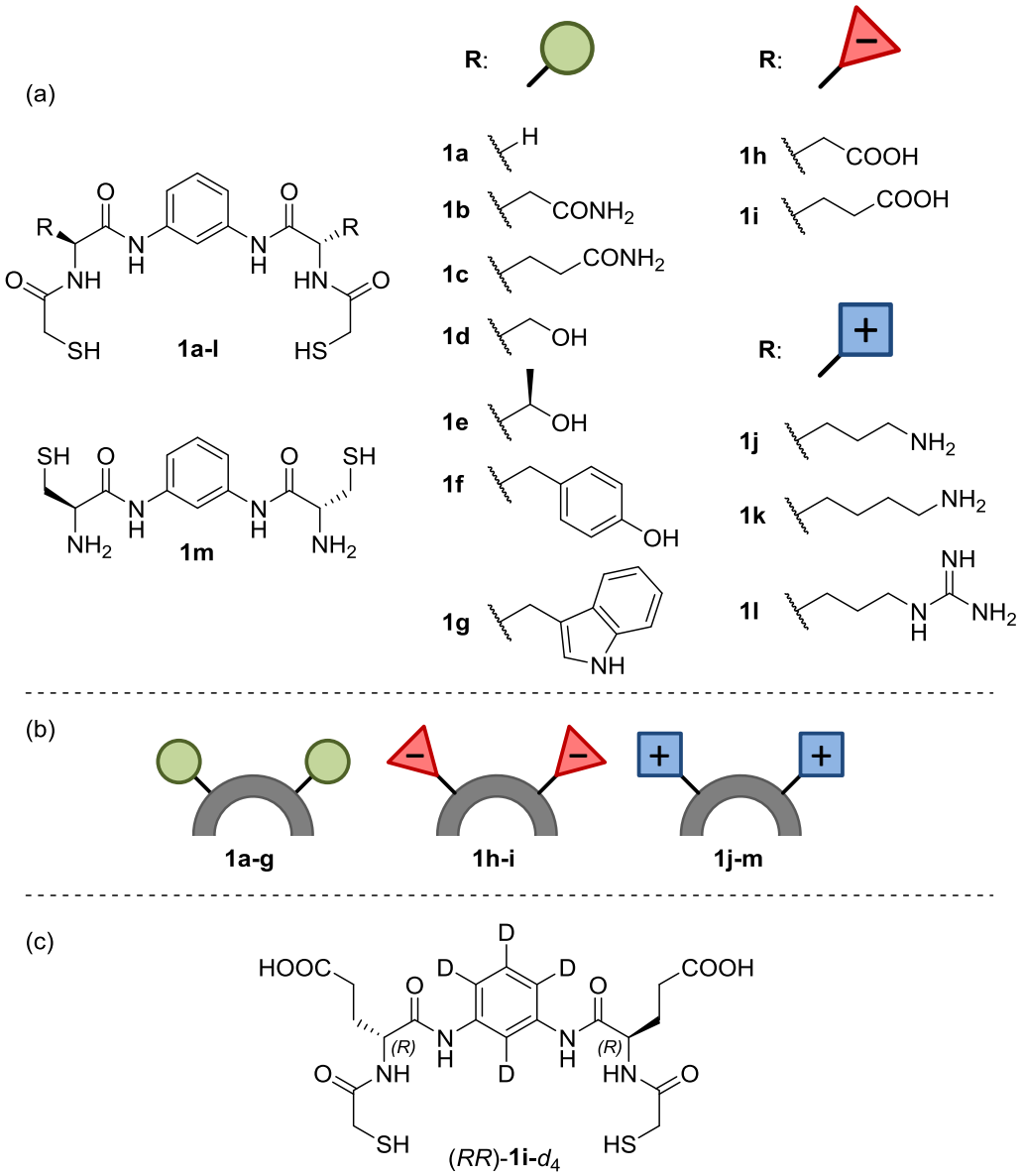


Figura R.2. Estructura química (a) i representació esquemàtica (b) dels tretze blocs constituents **1a-m**. (c) Estructura química de (*RR*)-**1i**-d_4, l'enantiòmer de **1i** marcat isotòpicament.

La funció de la *m*-fenilendiamina central és unir rígidament els dos braços de la molècula tot preorganitzant-la conformacionalment per formar estructures macrocícliques. La rigidesa de l'anell aromàtic força la separació dels dos braços de la molècula i dificulta la formació del corresponent producte d'autociclació. A més, la *m*-fenilendiamina també actua de cromòfor, facilitant l'anàlisi quantitativa de les quimioteques.

Els dos aminoàcids α de cadascun dels compostos **1a-m** actuen de font de diversitat estructural tot aportant diferents cadenes laterals (grups R, Figura R.2a). Així, el compost **1a** conté l'aminoàcid glicina i no incorpora cap cadena lateral; els compostos **1b-e** contenen els aminoàcids asparagina, glutamina, serina i treonina respectivament, i incorporen cadenes laterals polars no carregades; els compostos **1f-g** contenen els aminoàcids tirosina i triptòfan respectivament, i incorporen cadenes laterals hidrofòbiques; els compostos **1h-i** contenen els aminoàcids àcid aspàrtic i àcid glutàmic respectivament, i incorporen cadenes laterals carregades negativament a pH neutre; els compostos **1j-l** contenen els aminoàcids ornitina, lisina i arginina respectivament, i incorporen cadenes laterals carregades positivament a pH neutre; i finalment el compost **1m** conté l'aminoàcid cisteïna, i incorpora el grup funcional tiol. En aquest últim cas, l'absència del substituent mercaptoacetil fa que el compost **1m** tingui dues amines lliures.

Els dos aminoàcids de cadascun dels compostos **1a-m**, lluny de ser simplement la font de diversitat estructural, en realitat són la part més important pel que fa a les aplicacions per a les quals aquests blocs constituents es van dissenyar. La presència d'aquest motiu estructural fa que les molècules **1a-m** continguin informació estructural de tipus peptídic, fent que siguin molècules adients per a la preparació de DCLs bioinspirades (Capítol 3). Tal i com es mostra en la representació esquemàtica de la Figura R.2b, els grups funcionals de les cadenes laterals fan que, a pH neutre, els compostos **1a-m** tinguin diferents càrregues. Aquestes càrregues permeten l'estudi de l'efecte de les interaccions electrostàtiques en la composició de les DCLs (Capítol 4). Per últim, la presència d'aminoàcids fa que cadascun dels compostos **1b-m** estigui dotat de dos centres estereogènics. La quiralitat dels blocs constituents permet l'estudi de fenòmens directament relacionats amb les propietats estereoisomèriques dels membres de les DCLs (Capítol 5). Pel que fa a aquesta última aplicació, un últim compost ((*RR*)-**1i**-d_4, Figura R.2c) va ser inclòs a la llista de molècules objectiu. Aquest catorzè bloc constituent és la imatge especular del compost **1i** i conté dos àcids glutàmics amb configuració D. A més, incorpora una *m*-fenilendiamina marcada isotòpicament.

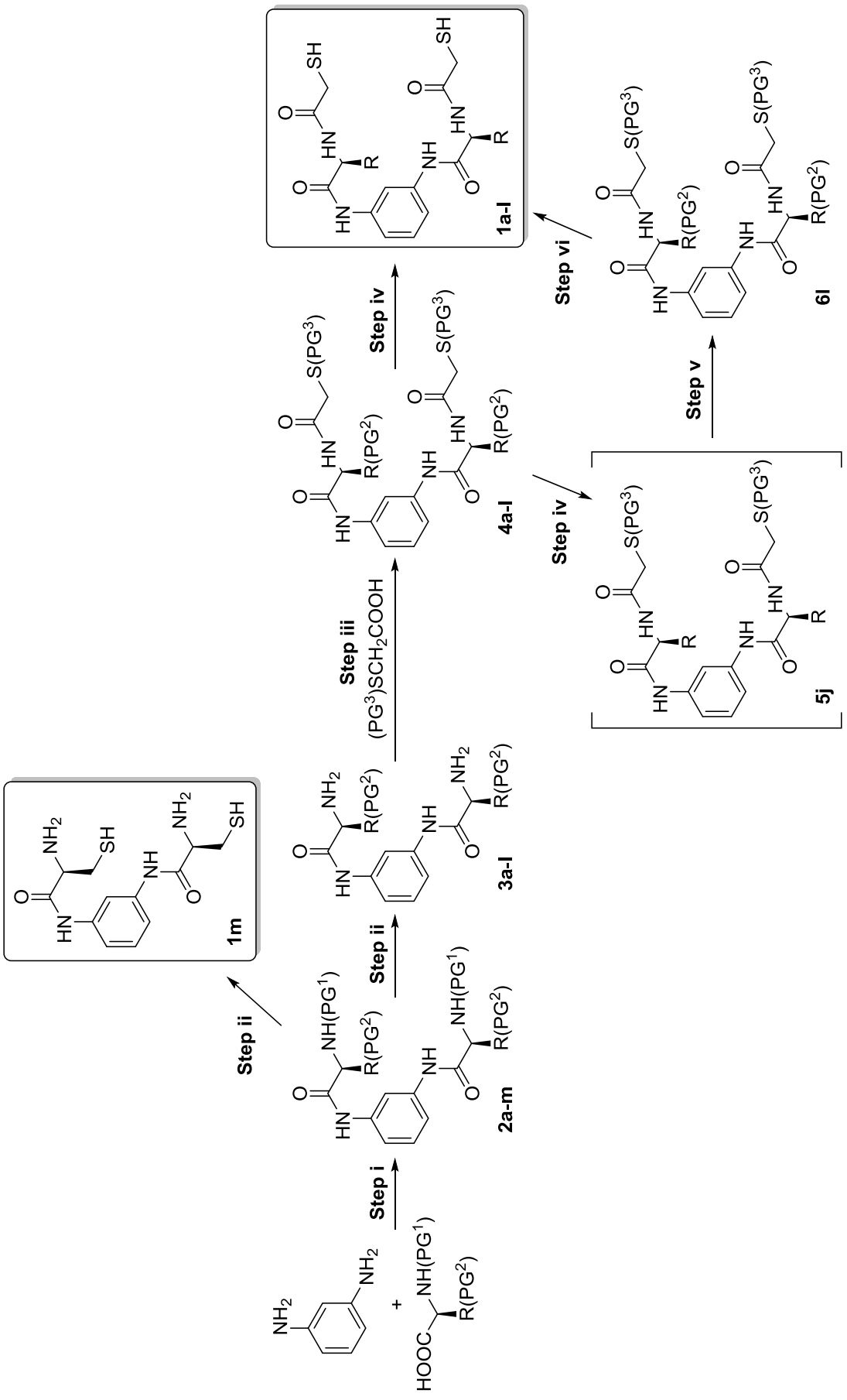
El substituent mercaptoacetil unit a cadascun dels dos aminoàcids dels compostos **1a-l** serveix per proveir les molècules dissenyades de dos grups tiol. Aquests permeten la unió reversible dels blocs constituents mitjançant la formació d'enllaços disulfur. En

el cas concret del compost **1m**, l'aminoàcid cisteïna ja incorpora el grup funcional tiol, de manera que el substituent mercaptoacetil no és necessari.

L'anàlisi retrosintètica dels ditiols **1a-m** és molt simple i consisteix a separar les tres parts de l'esquelet molecular (la *m*-fenilendiamina, els dos aminoàcids i els dos substituents mercaptoacetil). Així doncs, l'esquema retrosintètic comporta el trencament de tots els enllaços amida de les molècules i delimita el "problema" sintètic a la formació consecutiva d'aquest tipus d'unió covalent. La metodologia sintètica per formar poliamides mitjançant l'acoblament consecutiu d'amines i àcids carboxílics ha estat àmpliament investigada i desenvolupada en el camp de la síntesi de pèptids.[36]

La ruta sintètica proposada per a la síntesi dels blocs constituents **1a-m** consisteix a fer créixer simultàniament els dos braços de la molècula i requereix la utilització de tres grups protectors (PGs de "protecting grups"): PG^1, PG^2 i PG^3. Tal com es mostra en l'Esquema R.1, en un primer pas sintètic la *m*-fenilendiamina és diacilada amb dues unitats del corresponent aminoàcid degudament protegit, obtenint els intermedis **2a-m**. Pel cas concret de la síntesi del compost que conté l'aminoàcid cisteïna, en un segon i últim pas sintètic els grups protectors PG^1 i PG^2 són eliminats, obtenint el producte final **1m**. Per la resta de compostos, la ruta sintètica continua amb l'eliminació selectiva del grup protector PG^1, obtenint les amines lliures **3a-l**. Seguidament, aquestes són diacilades amb dues unitats d'àcid mercaptoacètic degudament protegit, obtenint els intermedis **4a-l**. En un quart i últim pas sintètic, els grups PG^2 i PG^3 són simultàniament eliminats per finalment obtenir els ditiols **1a-k**. Malauradament, aquest últim pas de desprotecció no va ser possible pel bloc constituent que conté l'aminoàcid arginina (**1l**), fent que per a la síntesi d'aquest compost fos necessari idear una ruta sintètica alternativa a través dels intermedis **5j** i **6l**.

El compost (*RR*)-**1i**-d_4 es va sintetitzar de la mateixa manera que el seu enantiòmer **1i**, però utilitzant diferents materials de partida: àcid glutàmic amb configuració D i *m*-fenilendiamina-2,4,5,6-d_4.



Esquema R.1. Ruta sintètica per a la preparació dels blocs constituents **1a-m**.

Conclusions

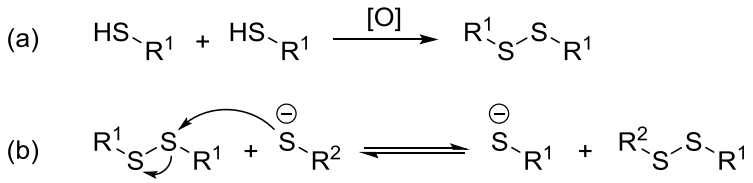
S'han dissenyat, sintetitzat i caracteritzat catorze ditiols amb estructura pseudopeptídica (**1a-m** i (*RR*)-**1i**-d_4) per tal de ser utilitzats com a blocs constituents per generar DCLs basades en la formació d'enllaços disulfur. L'estructura d'aquests compostos té un esquelet comú amb simetria C_2 i consta d'una *m*-fenilendiamina central que uneix dos braços idèntics, cadascun d'ells format per un aminoàcid. Unit al nitrogen d'aquest aminoàcid, els compostos **1a-l** i (*RR*)-**1i**-d_4 incorporen el substituent mercaptoacetil. La naturalesa pseudopeptídica dels blocs constituents sintetitzats fa que incorporin informació estructural de tipus peptídic, cadenes laterals amb diferents càrregues i informació quiral.

La síntesi d'aquests compostos s'ha dut a terme mitjançant una mateixa ruta sintètica que consta de quatre passos. Excepcionalment, els ditiols **1m** i **1l** s'han sintetitzat mitjançant rutes sintètiques de dos i sis passos respectivament.

CAPÍTOL 2

Introducció i objectius

Els enllaços disulfur, també anomenats ponts disulfur, són un tipus d'unió covalent molt important des del punt de vista biològic. Així, per exemple, en la química de pèptids i proteïnes aquest enllaç covalent té un paper cabdal en la formació d'estructures secundàries.[37-38] Generalment, els ponts disulfur s'obtenen a partir de l'oxidació irreversible de dos grups tiol (Esquema R.2a). Un cop formats, els disulfurs poden intercanviar-se entre ells mitjançant el procés reversible representat en la Figura R.2b. Tant la reacció de formació com la d'intercanvi de disulfurs necessiten la presència de quantitats catalítiques d'anions tiolat i, per tant, requereixen un pH lleugerament bàsic.



Esquema R.2. Processos d'oxidació de tiols (a) i d'intercanvi de disulfurs (b).

En nombrosos estudis aquests dos processos han estat utilitzats per generar DCLs en condicions de control termodinàmic. El mètode més emprat per preparar quimioteques dinàmiques de disulfurs consisteix a dissoldre els blocs constituents, *i.e.* els tiols, en aigua a pH lleugerament bàsic, mantenint la dissolució en contacte amb l'oxigen atmosfèric. En aquestes condicions, la formació de disulfurs sol ser bastant lenta i sovint són necessaris diversos dies per completar el procés d'oxidació. El requisit imprescindible per tal que la DCL resultant arribi a la situació d'equilibri termodinàmic és que la reacció d'intercanvi (reversible, Esquema R.2b) sigui més ràpida que la reacció d'oxidació (irreversible, Esquema R.2a).

L'objectiu principal d'aquest Capítol és trobar unes condicions experimentals en les quals els blocs constituents sintetitzats en el Capítol 1 puguin generar DCLs útils en condicions de control termodinàmic. La nostra hipòtesi és que la incorporació de dimetilsulfòxid (DMSO) en mescles aquoses serà beneficiosa tant pel que fa a la solubilitat dels membres de les DCLs com pel que fa a les propietats cinètiques i termodinàmiques dels processos de formació i intercanvi de disulfurs.

Resultats i discussió

Primerament es va estudiar l'efecte del DMSO en la cinètica del procés d'oxidació mitjançant el test d'Ellman.[39] Per fer-ho, es van escollir tres compostos amb grups funcionals representatius en les seves cadenes laterals: grups neutres (**1d**), grups àcids (**1h**) i grups bàsics (**1j**). De tots ells se'n va estudiar la velocitat d'oxidació a diferents valors de pH i a diferents percentatges de DMSO (*e.g.* oxidació de **1h**, Figura R.3).

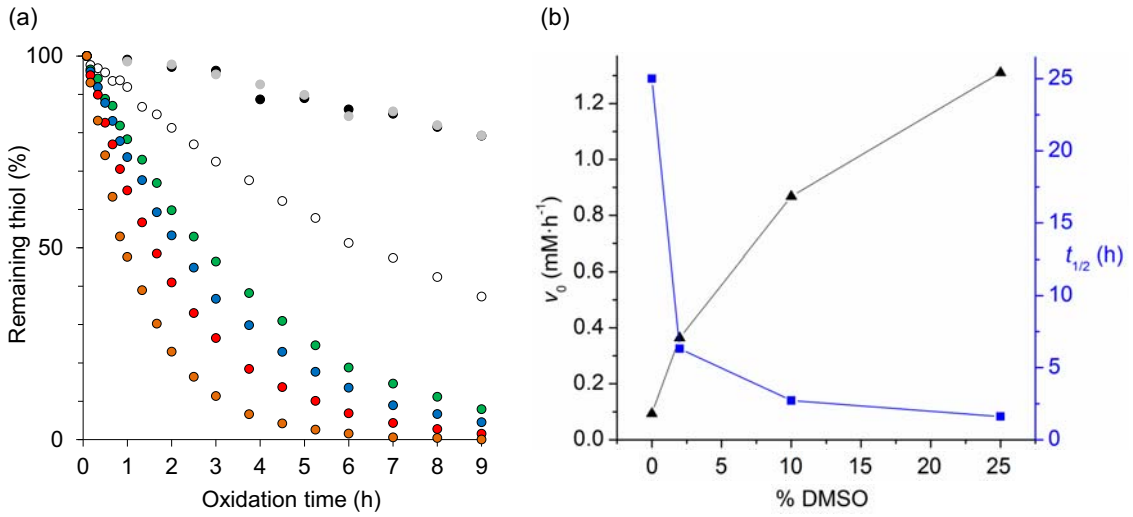
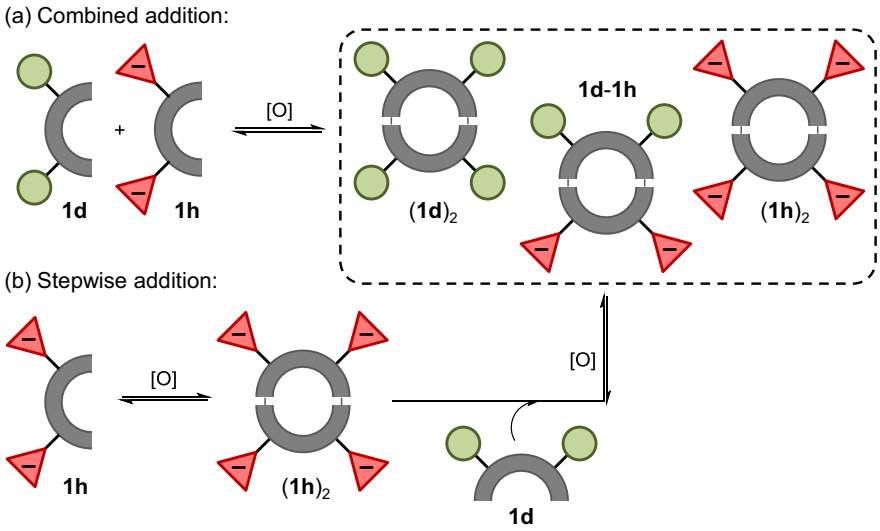


Figura R.3. (a) Representació de la quantitat de tiols restant (%) en funció del temps d'oxidació (h), pel procés d'oxidació de **1h** (2 mM) en diferents condicions de reacció: 0% DMSO a pH 7,5 (●); 100% DMSO (●); 2% (v/v) DMSO a pH 7,5 (○); 10% (v/v) DMSO a pH 7,5 (●); 10% (v/v) DMSO a pH 2.5 (●); 25% (v/v) DMSO a pH 7,5 (●); i 25% (v/v) DMSO a pH 2,5 (●). (b) Representació de la velocitat inicial (v_0, mM·h^{-1}, eix de l'esquerra, triangles negres) i el temps de vida mitja ($t_{1/2}$, h, eix de la dreta, quadrats blaus) en funció del contingut de DMSO, pel procés d'oxidació de **1h** (2 mM) a pH 7,5.

Es va veure que en mescles d'aigua/DMSO la reacció d'oxidació és més ràpida que en qualsevol d'aquests dos dissolvents purs i que, en l'interval 0%-25% (v/v) de DMSO, la velocitat de formació de disulfurs augmenta gradualment amb la proporció del codissolvent orgànic. Així, la presència de tan sols un 2% (v/v) de DMSO fa que la velocitat d'oxidació sigui quatre vegades més ràpida que en medi 100% aquós i la presència d'un 10% (v/v) de DMSO permet completar el procés de formació de disulfurs en poc més de 24 hores. Pel que fa a l'efecte del pH, en aigua l'oxidació només té lloc si el pH és lleugerament bàsic. En canvi, quan el dissolvent incorpora un 2%-25% (v/v) de DMSO, la reacció d'oxidació té lloc fins i tot a pH 2,5. La presència de cadenes laterals amb grups carregats negativament (**1h**) o positivament (**1j**) té un efecte menyspreable en la velocitat de formació de disulfurs.

La capacitat d'arribar a l'equilibri termodinàmic es va avaluar mitjançant DCLs generades a partir de tan sols dos blocs constituents (mescles binaries). En l'exemple representat en l'Esquema R.3a, els blocs constituents **1d** i **1h** es van oxidar junts, obtenint una mescla de disulfurs macrocíclics amb tres espècies majoritàries: els homodímers (**1d**)$_2$ i (**1h**)$_2$, i l'heterodímer **1d-1h**. Paral·lelament, una DCL similar es va preparar afegint el compost **1d** a una mostra de **1h** prèviament oxidada (Esquema R.3b). Si el procés d'intercanvi de disulfurs és més ràpid que la reacció d'oxidació, les DCLs obtingudes mitjançant els dos mètodes representats en l'Esquema R.3 haurien de presentar la mateixa composició, indicant que el sistema ha arribat a la situació l'equilibri termodinàmic. Si pel contrari el procés d'intercanvi és massa lent comparat amb la reacció d'oxidació, llavors les dues DCLs haurien de presentar una composició diferent, indicant que el sistema no ha arribat a la situació d'equilibri.



Esquema R.3. Representació esquemàtica del test de reversibilitat.

Aquest test de reversibilitat es va fer amb les parelles de blocs constituents **1d**+**1h** i **1d**+**1j**, a diferents percentatges de DMSO (10% i 25%) i a diferents pHs (2,5-6,5). L'anàlisi quantitativa de la composició final de les mescles va servir per determinar que, independentment de la proporció de DMSO i dels blocs constituents utilitzats, les mescles preparades a pH ≥ 4,5 arriben a l'equilibri termodinàmic.

Conclusions

S'han trobat condicions experimentals adequades per generar DCLs de disulfurs a partir de l'oxidació dels blocs constituents sintetitzats al Capítol 1. En aquest sentit, l'ús de DMSO com a codissolvent orgànic en mescles aquoses ha demostrat tenir una sèrie d'efectes beneficiosos. En primer lloc, el DMSO incrementa la solubilitat dels blocs constituents, permetent l'ús d'espècies no solubles en aigua. En segon lloc, la presència d'aquest codissolvent accelera l'oxidació dels tiols, reduint el temps de reacció necessari per formar les DCLs. Finalment, el DMSO accelera la reacció d'intercanvi de disulfurs, fent que els sistemes puguin arribar a l'equilibri termodinàmic fins i tot a valors de pH lleugerament àcids.

CAPÍTOL 3

Introducció i objectius

Els microorganismes halòfils són un tipus d'extremòfils que viuen en ambients amb concentracions salines molt elevades.[40] Per tal de reduir la pressió osmòtica exterior, aquests microorganismes incrementen la concentració de sal dins del citoplasma,[41] de manera que les seves biomolècules s'han hagut d'adaptar a unes condicions intracel·lulars d'extrema salinitat. Les proteïnes d'aquests extremòfils, les anomenades proteïnes halòfiles, han evolucionat tot maximitzant la seva estabilitat i activitat en ambients hipersalins.[42-43] Fruit d'aquest procés evolutiu, aquestes proteïnes presenten un seguit de modificacions estructurals respecte les seves homòlogues no halòfiles.[44-45] Els canvis estructurals afecten principalment a la composició d'aminoàcids superficials i es resumeixen en quatre tendències principals: i) increment del contingut d'aminoàcids àcids, ii) disminució del nombre de lisines, iii) disminució del contingut hidrofòbic, i iv) disminució de la superfície accessible al dissolvent (ASA de "accessible surface area").

L'objectiu d'aquest Capítol és investigar la capacitat potencial d'una DCL minimalista d'adaptar-se a un determinat estímul extern de manera similar o paral·lela al procés evolutiu d'un sistema biològic molt més complex. La nostra hipòtesi és que, com a resposta a l'increment de la concentració salina, una DCL generada a partir de blocs constituents amb informació estructural de tipus peptídic mostrarà tendències adaptatives similars als canvis estructurals descrits per l'evolució natural de les proteïnes halòfiles.

Resultats i discussió

Es va dissenyar una DCL minimalista generada a partir de la mescla de tres blocs constituents: el que conté àcid glutàmic (**1i**), aniònic a pH neutre, el que conté glutamina (**1c**), amb una estructura equivalent a l'anterior però sense càrrega, i el que conté serina (**1d**), com a bloc constituent innocent i no carregat. La quimioteca es va preparar dissolent quantitats equimolars dels tres compostos en un tampó aquós de fosfat (pH 7,5) amb un 25% (v/v) de DMSO. Un cop equilibrada, la DCL es va analitzar mitjançant HPLC-UV (Figura R.4a) i UPLC-MS, i es va veure que la mescla generada consta de tots els dímers possibles (sis), juntament amb petites quantitats de tots els trímers possibles (deu).

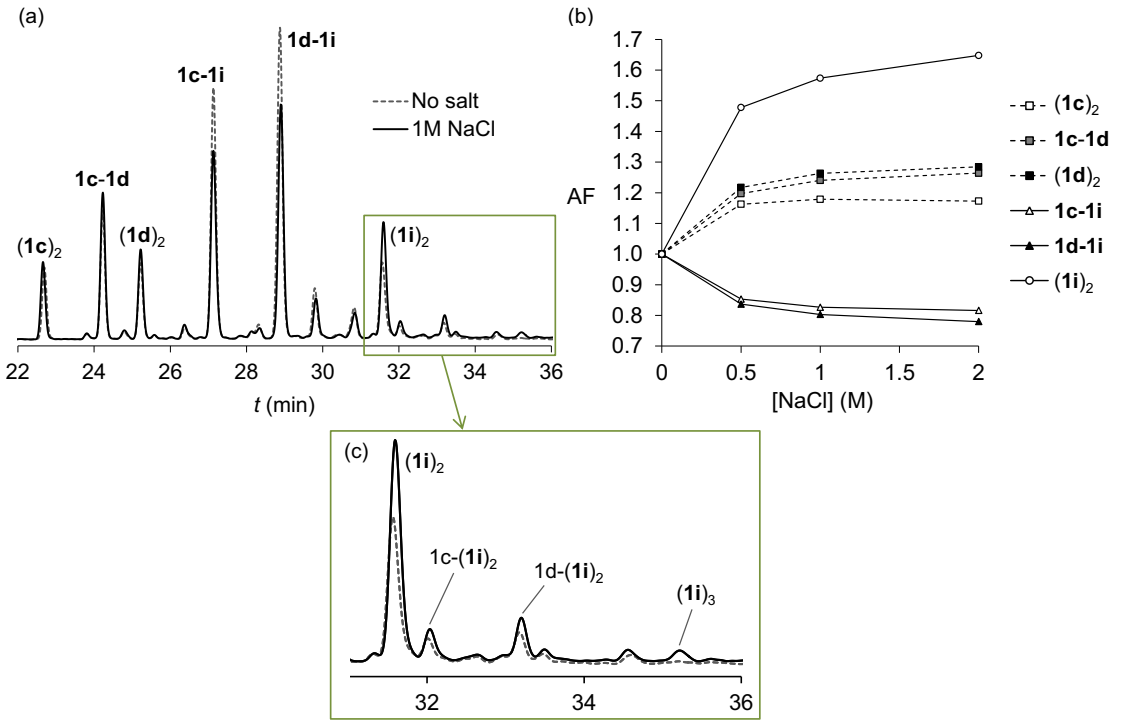


Figura R.4. (a) Perfils HPLC-UV (254 nm) de la DCL generada a partir de la mescla de **1c**+**1d**+**1i** (2 mM de cadascun) en tampó aquós de fosfat (pH 7,5) amb un 25% (v/v) de DMSO en absència de sal (línia discontínua de color gris) i en presència de NaCl 1,0 M (línia contínua de color negre); (b) representació del factor d'amplificació dels sis dímers en funció de la concentració de NaCl; i (c) ampliació de (a).

Paral·lelament, la mateixa DCL es va preparar afegint-hi diferents quantitats de NaCl (0,5-2,0 M). Els canvis en la composició de la quimioteca provocats per aquest estímul extern van ser quantificats mitjançant el factor d'amplificació (AF de "amplification factor"),[46] i el patró d'amplificacions obtingut en representar els valors d'AF dels sis dímers en funció de la concentració salina (Figura R.4b) va permetre identificar un important increment de l'estabilitat relativa del dímer (**1i**)$_2$. A continuació es van estudiar les característiques estructurals d'aquest membre de la quimioteca mitjançant modelatge molecular, ressonància magnètica nuclear i dicroisme circular. Es va concloure que diversos ponts d'hidrogen intramoleculars propicien que aquest homodímer presenti una estructura plegada en dissolució. A més, es va observar que les característiques conformacionals d'aquest membre de la quimioteca no es veuen afectades per l'increment de la salinitat del medi.

Pel que fa als trímers de la quimioteca, tots aquells que concentren més d'una unitat del bloc constituent **1i** (*i.e.* **1c**-(**1i**)$_2$, **1d**-(**1i**)$_2$ i (**1i**)$_3$) també són amplificats en presència de l'estímul extern (Figura R.4c). Per tant, en general, el sistema respon a l'augment de la salinitat del medi seleccionant aquells membres de la quimioteca que agrupen blocs

constituents **1i**, és a dir, aquells membres que concentren àcids glutàmics. Aquesta tendència adaptativa té una clara similitud amb el procés evolutiu descrit per les proteïnes halòfiles.

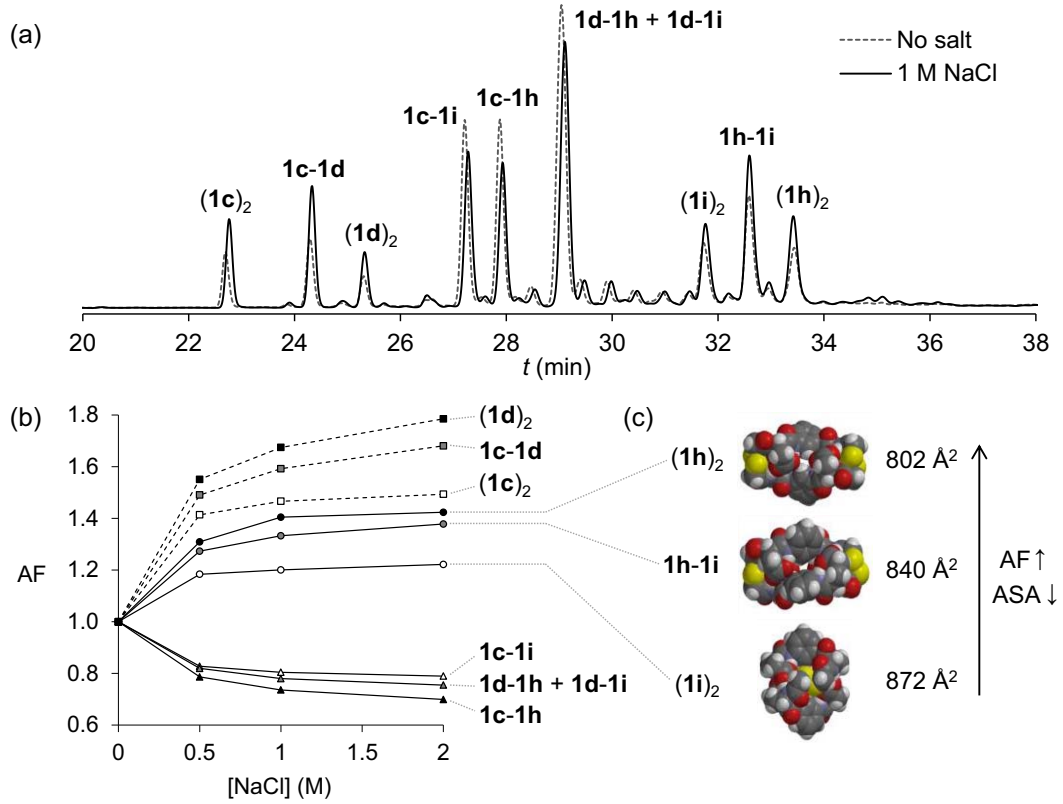


Figure R.5. (a) Perfils HPLC-UV (254 nm) de la DCL generada a partir de la mescla de **1c**+**1d**+**1h**+**1i** (2 mM de cadascun) en tampó aquós de fosfat (pH 7.5) amb un 25% (v/v) de DMSO en absència de sal (línia discontínua de color gris) i en presència de NaCl 1,0 M NaCl (línia contínua de color negre); (b) representació del factor d'amplificació dels deu dímers en funció de la concentració de NaCl; i (c) models CPK dels mínims globals calculats per (**1h**)$_2$, **1h-1i** i (**1i**)$_2$ amb els corresponents valors de ASA (Å^2).

L'efecte de l'increment de la concentració salina també es va investigar en una DCL més complexa que contenia, a més dels tres blocs constituents utilitzats anteriorment (**1c**, **1d** i **1i**), el ditiol derivat de l'àcid aspàrtic (**1h**). Es va veure que l'augment de la salinitat del medi provoca l'amplificació de tots aquells membres que agrupen els blocs constituents **1i** i/o **1h**, corroborant la selecció de les espècies que concentren cadenes laterals àcides (Figura R.5a). A més, es va veure que l'amplificació dels derivats de l'àcid aspàrtic és major que la dels derivats de l'àcid glutàmic, indicant que els primers presenten una millor adaptació a l'increment de la salinitat (Figura R.5b). Aquesta tendència també ha estat descrita per l'evolució de les proteïnes halòfiles i es creu que està relacionada amb la minimització de la superfície accessible al dissolvent (Figura R.5c).

Per últim, es va investigar l'efecte de diferents sals i es va veure que el clorur sòdic, el clorur potàssic i el nitrat sòdic provoquen canvis gairebé idèntics en la composició de la quimioteca. Per tant, l'efecte d'incrementar la salinitat del medi no respon a cap interacció específica del catió o l'anió, sinó a l'increment de la força iònica del medi.

Conclusions

S'ha utilitzat satisfactòriament una DCL minimalista de disseny bioinspirat per reproduir tendències adaptatives pròpies de processos evolutius biològics. Així, en una DCL de pseudopèptids macrocíclics, l'increment de la concentració de sal provoca la selecció i amplificació d'aquelles espècies que concentren aminoàcids àcids. A més, els derivats de l'àcid aspàrtic presenten una millor adaptació salina que els derivats de l'àcid glutàmic. Estudis estructurals indiquen que les espècies seleccionades presenten una conformació plegada en dissolució i revelen una millor adaptació salina per aquells membres que tenen una menor superfície accessible al dissolvent. El procés adaptatiu és causat per l'increment de la força iònica i té una notable similitud amb l'evolució natural de les proteïnes dels microorganismes halòfils per sobreviure en medis hipersalins.

CAPÍTOL 4

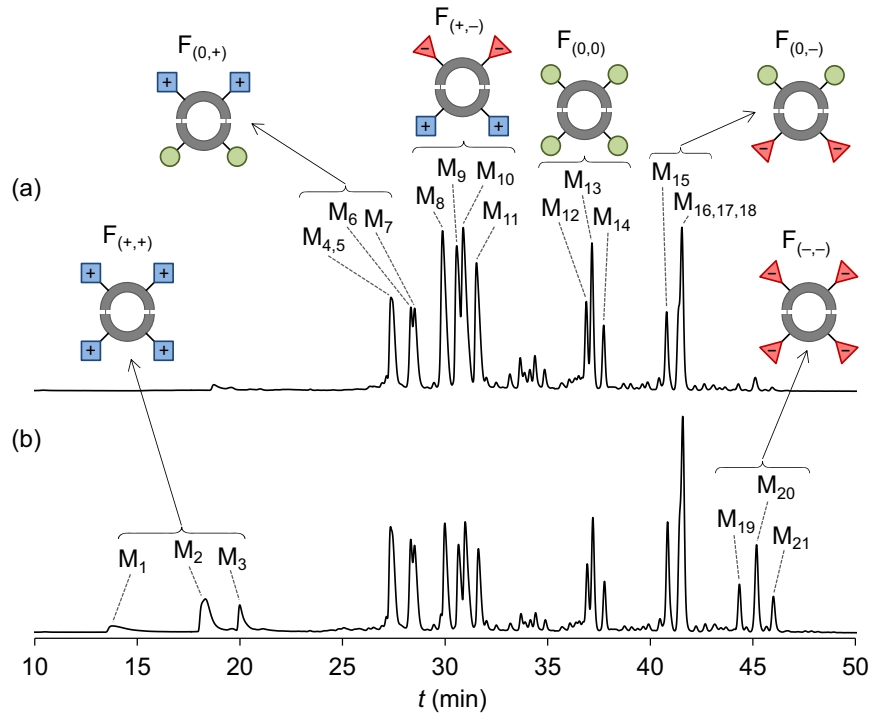
Introducció i objectius

Si bé l'adició d'una plantilla molecular és probablement la manera més habitual de pertorbar la composició d'una quimioteca dinàmica, en alguns casos altres factors químics i físics també s'han emprat amb aquesta finalitat. Un d'aquests factors és l'increment de la força iònica resultant de la dissolució d'una sal inorgànica en el medi, tal i com s'ha vist en el capítol anterior. Aquest estímul extern ha estat relativament poc estudiat i gairebé sempre s'ha utilitzat amb l'únic propòsit d'afavorir les interaccions hidrofòbiques entre els components de les DCLs. Així, per exemple, l'adició de nitrat sòdic ha demostrat ser una estratègia molt útil per afavorir la formació d'aquelles molècules que minimitzen la superfície de contacte amb el dissolvent, i s'ha utilitzat en diverses ocasions per preparar estructures amb elevada complexitat topològica.[47-48] Recentment, l'increment de la força iònica també s'ha emprat per induir la separació de fases líquid-líquid en una mescla binària de dissolvents, provocant la redistribució dels components d'una DCL.[15]

L'objectiu principal d'aquest capítol és estudiar l'efecte d'incrementar la força iònica en la composició d'una DCL complexa formada per espècies amb diferents càrregues. La nostra hipòtesi és que la caracterització del sistema mitjançant experiments de deconvolució dinàmica permetrà desxifrar les normes que governen el comportament global del sistema i identificar les relacions coadaptatives que s'estableixen entre els membres de la xarxa molecular.

Resultats i discussió

Es va dissenyar una DCL complexa generada a partir de la mescla de sis blocs constituents amb diferents càrregues a pH neutre: dos amb cadenes laterals no carregades (**1b** i **1d**), dos amb cadenes laterals carregades negativament (**1h** i **1i**) i dos amb cadenes laterals carregades positivament (**1j** i **1k**). La quimioteca es va preparar dissolent quantitats equimolars dels sis compostos en un tampó aquós de bis-Tris (pH 6,5) amb un 25% (v/v) de DMSO. Paral·lelament, la mateixa DCL també es va preparar afegint-hi diferents quantitats de NaCl (0,5-2,0 M). Un cop equilibrades, les DCLs es van analitzar mitjançant HPLC-UV (Figura R.6) i UPLC-MS.



$F_{(+,+)}$	$F_{(0,+)}$	$F_{(+,–)}$	$F_{(0,0)}$	$F_{(0,–)}$	$F_{(–,–)}$
M_1: (**1j**)$_2$	$M_{4,5}$: **1b-1j**	M_8: **1i-1j**	M_{12}: (**1b**)$_2$	M_{15}: **1b-1i**	M_{19}: (**1i**)$_2$
M_2: **1j-1k**	$M_{4,5}$: **1d-1j**	M_9: **1h-1j**	M_{13}: **1b-1d**	$M_{16\text{-}18}$: **1b-1h**	M_{20}: **1h-1i**
M_3: (**1k**)$_2$	M_6: **1b-1k**	M_{10}: **1i-1k**	M_{14}: (**1d**)$_2$	$M_{16\text{-}18}$: **1d-1h**	M_{21}: (**1h**)$_2$
	M_7: **1d-1k**	M_{11}: **1h-1k**		$M_{16\text{-}18}$: **1d-1i**	

Figura R.6. Perfils HPLC-UV (254 nm) de la DCL generada a partir de la mescla dels blocs constituents **1b**+**1d**+**1h**+**1i**+**1j**+**1k** (0,5 mM de cadascun) en tampó aquós de bis-Tris (pH 6,5) amb un 25% (v/v) DMSO en absència de sal (a) i en presència de NaCl 1,0 M (b).

Es va obtenir una DCL molt diversa formada per 21 dímers ($M_{1\text{-}21}$, Figura R.6) i es va veure que l'adaptació salina de cadascun dels membres de la quimioteca es pot classificar en sis famílies atenent a la càrrega dels corresponents blocs constituents. Així, tots els membres d'una mateixa família responen de manera similar a l'increment de la força iònica. L'anàlisi quantitativa de la composició de la DCL a diferents concentracions de sal va evidenciar que, en absència de l'estímul extern (Figura R.6a), el sistema optimitza les interaccions electrostàtiques afavorint els membres que agrupen càrregues de diferent signe ($F_{(+,–)}$) i desafavorint els membres que agrupen càrregues d'igual signe ($F_{(+,+)}$ i $F_{(–,–)}$). L'increment de la força iònica del medi fa que aquestes interaccions electrostàtiques siguin apantallades i que la composició de la quimioteca s'aproximi a la proporció estadística (Figura R.6b).

Després d'analitzar el comportament de les sis famílies de dímers i veure que les interaccions electrostàtiques són el factor que més influencia el comportament de la

DCL, diversos experiments de deconvolució dinàmica van ser dissenyats amb la finalitat d'estudiar la relació entre les propietats estructurals i el procés adaptatiu de cadascun dels membres. Així, l'efecte d'incrementar la força iònica va ser avaluat quantitativament en les 15 possibles subquimioteques de només dos blocs constituents (mescles binàries, *e.g.* Esquema R.3a). A més, les preferències conformacionals dels homodímers derivats de l'àcid glutàmic ((**1i**)$_2$) i de l'ornitina ((**1j**)$_2$) van ser estudiades mitjançant diversos experiments de ressonància magnètica nuclear i simulacions de dinàmica molecular. El conjunt de resultats obtinguts evidencia dues tendències adaptatives principals. En primer lloc, per les espècies formades a partir de la unió de blocs constituents amb càrregues del mateix signe, els membres amb cadenes laterals més curtes s'adapten millor a l'increment de la força iònica (*e.g.* (**1h**)$_2$ presenta una millor adaptació salina que (**1i**)$_2$). En segon lloc, els homodímers amb càrrega negativa s'adapten millor a l'increment de la força iònica que els homodímers amb càrrega positiva (*e.g.* (**1i**)$_2$ presenta una millor adaptació salina que (**1j**)$_2$). Aquestes dues tendències estan directament relacionades amb la distància entre les diferents càrregues de cada membre de la quimioteca.

Per últim, el coneixement derivat de l'estudi del sistema complex va permetre dissenyar sistemes més simples per tal d'investigar les relacions de competència i cooperació que s'estableixen entre els membres de la quimioteca. L'estudi d'aquests sistemes va permetre corroborar les tendències adaptatives observades en el sistema complex.

Conclusions

S'ha estudiat l'efecte d'incrementar la força iònica en la composició d'una DCL complexa formada per 21 dímers macrocíclics amb diferents càrregues. El procés adaptatiu s'ha caracteritzat mitjançant experiments de deconvolució dinàmica i l'estudi estructural d'alguns membres seleccionats de la quimioteca. S'ha vist que, en absència de sal, el sistema optimitza les interaccions electrostàtiques i presenta una composició clarament diferent a la distribució estadística. L'increment de la força iònica apantalla aquestes interaccions, fent que la composició de la quimioteca s'aproximi a la proporció estadística. La resposta dels membres de la quimioteca a l'increment de la salinitat del medi es pot classificar en diferents famílies segons les càrregues del membres. S'ha vist que el comportament de cada membre de la quimioteca està determinat per la seva

informació estructural i per les múltiples relacions coadaptatives que aquest estableix amb la resta de membres que la xarxa molecular.

CAPÍTOL 5

Introducció i objectius

Moltes biomolècules, com per exemple l'ARN, l'ADN i els polisacàrids, estan formades per monòmers quirals que presenten una sola configuració: aminoàcids L i sucres D. Aquesta homogeneïtat de les biomolècules, anomenada homoquiralitat biològica, és una propietat que presenten tots els organismes vius i té un paper fonamental en la majoria de processos biomoleculars. L'origen de l'homoquiralitat biològica és un problema no resolt sobre el qual s'han formulat múltiples teories i hipòtesis.[49-51] Alguns d'aquests plantejaments teòrics han estat recolzats per diversos estudis experimentals.[52-55]

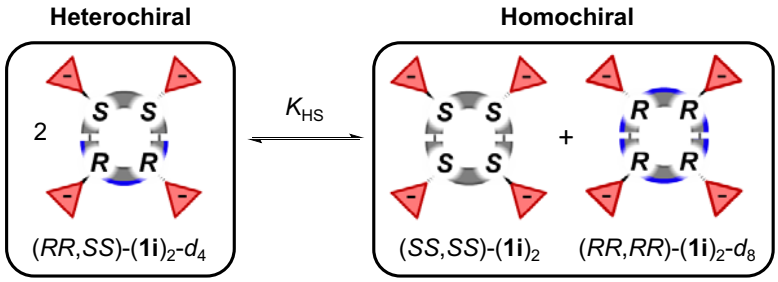
Un fenomen que es creu que podria estar estretament lligat a l'origen de l'homoquiralitat biològica és l'autoordenació quiral ("chiral self-sorting" en anglès). Aquest fenomen és el resultat d'un procés de reconeixement quiral entre els components d'una mescla estereoisomèrica, i pot donar lloc a processos d'autoreconeixement o d'autodiscriminació,[56] depenent de si el reconeixement es dóna preferentment entre molècules amb una mateixa configuració (generant espècies homoquirals) o entre diferents enantiòmers (generant espècies heteroquirals).

L'objectiu principal d'aquest Capítol és investigar el procés de transmissió de la informació quiral en una DCL minimalista formada per compostos flexibles i de naturalesa pseudopeptídica. La nostra hipòtesi és que les diferències estructurals entre membres diastereoisomèrics d'una quimioteca dinàmica podran donar lloc a l'aparició de processos d'autoordenació quiral.

Resultats i discussió

Es va dissenyar una DCL minimalista generada a partir de la mescla de dos blocs constituents pseudoenantiomèrics: (*SS*)-**1i** i (*RR*)-**1i**-d_4. La quimioteca es va preparar dissolent quantitats equimolars dels dos compostos en un tampó aquós de fosfat (pH 8,0) amb un 1,2% (v/v) de DMSO. Un cop equilibrada, la DCL es va analitzar mitjançant HPLC-UV quiral i UPLC-MS, i es va veure que la xarxa molecular generada consta dels tres membres representats en l'Esquema R.4: un dímer heteroquiral i dos dímers homoquirals. L'estratègia de marcar isotòpicament un dels dos blocs

constituents ((*RR*)-**1i**-d_4) va demostrar ser de gran utilitat, permetent la directa identificació dels tres dímers de la quimioteca mitjançant UPLC-MS.



Esquema R.4. Reacció d'intercanvi entre dímers homo- i heteroquirals.

$$K_{HS} = \frac{\left[(SS,SS)\text{-}(\mathbf{1i})_2\right]\cdot\left[(RR,RR)\text{-}(\mathbf{1i})_2\text{-}d_8\right]}{\left[(RR,SS)\text{-}(\mathbf{1i})_2\text{-}d_4\right]^2} \qquad \textbf{(Equació R.1)}$$

Es va avaluar quantitativament la composició de la DCL mitjançant la constant d'intercanvi K_{HS} (Equació R.1) i es va veure que l'equilibri representat en l'esquema R.4 està lleugerament desplaçat cap a la dreta, indicant un cert grau d'autoreconeixement.

A continuació es va investigar l'efecte de la polaritat del medi en el grau autoordenació homoquiral. Amb aquesta finalitat, la mateixa DCL es va preparar en un medi més polar, incorporant clorur de guanidini (Gnd·HCl), i en diversos medis més apolars, substituint el tampó de fosfat per aigua basificada amb hidròxid de tetrabutilamoni i incorporant diferents quantitats d'acetonitril (AN). Es va calcular la constant d'intercanvi K_{HS} de cadascuna de les quimioteques (Figura R.7d) i es va veure que la disminució de la polaritat del medi clarament afavoreix el procés de selecció homoquiral. Així, per exemple, en presència d'un 92% (v/v) d'AN la constant d'intercanvi és deu vegades més gran que en la situació teòrica d'absència d'autoreconeixement. Aquests resultats suggereixen que la selecció homoquiral observada és deguda a l'efecte d'interaccions de tipus polar.

Complementàriament, algunes DCLs seleccionades es van preparar a diferents temperatures (2, 22 i 45 °C), diferents valors de pH (2,5-7,5) i diferents concentracions dels blocs constituents (0,5-2,0 mM). Es va veure que el grau d'autoreconeixement augmenta amb la temperatura. Aquest resultat implica una contribució entròpica positiva al procés de selecció homoquiral. Pel que fa al pH del medi, el valor de la constant d'intercanvi K_{HS} augmenta amb el pH, indicant que les espècies completament

desprotonades són les que presenten un major grau d'autoreconeixement. Finalment, es va observar que les DCLs preparades a diferent concentració dels blocs constituents presenten una distribució pràcticament idèntica, indicant que en el procés d'autoreconeixement no hi intervenen interaccions intermoleculars.

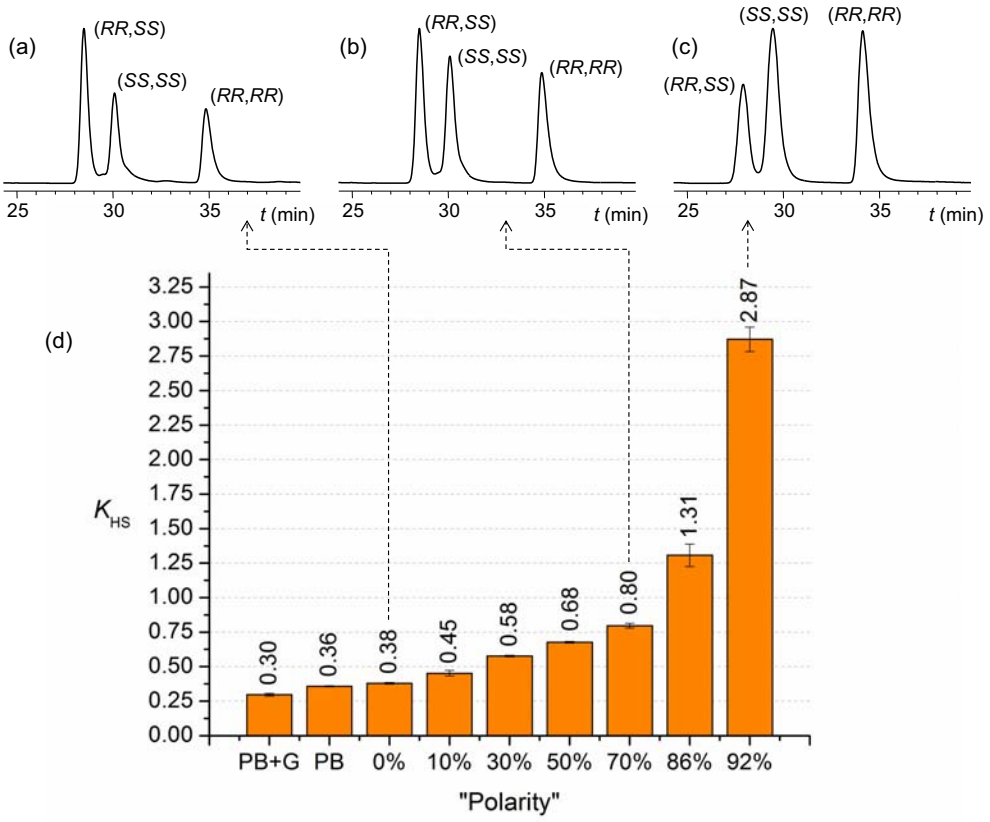


Figura R.7. (a) Perfils HPLC-UV quirals (254 nm) de la mescla binària de (*SS*)-**1i** + (*RR*)-**1i**-d_4 (0,5 mM de cadascun) preparada a 22 °C en (a) aigua basificada + 1,2% (v/v) DMSO; (b) aigua basificada + 70% (v/v) AN + 1,2% (v/v) DMSO; i (c) aigua basificada + 92% (v/v) AN + 1,2% (v/v) DMSO. (d) Representació del valor de la constant K_{HS} en funció d'una escala qualitativa de "polaritat". (PB+G = tampó aquós de fosfat (pH 8.0) amb Gnd·HCl 1,0 M; PB = tampó aquós de fosfat (pH 8.0); els valors 0, 10, 30, 50, 70, 86 i 92% corresponen a les mescles binàries preparades amb aquest percentatge d'AN).

Per últim, les propietats conformacionals dels membres de la quimioteca van ser avaluades mitjançant experiments de ressonància magnètica nuclear. Els resultats preliminars obtinguts suggereixen que els dímers homo- i heteroquirals presenten conformacions significativament diferents en dissolució. En un futur, estudis estructurals complementaris, incloent modelatge molecular, serviran per acabar d'entendre el procés d'autoordenació homoquiral a nivell molecular.

Conclusions

S'ha utilitzat satisfactòriament una DCL minimalista formada per pseudopèptids dimèrics homo- i heteroquirals per tal d'estudiar el fenomen d'autoordenació quiral. En aigua tamponada, la composició de la quimioteca presenta una lleugera preferència pels dímers homoquirals, indicant un cert grau d'autoreconeixement. L'autoordenació homoquiral es veu clarament accentuada quan la polaritat del medi disminueix per efecte de la presència d'un codissolvent orgànic, indicant que el procés de selecció és degut a interaccions intramoleculars de tipus polar. A més, el grau de selectivitat homoquiral també incrementa amb la temperatura, indicant una contribució entròpica positiva al procés de selecció. Experiments preliminars de ressonància magnètica nuclear suggereixen que els dímers homo- i heteroquirals presenten conformacions significativament diferents en dissolució. S'ha previst realitzar un estudi estructural més complert per tal d'entendre millor el procés d'autoordenació homoquiral a nivell molecular.

BIBLIOGRAFIA

(1) Kindermann, M.; Stahl, I.; Reimold, M.; Pankau, W. M.; von Kiedrowski, G. *Angew. Chem. Int. Ed.* **2005**, *44*, 6750.
(2) Li, J.; Nowak, P.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 9222.
(3) Hunt, R. A. R.; Otto, S. *Chem. Commun.* **2011**, *47*, 847.
(4) Ludlow, R. F.; Otto, S. *Chem. Soc. Rev.* **2008**, *37*, 101.
(5) Steed, J. W.; Atwood, J. L. *Supramolecular chemistry*; John Wiley & Sons, 2013.
(6) Beer, P. D.; Gale, P. A.; Smith, D. K. *Supramolecular chemistry*; Oxford University Press, 1999.
(7) Otto, S.; Severin, K. In *Creative Chemical Sensor Systems*; Schrader, T., Ed.; Springer Berlin Heidelberg: 2007; Vol. 277, p 267.
(8) Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H. *Angew. Chem. Int. Ed.* **2014**, *53*, 3259.
(9) Herrmann, A. *Chem. Soc. Rev.* **2014**, *43*, 1899.
(10) López-Senín, P.; Gómez-Pinto, I.; Grandas, A.; Marchán, V. *Chem. Eur. J.* **2011**, *17*, 1946.
(11) Faggi, E.; Moure, A.; Bolte, M.; Vicent, C.; Luis, S. V.; Alfonso, I. *J. Org. Chem.* **2014**, *79*, 4590.
(12) Riddell, I. A.; Smulders, M. M. J.; Clegg, J. K.; Hristova, Y. R.; Breiner, B.; Thoburn, J. D.; Nitschke, J. R. *Nat. Chem.* **2012**, *4*, 751.
(13) You, L.; Berman, J. S.; Anslyn, E. V. *Nat. Chem.* **2011**, *3*, 943.
(14) Buryak, A.; Severin, K. *Angew. Chem. Int. Ed.* **2005**, *44*, 7935.
(15) Hafezi, N.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 12861.
(16) Perez-Fernandez, R.; Pittelkow, M.; Belenguer, A. M.; Lane, L. A.; Robinson, C. V.; Sanders, J. K. M. *Chem. Commun.* **2009**, 3708.
(17) Fanlo-Virgós, H.; Alba, A.-N. R.; Hamieh, S.; Colomb-Delsuc, M.; Otto, S. *Angew. Chem. Int. Ed.* **2014**, *53*, 11346.
(18) Dydio, P.; Breuil, P.-A. R.; Reek, J. N. H. *Isr. J. Chem.* **2013**, *53*, 61.
(19) Tauk, L.; Schröder, A. P.; Decher, G.; Giuseppone, N. *Nat. Chem.* **2009**, *1*, 649.
(20) Chang, T.; Rozkiewicz, D. I.; Ravoo, B. J.; Meijer, E. W.; Reinhoudt, D. N. *Nano Lett.* **2007**, *7*, 978.
(21) Kovaříček, P.; Lehn, J.-M. *J. Am. Chem. Soc.* **2012**, *134*, 9446.
(22) Campaña, A. G.; Carlone, A.; Chen, K.; Dryden, D. T. F.; Leigh, D. A.; Lewandowska, U.; Mullen, K. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 5480.
(23) von Delius, M.; Geertsema, E. M.; Leigh, D. A. *Nat. Chem.* **2010**, *2*, 96.
(24) Campbell, V. E.; de Hatten, X.; Delsuc, N.; Kauffmann, B.; Huc, I.; Nitschke, J. R. *Nat. Chem.* **2010**, *2*, 684.
(25) Kuriyan, J.; Konforti, B.; Wemmer, D. *The molecules of life: Physical and chemical principles*; Garland Science, 2012.
(26) Whitford, D. *Proteins: structure and function*; John Wiley & Sons, 2013.
(27) Luis, S. V.; Alfonso, I. *Acc. Chem. Res.* **2014**, *47*, 112.
(28) Paradowska, J.; Pasternak, M.; Gut, B.; Gryzło, B.; Mlynarski, J. *J. Org. Chem.* **2012**, *77*, 173.
(29) Bauke Albada, H.; Rosati, F.; Coquière, D.; Roelfes, G.; Liskamp, R. M. J. *Eur. J. Org. Chem.* **2011**, *2011*, 1714.
(30) Rodríguez-Llansola, F.; Escuder, B.; Miravet, J. F. *J. Am. Chem. Soc.* **2009**, *131*, 11478.
(31) Micale, N.; Scarbaci, K.; Troiano, V.; Ettari, R.; Grasso, S.; Zappalà, M. *Med. Res. Rev.* **2014**, *34*, 1001.

(32) Zervoudi, E.; Saridakis, E.; Birtley, J. R.; Seregin, S. S.; Reeves, E.; Kokkala, P.; Aldhamen, Y. A.; Amalfitano, A.; Mavridis, I. M.; James, E. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 19890.
(33) Kokkonen, P.; Rahnasto-Rilla, M.; Kiviranta, P. H.; Huhtiniemi, T.; Laitinen, T.; Poso, A.; Jarho, E.; Lahtela-Kakkonen, M. *ACS Med. Chem. Lett.* **2012**, *3*, 969.
(34) Moure, A.; Luis, S. V.; Alfonso, I. *Chem. Eur. J.* **2012**, *18*, 5496.
(35) Alfonso, I.; Burguete, I.; Luis, S. V.; Miravet, J. F.; Seliger, P.; Tomal, E. *Org. Biomol. Chem.* **2006**, *4*, 853.
(36) Valeur, E.; Bradley, M. *Chem. Soc. Rev.* **2009**, *38*, 606.
(37) Sevier, C. S.; Kaiser, C. A. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 836.
(38) Zhou, N. E.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1993**, *32*, 3178.
(39) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70.
(40) Lanyi, J. K. *Bacteriol. Rev.* **1974**, *38*, 272.
(41) Pieper, U.; Kapadia, G.; Mevarech, M.; Herzberg, O. *Structure* **1998**, *6*, 75.
(42) Mevarech, M.; Frolow, F.; Gloss, L. M. *Biophys. Chem.* **2000**, *86*, 155.
(43) Madern, D.; Ebel, C.; Zaccai, G. *Extremophiles* **2000**, *4*, 91.
(44) Paul, S.; Bag, S.; Das, S.; Harvill, E.; Dutta, C. *Genome Biol.* **2008**, *9*, 1.
(45) Fukuchi, S.; Yoshimune, K.; Wakayama, M.; Moriguchi, M.; Nishikawa, K. *J. Mol. Biol.* **2003**, *327*, 347.
(46) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Angew. Chem. Int. Ed.* **2007**, *46*, 8858.
(47) Cougnon, F. B. L.; Ponnuswamy, N.; Jenkins, N. A.; Pantoş, G. D.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2012**, *134*, 19129.
(48) Cougnon, F. B. L.; Jenkins, N. A.; Pantoş, G. D.; Sanders, J. K. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 1443.
(49) Brewer, A.; Davis, A. P. *Nat. Chem.* **2014**, *6*, 569.
(50) Calvin, M. *Chemical evolution: molecular evolution towards the origin of living systems on the earth and elsewhere*; Oxford University Press, 1969.
(51) Frank, F. C. *Biochim. Biophys. Acta* **1953**, *11*, 459.
(52) Steendam, R. R. E.; Verkade, J. M. M.; van Benthem, T. J. B.; Meekes, H.; van Enckevort, W. J. P.; Raap, J.; Rutjes, F. P. J. T.; Vlieg, E. *Nat. Commun.* **2014**, *5*.
(53) Dressel, C.; Reppe, T.; Prehm, M.; Brautzsch, M.; Tschierske, C. *Nat. Chem.* **2014**, *6*, 971.
(54) Klussmann, M.; Iwamura, H.; Mathew, S. P.; Wells, D. H.; Pandya, U.; Armstrong, A.; Blackmond, D. G. *Nature* **2006**, *441*, 621.
(55) Soai, K.; Shibata, T.; Morioka, H.; Choji, K. *Nature* **1995**, *378*, 767.
(56) Safont-Sempere, M. M.; Fernández, G.; Würthner, F. *Chem. Rev.* **2011**, *111*, 5784.