UNIVERSITAT DE BARCELONA

Study and applications of dynamic chemical networks of pseudopeptidic compounds

Maria Lafuente Fabra

DEPARTAMENT DE QUÍMICA ORGÀNICA

FACULTAT DE QUÍMICA
UNIVERSITAT DE BARCELONA

Programa de doctorat de Química Orgànica

Study and applications of dynamic chemical networks of pseudopeptidic compounds

Memòria presentada per **Maria Lafuente Fabra**
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)

Directors:
Dr. Ignacio Alfonso Rodríguez
Dr. Jordi Solà Oller
Departament de Química Biològica i Modelització Molecular (IQAC-CSIC)

Tutor:
Dr. Pedro Romea García

Departament de Química Orgànica
Facultat de Química (UB)

Doctorand:

Maria Lafuente Fabra
Departament de Química Biològica i Modelització Molecular (IQAC-CSIC)

Als meus estimats pares.

ACKNOWLEDGMENTS

En primer lloc, vull agrair especialment als meus directors de tesis, el Dr. Ignacio Alfonso i el Dr. Jordi Solà, pel seu suport acadèmic i per haver-me donat l'oportunitat de realitzar aquest projecte en el seu grup. Gràcies per haver-me ajudat i guiat durant tota aquesta etapa professional i per haver-me donat suficient llibertat per equivocar-me i aprendre, però també perquè sempre m'he sentit recolzada i acompanyada durant tot aquest camí. Sempre heu tingut una paraula encoratjadora i positiva, i m'heu contagiat aquesta motivació durant tota la tesis. M'heu transmès molts coneixements però també una manera de treballar i d'enfocar els projectes que em serà útil durant tota la meva vida professional.

Gràcies a tots els meus companys del Lab.309, els que sempre hi heu sigut i als que ja han marxat, que heu compartit amb mi el dia a dia durant aquests quatre anys. A tots, gràcies pels esmorzars, les estones de riures compartits, les anècdotes i també pels sopars de laboratori, festes i congressos. Perquè sense vosaltres no hagués estat el mateix. Acaba aquesta etapa i me'nduc una pila de records que sempre portaré amb mi. Quan vaig arribar em vàreu rebre amb els braços oberts i em vàreu ajudar amb tot, tant personal com professionalment. En especial als meus companys, que ja son amics, de Mepram: Anna, Asun, Cristian i Lucia. Gràcies per tot aquest temps. També vull enrecordar-me del Joan, el meu sensei durant la primera etapa de la tesi. Gràcies a tots, de veritat!

També voldria expressar el meu agraïment a tothom de l'IQAC que m'ha ajudat en algun moment en aquest projecte. En particular m'agradaria donar les gràcies al Professor Angel Messeguer i al Dr. Ciril Jimeno pels consells rebuts i el seu interès envers la meva feina. A la Dra. Yolanda Pérez per l'ajut amb els experiments de RMN. A la Dra. Maria José Gómara per per donar-me facilitats per utilitzar les seves instal·lacions.

I would like to thank Prof. David Amabilino (University of Nothingham) for giving me the opportunity to work in his research group. It was an amazing experience for me to work in another leading research group and get the chance to learn a different way to work and used different techniques. I would also like to thank all the members of the Amabilino group for making my short stay in Nottingham unforgettable, for your summer visits to Barcelona and for inviting me to come back so many times.

A Lula, porque después de muchos años de amistad, de compartir colegio e infinitas cosas más, me animaste a tomar la decisión de realizar el doctorado. Me acordaré siempre de aquel momento. Tú lo sabes bien. Te has seguido interesando siempre por saber de qué iba todo esto y, aunque siempre me ha dado pereza explicártelo, te has salido con la tuya. Gracias Luls.

Als Chemical Brothers. Amb vosaltres tot es més fàcil. Vàrem començar l'aventura de la Química junts i molts també vàreu triar el camí del doctorat. Han passat ja més de deu anys i hem compartit infinits moments, des d'hores d'estudi fins a bussejos a les platges del sud-est asiàtic. Crec que pocs grups poden presumir de tantes aventures. El temps passa i seguim estant igual d'units i amb la mateixes ganes de celebrar, i tot i que molts avui en dia viviu a l'estranger, seguim traient temps per reunir-nos. Ho considero una gran sort. A l'Ana i la Mire, que sempre hi son, per a tot.

A l'Adri que ha compartit amb mi tota aquesta etapa i no ha deixat d'animar-me ni un sol moment. Gràcies per intentar que sempre estigui rient, saps que se't dona molt bé. Sempre pendent del que fem al laboratori, llegint-te els articles i fent preguntes de tot, explicant-li als amics el que feia i aportant idees esbojarrades però amb molta il·lusió, crec que ningú hauria pogut mostra més interès en tot el que faig com ho fas tu. Gràcies per alegrar-te dels meus èxits i donar-me suport sempre, per ajudar-me tant durant l'escriptura però sobretot gràcies per estar sempre al meu costat.

I per últim, als meus estimats pares, a qui dedico aquesta tesis. Gràcies per tota l'estima i suport, pels valors transmesos i l'educació rebuda, i per haver-me ajudat i animat a formar-me sempre. Sense ells res de tot això hagués estat possible. Gràcies per fer-me creure que sóc capaç d'afrontar qualsevol repte. Per ajudar-me en tot sense demanar res. Gràcies per contribuir sempre a la meva felicitat.

The present doctoral thesis was carried out at the Institute of Advanced Chemistry of Catalonia (IQAC), which belongs to the Spanish National Research Council (CSIC).

This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO, CTQ2012-38543-C03-03 and CTQ2015-70117-R projects), la Generalitat de Catalunya (AGAUR, 2014SGR231) and the European Union (COST action CM1304 and FP7-PEOPLE-2012-CIG-321659). I am also grateful for personal financial support with a predoctoral fellowship (MINECO, FPI: BES-2013-063128).

ABBREVIATIONS

AHBO	Azido-substituted hydroxyphenylbenzoxazole derivate
BB	Building block
bis-Tris	Bis(2-hydroxyethyl)aminotris(hydroxymethyl)methane
Boc	*Tert*-butyloxycarbonyl
CD	Circular dichroism
Cys	L-Cysteine
CyA	Cysteamine
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
DTNB	5,5'-dithiobis(2-nitrobenzoic acid)
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	1,2-ethanedithiol
EtOAc	Ethyl acetate
EtOH	Ethanol
ESI	Electrospray ionization
FIA	Flow injection analysis
Fmoc	Fluorenylmethyloxycarbonyl
GSH	Glutathione
HOBt	1-hydroxybenzotriazole
HSQC	Heteronuclear single quantum correlation
HMBC	Heteronuclear multiple bond correlation
Hcy	Homocysteine
HRMS	High resolution mass spectrometry
LC	Liquid chromatography
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization Time-of-flight
MeOH	Methanol
MS	Mass spectrometry
NAC	Acetylcysteine
NMR	Nuclear magnetic resonance
RP-HPLC	Reversed phase high performance liquid chromatography
TFA	Trifluoroacetic acid
TIS	Triisobutylsilane
TLC	Thin layer chromatography
TMS	Trimethylsilane
TOF	Time-of-flight
Trt	Triphenylmethyl
UPLC	Ultra performance liquid chromatography

UPLC-TOF MS	Ultra performance liquid chromatography Time-of-flight mass spectrometry
UV-Vis	Ultraviolet-visible

TABLE OF CONTENTS

Barcelona, June the 6th, 2018

To whom it may concern:

This section is included following mandatory guidelines and serves to certify the contribution of María Lafuente Fabra to each publication included in this PhD book.

Publications:

Solà, J.; Lafuente, M.; Atcher, J.; Alfonso, I. "Constitutional self-selection from dynamic combinatorial libraries in aqueous solution through supramolecular interactions" *Chem. Commun.* **2014**, *50*, 4564- 4566, DOI: 10.1039/c4cc00245h, IF: 6.834, Q1.

María Lafuente performed most of the experimental work and also the interpretation of the data. The initial experiments were carried out by J. Solà, and then the project was taken by María as the main goal for her PhD Thesis.

Lafuente, M.; Atcher, J.; Solà, J.; Alfonso, I. "Adaptive correction from virtually complex dynamic libraries: the role of non-covalent interactions in structural selection and folding" *Chem. Eur. J*. 2015, 21, 17002-17009, DOI: 10.1002/chem.201501415, IF: 5.771, Q1.

In this publication María carried out all the experimental part and most of the planning and interpretation of the results.

Lafuente, M; Solà, J.; Alfonso, I. "A dynamic chemical network for cystinuria diagnosis" *Angew. Chem. Int. Ed.* **2018**, accepted in press, 10.1002/anie.201802189, IF: 11.994, Q1.

All the experimental work in this communication was carried out by María Lafuente, including the synthesis of the building blocks, the design and optimization of the sensing systems, as well as the measurements with the biofliud.

Lafuente, M; Solà, J.; Alfonso, I. "Structurally selective assembly of a specific macrobicycle from a dynamic library of pseudopeptidic disulfides", manuscript to be submitted.

All the experimental work included in this full paper was performed by María Lafuente. I must highlight that in this last paper, María showed a high degree of maturity and

independent thinking. This manuscript will be submitted for publication soon, probably just after the defense.

We also certify that the data included in this PhD was not used in any other document for the achievement of any academic degree (PhD, Master degree) and will not be used by a future student for any similar objective. Thus, the results here described have been obtained exclusively for the PhD of María Lafuente.

Sincerely,

PhD Advisor	PhD Advisor
Ignacio Alfonso Rodríguez	Jordi Solà i Oller

CHAPTER I

General introduction

CHAPTER I: General introduction

1.1. Supramolecular chemistry

Supramolecular chemistry[1-2] has been defined[3-4] as the 'chemistry beyond the molecule' referring to the chemistry of molecular assemblies and intermolecular bonds, covering the structures and functions of the entities formed by association of two or more chemical species. In contrast to molecular chemistry that focuses on the covalent bond, supramolecular chemistry is based on the study of weaker and reversible non-covalent interactions within and between molecules, and on the properties of the resulting multimolecular complexes. These non-covalent interactions include electrostatic effects, hydrogen bonding, metal coordination, van der Waals forces, π-π stacking, solvophobic effects and halogen bonding.

The first studies of architectures monitored by specific noncovalent interactions and the early indications of a future supramolecular chemistry were established in 1894 when Fischer suggested that enzyme-substrate interactions take the form of a "lock and key", the fundamental principles of molecular recognition and host-guest chemistry. The use of these principles led to an increasing understanding of protein structure and other biological processes. Consequently, chemists began to recognize and study synthetic structures based on noncovalent interactions and the interaction models between the receptor and the substrate were evolving over time. However, the importance of supramolecular chemistry was not established until in 1987, when Donald J. Cram, Jean-Marie Lehn, and Charles J. Pedersen were jointly awarded with the Nobel Prize for Chemistry in recognition of their work on "host-guest" assemblies, in particular for the development of selective "host-guest" complexes, in which a host molecule recognizes and selectively binds a certain guest. The rapid expansion of supramolecular chemistry over the past years has resulted in an enormous diversity of chemical systems, crossing from biological chemistry to materials science; and from synthesis to catalysis. Synthetic supramolecular architectures have found their inspiration in biological molecules, and can be used to manipulate and study the biological systems they were inspired from. Gaining control over these supramolecular interactions is crucial to understanding and targeting many biological processes.[5] Work in modern supramolecular chemistry encompasses not just host-guest systems but also synthesis and operation of molecular devices and machines, molecular recognition, molecular self-assembly, dynamic covalent chemistry and systems capable of spontaneously generating well-defined functional supramolecular architectures inspired by natural systems from their components. It is highly interdisciplinary in nature and attracts not just chemists but biologists, environmental scientists, physicists, biochemists, theoreticians, crystallographers.

At a fundamental level, supramolecular chemistry is an intrinsically dynamic process due to the lability of the noncovalent interactions connecting the molecular components. The interactions that connect the molecules of a supramolecular entity are reversibly dissociated and associated, allowing supramolecular species to readily exchange their component parts. The same can occur for "molecular chemistry", when a molecule contains covalent bonds that can form and break reversibly, its building blocks (BBs) can be continuously exchanged and reorganized. These are the features that define constitutional dynamic chemistry on both the molecular and supramolecular levels that paves the way towards an adaptive and evolutionary chemistry, a further step towards generating complexity within the field of systems chemistry.

1.2. Systems chemistry

The complexity around us in nature is overwhelming. The traditional research methods do not seem to be powerful enough to capture this diversity and it seems that a paradigm shift is needed, because there is a growing interest in complex networks in chemistry, biology, engineering and physics, among others. The field of systems chemistry deals with complex chemical systems that show properties that go beyond the sum of its components. Such properties are commonly found in systems that have many components that interact with each other and consist 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]

Research into complex networks[11-12] is well established in most major scientific disciplines including engineering, economics, biology, mathematics and physics, but not so far in chemistry. Commonly, chemists have focused for a long time on the synthesis of individual molecules with particular sought-after properties which may then be isolated for further study and exploitation, while complex dynamic mixtures have been intentionally avoided. However, with the recent rapid development of analytical tools, this situation has changed and the study of complex mixtures has already resulted in some useful applications.[13-14] Recently, systems chemistry has irrupted as an emerging discipline that looks at complex mixtures of interacting molecules and attempts to capture this complexity, which is prevalent in the life sciences. Complex mixtures can give rise to interesting and desirable emergent properties, properties that result from the interactions between components and cannot be attributed to any of these components acting in isolation such as non-linearity, emergence, spontaneous order, adaptation, and feedbacks loops, among others.[15-16] Complementary to this, a complex mixture contains information about all its constituents and the study of the complete mixture should in principle allow us to obtain properties of interest of all these molecules simultaneously, provided we can find a way of deconvoluting the results of such investigations.[17]

While multicomponent reactions have been studied for centuries, the idea of deliberately analyzing mixtures and reaction networks is more recent. The first mentions of systems chemistry as a field date from 2005, which it was first used by von Kiedrowski in 2005 in a publication describing the kinetic and computational analysis of a nearly exponential organic replicator.[18] In this paper, von Kiedrowski claims new research "*could open a door to a field that may be termed systems chemistry, namely, the design of prespecified dynamic behavior*." In the same year, a systems chemistry workshop was held in Venice, where a more detailed definition was given describing 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*".[19]

"Complexity" is a concept highly dependent on context and chemistry has its own understandings of it.[20] 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).[21]

Even though systems chemistry is a relatively young sub-discipline of chemistry, in a certain way it is not a totally new field, under this term are included a few research areas that already have been well established. In this regard, dynamic combinatorial chemistry has been central to developing complex systems.[22-24] Other areas of development include self-replication,[25-30] prebiotic chemistry,[31] molecular logics[32-34] and oscillating reactions,[35-36] among others.

1.3. Dynamic Combinatorial Chemistry

The Dynamic Combinatorial Chemistry (DCC)[37-39] has emerged as a powerful tool to create and study chemical complexity as it allows easy access to molecular networks. DCC proposes the creation of a mixture of compounds (dynamic combinatorial library) inter-connected through reversible chemical processes with individual properties that may be explored through the library's response to the stabilizing influences of external stimuli. A dynamic combinatorial library (DCL) is generated by combining BBs, functionalized such that they can react with each other through reversible covalent reactions or specific noncovalent interactions, to form a mixture of interconverting library members. The distribution of all molecules in such a network is typically, but not necessarily,[a,22] governed by thermodynamics. The changes of concentration of species in a DCL contain valuable information about changes in stability. The library composition is not fixed and the introduction of any external stimulus that can alter the relative stability of the library member will influence the product distribution, this means, the possibility of the chemical system to adapt to environmental changes.[40] Any change in the equilibrium composition will ideally lead to an increased production (an overexpression or amplification) of the stabilized library member at the expense of the other species in the mixture.

The DCC has proved valuable in identifying unexpected molecules with remarkable binding and molecular recognition properties, and in providing effective synthetic routes to complex species.[23] DCC has emerged as a new approach to the self-organisation of molecular libraries, more sophisticated and efficient than "conventional" synthetic chemistry, due to the fact that the library synthesis and the affinity screening step have been combined in a single process. This new methodology includes not only combinatorial elements but also selection, proofreading, self-correction and amplification elements. The DCC relies on a reversible connection process for the spontaneous and continuous generation of all possible combinations of a set of basic components, thus making virtually available all structural and interactional features that these combinations may present resulting in a dynamic mixture of products. The template (guest) molecule would be allowed to select by itself the most suitable host from a mixture of a virtual combinatorial library, this means, from all the possible combinations of the available components (hosts).[41-42] The stabilization of a particular library member through noncovalent interactions with an added template will ideally lead to an

[a] Although DCC was initially defined as combinatorial chemistry under thermodynamic control, actually have been explored systems that are not at equilibrium. Examples include self-replication and molecular machines.

increased production (or amplification) of the stabilized library member[43] minimizing the energy of the system. In this way, a library may be probed for species with affinity for a given target molecule.[44]

The basic principles of DCC were developed in the 1990s[45] by the groups of Sanders and Lehn. In the Sanders group, early experiments resulted into a study that demonstrated modest amplifications of specific macrocycles in the presence of alkali metal ions.[46] Lehn developed the dynamic combinatorial approach as a result of his work on metal helicates, observing that the major product in a dynamic mixture of helicates was determined by the nature of the counterion that binds in the center of the helicates.[47]

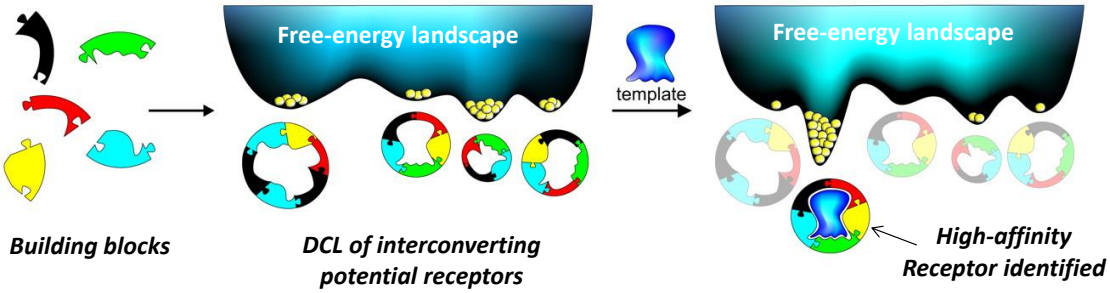


Figure 1.1. Schematic representation of a DCL and its free energy landscape, showing how the introduction of a template can lead to the amplification of the receptor with the highest affinity for the guest (Figure modified from reference [48]).

Although this concept was only recently articulated as a general approach to the synthesis of chemical species exhibiting molecular recognition capabilities, many of the principles and practices that characterize DCC had been in place for several decades. DCC may be viewed as the intersection of two pre-existing approaches to synthesis: thermodynamically controlled templated synthesis and combinatorial chemistry. Moreover, this approach will in principle minimize the synthetic efforts compared with conventional synthetic chemistry due to small number of BBs can lead to a wide range of large, complex products.

Therefore, the identification and synthesis of receptors for small molecules has been one of the first applications in DCC,[48] its study currently extends much further. After successfully using this methodology to generate effective ligands for biomolecules,[49-51] molecular cages[52-53] and sensors;[54-55] DCC started to have an impact on some adjacent areas like catalysis,[56-57] multiphase systems,[58-59] surface chemistry,[60-61] dynamic combinatorial materials[62-63] and interlocked structures.[64-67] In nearly all 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[68] and molecular machines.[69-72]

DCC embraces diversity and complexity as an efficient mean to discover new molecules or supramolecular assemblies with unanticipated recognition properties. The DCC considers a complex system of multiple constantly interconverting compounds as an opportunity to discover new potential properties and uses. There is the prospect of discovering not only new receptors, catalysts, inhibitors, sensors, or materials, but also unexplored interactions and systems behaviors.

1.4. Dynamic covalent chemistry

As described in the previous section, in DCC the generation of a mixture of interconverting library members can be through reversible covalent reactions or specific noncovalent interactions. Dynamic covalent chemistry (DCvC)[73-75] deals with reversible covalent reactions that allow the free exchange of molecular components to achieve thermodynamic minimum of the system at equilibrium. DCvC combines the advantages typically associated with noncovalent synthesis (the formation of molecular structures using weak noncovalent interactions), such as spontaneous formation, error correction, and proof reading, with the robustness of covalent bonds.

In dynamic reaction mixtures, multiple products exist in equilibrium. Reversible assembly of molecular components generates products and semi-stable intermediates. Reactions can proceed along kinetic or thermodynamic pathways. The principal requirement for covalent bonds[74] from the thermodynamic point of view is that the bonds should be stable enough to hold molecular structures detectable and even isolable but still should have a dynamic behavior. From the kinetic point of view, short equilibration times are generally preferred for practical purposes. Additionally, mild reaction conditions are beneficial for preserving the integrity of the bond and to maintain delicate non covalent interactions of interest in the system. Bio-compatible aqueous reaction conditions are highly desired for the design of bio-inspired DCLs. Moreover, the library composition is responsive to reaction environments, such as reaction medium (solvent), physical factors (temperature, light, electric, mechanical stress, etc.), or the presence of stabilization sources (templates, metal ions, protons, etc.), so it is adaptive chemistry. The most widespread applications of DCvC have been derived from its adaptability feature. Finally, the exchange process should be easily turned off[b] to obtain kinetically inert products. The possibility of "freezing" the exchange allows the analysis, separation and handling of the library members without fear of their decomposition or further re-equilibration. Satisfying these requirements there are many reversible reactions suitable for DCvC;[37, 74] some of the most common ones are listed below in Figure 1.2.

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

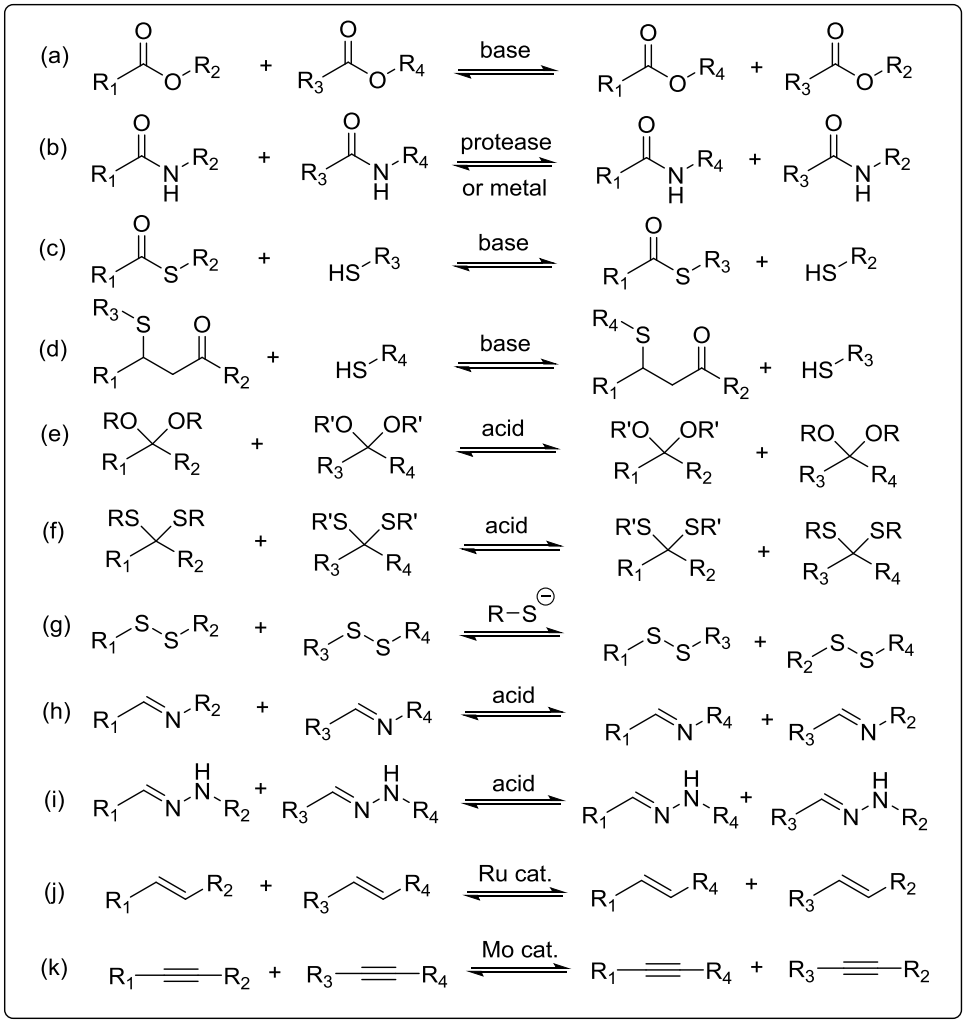


Figure 1.2. Most common reversible reactions used in DCvC. Examples of acyl transfer: (a) transesterification, (b) transamidation (c) transthioesterification and (d) Michael/retro-Michael reaction. Acetal exchange and related: (e) acetal exchange and (f) thioacetal exchange. Example of nucleophilic substitution: (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. (Figure modified from reference [37] and [39]).

A key feature of dynamic covalent bonds is the bond exchange symmetry.[76] The covalent bonds can be symmetric connections, where there is full symmetry in the exchanging bond, no directionality exists (*e.g.* disulfide bonds, Figures 1.2g and 1.3a), or directional unsymmetrical exchange, where the two functional groups constituting the exchanging reaction are different and self-inert (*e.g.* imine bonds made from the condensation of aldehydes and amines, Figure 1.2h). The combination of both exchange modes result into a trans-symmetric exchange where two connected reversible reactions work in concert so that the functional group in an unsymmetric reaction is reversibly transformed into the other.[77]

The exchange symmetry of the bonds has implications for the complexity of dynamic systems. Thus, the symmetry of the bond dictates whether the BBs 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.

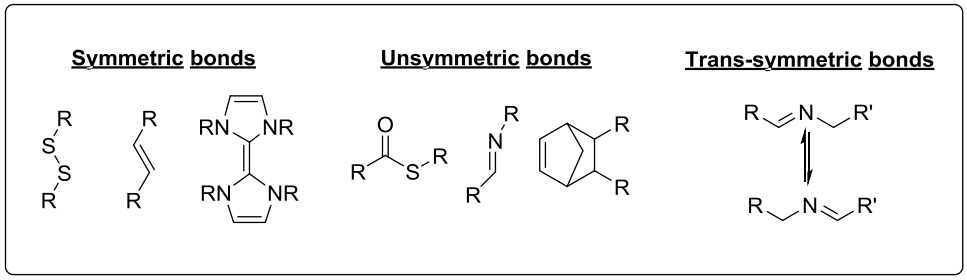


Figure 1.3. Different examples of (a) a symmetrical bond, (b) an unsymmetrical bond and (c) a trans-symmetric bond (Figure extracted from reference [76]).

1.5. Disulfide bond

Disulfide bonds are the covalent union between two substituted sulfur atoms, with the structure R−S−S−R'. The linkage is also called disulfide bridge and is usually derived by the coupling of two thiol groups. Disulfides where the two R groups are the same are called symmetric, conversely when the two R groups are not identical, the compound is said to be an asymmetric or mixed disulfide. The disulfide bonds are relatively strong, with typical bond dissociation energy of 60 kcal·mol^{-1} (251 kJ·mol^{-1}).[78] However, the disulfide bond is often the "weak link" in many molecules being about 40% weaker than C-C and C-H bonds. The distance between the two sulfur atoms is about 2.05 Å. The disulfide bond shows a conformational preference for dihedral C–S–S–C angles approaching 90°, although the rotation about the S−S axis is subject to a low energy barrier (ca. 7.5 kcal·mol^{-1}).[79] When the angle approaches 0° or 180°, then the disulfide is a significantly better oxidant.

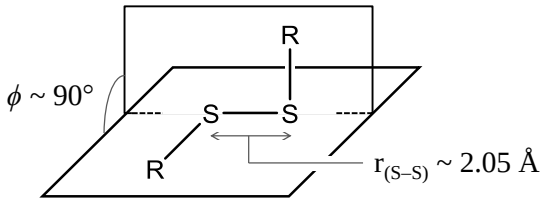


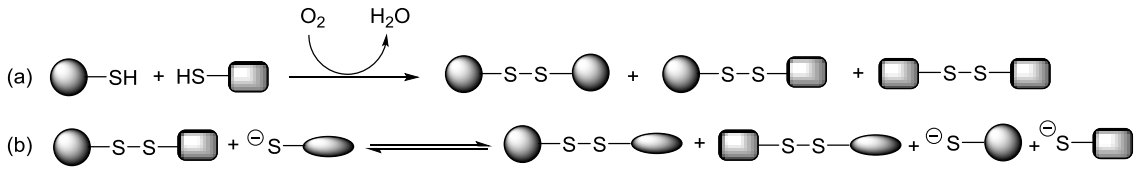
Figure 1.4. Representation of the dihedral angle (φ) and the bond length between the two sulfur atoms ($r_{(S-S)}$) of the disulfide bond.

Disulfide bonds are prevalent in biology and found in a range of biological systems, e.g. the sulfur atoms of protein cysteine residues. Disulfide bonds have a very important role in proteins where they play a key role in the development and stabilization of secondary structures and can fulfill a wide range of functions[80-81] including regulating protein activity or function and promoting protein conformational stability. Otherwise, disulfide bonds are involved in other important biochemical processes, e.g. the triggering event in the cleavage of DNA by calichemicin and esperamicin, etc.[82]

1.6. Thiol oxidation and disulfide exchange processes

Disulfide exchange is one of the most widely used reactions in DCC. As the thiol group is encountered in the amino acid cysteine, disulfide chemistry is common in biological systems. There are usually two steps involved in the preparation of DCLs of disulfides (Scheme 1.1). The

first step is the irreversible oxidation of the thiols to disulfides (Scheme 1.1a). 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.[80, 83-84] In aqueous media, it takes place at neutral or mildly basic conditions in the presence of air.[85-86] 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 presence of a catalytic amount of deprotonated thiolate is required to initiate the oxidation process.[c] 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.[87-89]



Scheme 1.1. Mechanism for disulfide exchange, (a) thiol oxidation and (b) disulfide exchange processes.

The second reaction described in Scheme 1.1 is the thiol-disulfide exchange process, also called disulfide metathesis or simply disulfide exchange. This reaction play critical roles in many aspects of cellular function and diverse biological processes, such as folding of proteins, enzyme catalysis, protection against oxidative damage, stabilization of extracellular proteins among others.[90-91] The essence of the disulfide exchange mechanism[92] involves three steps: initial ionization of the thiol to thiolate anion, nucleophilic attack of the thiolate anion on the sulfur atom of the disulfide moiety to cleave the original S–S bond and create another thiolate, and finally the protonation of the generated thiolate anion. Although it involves the cleavage and formation of a strong covalent bond, the process is reversible and allows the disulfides to self-correct to reach the most stable final situation. Similarly to the disulfide formation process (Scheme 1.1a), since exchange requires a deprotonated thiol, the reaction is highly pH-sensitive, allowing the process to be ‘‘frozen’’ by lowering the pH. In most cases, neutral or slightly basic conditions are needed to generate a sufficient thiolate concentration for the exchange process to occur.

Due to the controllability of this exchange reaction and the synthetic ease with which diverse thiol-bearing BBs can be fashioned, disulfide exchange has become one of the most commonly used reactions in DCC.[85]

1.7. 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) under the supervision of Dr. Ignacio Alfonso and Dr. Jordi Solà. The group has wide expertise in the use of peptidic and pseudopeptidic compounds for different applications, the study of processes of molecular recognition and self-assembly.[93-94] In the last years, a new

[c] Typically, the pK_a of thiols is *ca*. 8-10, although it can be considerably lower for aromatic thiols.

research line dedicated to the study of DCvC and disulfide-based dynamic combinatorial libraries has been initiated,[95-97] progress has been made in the preparation of these systems by designing simple BBs that are capable of combining to form pseudopeptidic macrocycles.

As evidenced in the following general objectives, the main idea and motivation behind this thesis is to advance in the study of dynamic combinatorial libraries with pseudopeptidic compounds previously synthetized by our group. These compounds have been combined in DCvC processes with the purpose of investigating a wide variety of structural diversity (linear, cyclic, polycyclic) and analyzing the adaptive properties of dynamic libraries in response to environment and structural effects arising from the BBs. The systems that gave rise to efficient molecular recognition processes were studied in detail. Using the knowledge acquired, we then decided to modify the corresponding buildings blocks to carry out dynamic systems with a practical application.

1.8. References

[1] J. W. Steed, J. L. Atwood, *Supramolecular chemistry*, John Wiley & Sons, **2013**.
[2] P. D. Beer, P. A. Gale, D. K. Smith, *Supramolecular chemistry*, Oxford University Press, **1999**.
[3] J. M. Lehn, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 89-112.
[4] J. M. Lehn, *Science* **1985**, *227*, 849-856.
[5] D. A. Uhlenheuer, K. Petkau, L. Brunsveld, *Chem. Soc. Rev.* **2010**, *39*, 2817-2826.
[6] J. M. Montoya, S. L. Pimm, R. V. Solé, *Nature* **2006**, *442*, 259-264.
[7] S. Azaele, S. Pigolotti, J. R. Banavar, A. Maritan, *Nature* **2006**, *444*, 926-928.
[8] S. P. Borgatti, A. Mehra, D. J. Brass, G. Labianca, *Science* **2009**, *323*, 892-895.
[9] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, D. R. White, *Science* **2009**, *325*, 422-425.
[10] N. Economides, *Int. J. Ind. Organ.* **1996**, *14*, 673-699.
[11] S. H. Strogatz, *Nature* **2001**, *410*, 268-276.
[12] R. Albert, A.-L. Barabási, *Rev. Mod. Phys.* **2002**, *74*, 47-97.
[13] C. Schalley, *Analytical Methods in Supramolecular Chemistry*, Wiley, **2007**.
[14] J. R. Nitschke, *Nature* **2009**, *462*, 736-738.
[15] O. Š. Miljanić, *Chem* **2017**, *2*, 502-524.
[16] E. Mattia, S. Otto, *Nat. Nanotechnol.* **2015**, *10*, 111-119.
[17] A. M. Valdivielso, F. Puig-Castellví, J. Atcher, J. Solà, R. Tauler, I. Alfonso, *Chem. Eur. J.* **2017**, *23*, 10789-10799.
[18] M. Kindermann, I. Stahl, M. Reimold, W. M. Pankau, G. v. Kiedrowski, *Angew. Chem. Int. Ed.* **2005**, *44*, 6750-6755.
[19] J. Stankiewicz, L. H. Eckardt, *Angew. Chem. Int. Ed.* **2006**, *45*, 342-344.
[20] G. M. Whitesides, R. F. Ismagilov, *Science* **1999**, *284*, 89-92.
[21] J. M. Lehn, *Angew. Chem. Int. Ed.* **2013**, *52*, 2836-2850.
[22] J. Li, P. Nowak, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 9222-9239.
[23] R. A. R. Hunt, S. Otto, *Chem. Commun.* **2011**, *47*, 847-858.
[24] R. F. Ludlow, S. Otto, *Chem. Soc. Rev.* **2008**, *37*, 101-108.
[25] J. W. Sadownik, D. Philp, *Angew. Chem. Int. Ed.* **2008**, *47*, 9965-9970.
[26] B. Rubinov, N. Wagner, H. Rapaport, G. Ashkenasy, *Angew. Chem. Int. Ed.* **2009**, *48*, 6683-6686.
[27] J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans, S. Otto, *Science* **2010**, *327*, 1502-1506.
[28] M. Colomb-Delsuc, E. Mattia, J. W. Sadownik, S. Otto, *Nat. Commun.* **2015**, *6*, 7427.
[29] P. Nowak, M. Colomb-Delsuc, S. Otto, J. Li, *J. Am. Chem. Soc.* **2015**, *137*, 10965-10969.
[30] G. Clixby, L. Twyman, *Org. Biomol. Chem.* **2016**, *14*, 4170-4184.
[31] K. Ruiz-Mirazo, C. Briones, A. de la Escosura, *Chem. Rev.* **2014**, *114*, 285-366.
[32] N. Wagner, G. Ashkenasy, *Chem. Eur. J.* **2009**, *15*, 1765-1775.
[33] A. P. de Silva, S. Uchiyama, *Nat. Nanotechnol.* **2007**, *2*, 399-410.
[34] C. P. Collier, E. W. Wong, M. Belohradský, F. M. Raymo, J. F. Stoddart, P. J. Kuekes, R. S. Williams, J. R. Heath, *Science* **1999**, *285*, 391-394.
[35] I. R. Epstein, J. A. Pojman, O. Steinbock, *Chaos* **2006**, *16*, 037101.
[36] I. Ferino, E. Rombi, *Catal. Today* **1999**, *52*, 291-305.
[37] J. N. Reek, S. Otto, *Dynamic combinatorial chemistry*, John Wiley & Sons, **2010**.
[38] F. B. L. Cougnon, J. K. M. Sanders, *Acc. Chem. Res.* **2012**, *45*, 2211-2221.
[39] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652-3711.
[40] J. M. Lehn, *Chem. Soc. Rev.* **2007**, *36*, 151-160.
[41] J. M. Lehn, *Chem. Eur. J.* **1999**, *5*, 2455-2463.
[42] I. Huc, J. M. Lehn, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2106-2110.
[43] P.T. Corbett, J. K. M. Sanders, S. Otto, *Chem. Eur. J.* **2008**, *14*, 2153-2166.
[44] M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, *44*, 2455-2488.
[45] P. A. Brady, R. P. Bonar-Law, S. J. Rowan, C. J. Suckling, J. K. M. Sanders, *Chem. Commun.* **1996**, 319-320.
[46] P. A. Brady, J. K. M. Sanders, *J. Chem. Soc., Perkin Trans. 1* **1997**, 3237-3254.

[47] B. Hasenknopf, J. M. Lehn, N. Boumediene, A. Dupont-Gervais, A. Van Dorsselaer, B. Kneisel, D. Fenske, *J. Am. Chem. Soc.* **1997**, *119*, 10956-10962.
[48] S. Otto, K. Severin, in *Creative Chemical Sensor Systems* (Ed.: T. Schrader), Springer Berlin Heidelberg, Berlin, Heidelberg, **2007**, pp. 267-288.
[49] M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, *53*, 3259-3263.
[50] A. Herrmann, *Chem. Soc. Rev.* **2014**, *43*, 1899-1933.
[51] P. López-Senín, I. Gómez-Pinto, A. Grandas, V. Marchán, *Chem. Eur. J.* **2011**, *17*, 1946-1953.
[52] E. Faggi, A. Moure, M. Bolte, C. Vicent, S. V. Luis, I. Alfonso, *J. Org. Chem.* **2014**, *79*, 4590-4601.
[53] I. A. Riddell, M. M. Smulders, J. K. Clegg, Y. R. Hristova, B. Breiner, J. D. Thoburn, J. R. Nitschke, *Nat. Chem.* **2012**, *4*, 751-756.
[54] L. You, J. S. Berman, E. V. Anslyn, *Nat. Chem.* **2011**, *3*, 943-948.
[55] A. Buryak, K. Severin, *Angew. Chem. Int. Ed.* **2005**, *44*, 7935-7938.
[56] H. Fanlo-Virgós, A.-N. R. Alba, S. Hamieh, M. Colomb-Delsuc, S. Otto, *Angew. Chem. Int. Ed.* **2014**, *53*, 11346-11350.
[57] P. Dydio, P. A. R. Breuil, J. N. H. Reek, *Isr. J. Chem* **2013**, *53*, 61-74.
[58] N. Hafezi, J.-M. Lehn, *J. Am. Chem. Soc.* **2012**, *134*, 12861-12868.
[59] R. Perez-Fernandez, M. Pittelkow, A. M. Belenguer, L. A. Lane, C. V. Robinson, J. K. M. Sanders, *Chem. Commun.* **2009**, 3708-3710.
[60] L. Tauk, A. P. Schröder, G. Decher, N. Giuseppone, *Nat. Chem.* **2009**, *1*, 649-656.
[61] T. Chang, D. I. Rozkiewicz, B. J. Ravoo, E. Meijer, D. N. Reinhoudt, *Nano Lett.* **2007**, *7*, 978-980.
[62] J. M. Lehn, *Angew. Chem. Int. Ed.* **2015**, *54*, 3276-3289.
[63] E. Moulin, G. Cormos, N. Giuseppone, *Chem. Soc. Rev.* **2012**, *41*, 1031-1049.
[64] N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş, J. K. M. Sanders, *Science* **2012**, *338*, 783-785.
[65] F. B. L. Cougnon, N. A. Jenkins, G. D. Pantoş, J. K. M. Sanders, *Angew. Chem. Int. Ed.* **2012**, *51*, 1443-1447.
[66] M. E. Belowich, C. Valente, R. A. Smaldone, D. C. Friedman, J. Thiel, L. Cronin, J. F. Stoddart, *J. Am. Chem. Soc.* **2012**, *134*, 5243-5261.
[67] R. T. S. Lam, A. Belenguer, S. L. Roberts, C. Naumann, T. Jarrosson, S. Otto, J. K. M. Sanders, *Science* **2005**, *308*, 667-669.
[68] S. Otto, *Acc. Chem. Res.* **2012**, *45*, 2200-2210.
[69] P. Kovaříček, J.-M. Lehn, *J. Am. Chem. Soc.* **2012**, *134*, 9446-9455.
[70] A. G. Campaña, A. Carlone, K. Chen, D. T. F. Dryden, D. A. Leigh, U. Lewandowska, K. M. Mullen, *Angew. Chem. Int. Ed.* **2012**, *51*, 5480-5483.
[71] M. Von Delius, E. M. Geertsema, D. A. Leigh, *Nat. Chem.* **2010**, *2*, 96-101.
[72] V. E. Campbell, X. De Hatten, N. Delsuc, B. Kauffmann, I. Huc, J. R. Nitschke, *Nat. Chem.* **2010**, *2*, 684-687.
[73] Y. Jin, Q. Wang, P. Taynton, W. Zhang, *Acc. Chem. Res.* **2014**, *47*, 1575-1586.
[74] Y. Jin, C. Yu, R. J. Denman, W. Zhang, *Chem. Soc. Rev.* **2013**, *42*, 6634-6654.
[75] S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders, J. F. Stoddart, *Angew. Chem. Int. Ed.* **2002**, *41*, 898-952.
[76] W. Zhang, Y. Jin, *Dynamic Covalent Chemistry: Principles, Reactions, and Applications*, John Wiley & Sons, **2017**.
[77] F. Schaufelberger, L. Hu, O. Ramström, *Chem. Eur. J.* **2015**, *21*, 9776-9783.
[78] R. Singh, G. M. Whitesides, in *Sulphur-Containing Functional Groups (1993)*, John Wiley & Sons. Inc., **2010**, p. p 633.
[79] D. A. Dixon, D. J. Zeroka, J. J. Wendoloski, Z. R. Wasserman, *J. Phys. Chem.* **1985**, *89*, 5334-5336.
[80] C. S. Sevier, C. A. Kaiser, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 836-847.
[81] N. E. Zhou, C. M. Kay, R. S. Hodges, *Biochemistry* **1993**, *32*, 3178-3187.
[82] J. Golik, J. Clardy, G. Dubay, G. Groenewold, H. Kawaguchi, M. Konishi, B. Krishnan, H. Ohkuma, K. Saitoh, T. W. Doyle, *J. Am. Chem. Soc.* **1987**, *109*, 3461-3462.
[83] R. Noiva, *Protein Expr. Purif.* **1994**, *5*, 1-13.
[84] J. C. Bardwell, J. O. Lee, G. Jander, N. Martin, D. Belin, J. Beckwith, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1038-1042.
[85] S. P. Black, J. K. M. Sanders, A. R. Stefankiewicz, *Chem. Soc. Rev.* **2014**, *43*, 1861-1872.
[86] J. P. Tam, C. R. Wu, W. Liu, J. W. Zhang, *J. Am. Chem. Soc.* **1991**, *113*, 6657-6662.

[87] T. M. Postma, F. Albericio, *Eur. J. Org. Chem.* **2014**, *2014*, 3519-3530.
[88] D. J. Cline, C. Thorpe, J. P. Schneider, *Anal. Biochem.* **2004**, *335*, 168-170.
[89] L. Chen, I. Annis, G. Barany, *Current Protocols in Protein Science* **2001**, *23*, 18.6.1-18.6.19.
[90] T. V. DeCollo, W. J. Lees, *J. Org. Chem.* **2001**, *66*, 4244-4249.
[91] W. J. Wedemeyer, E. Welker, M. Narayan, H. A. Scheraga, *Biochemistry* **2000**, *39*, 4207-4216.
[92] P. A. Fernandes, M. J. Ramos, *Chem. Eur. J.* **2004**, *10*, 257-266.
[93] S. V. Luis, I. Alfonso, *Acc. Chem. Res.* **2013**, *47*, 112-124.
[94] I. Alfonso, *Chem. Commun.* **2016**, *52*, 239-250.
[95] J. Atcher, I. Alfonso, *RSC Advances* **2013**, *3*, 25605-25608.
[96] J. Atcher, A. Moure, I. Alfonso, *Chem. Commun.* **2013**, *49*, 487-489.
[97] J. Atcher, A. Moure, J. Bujons, I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 6869-6878.

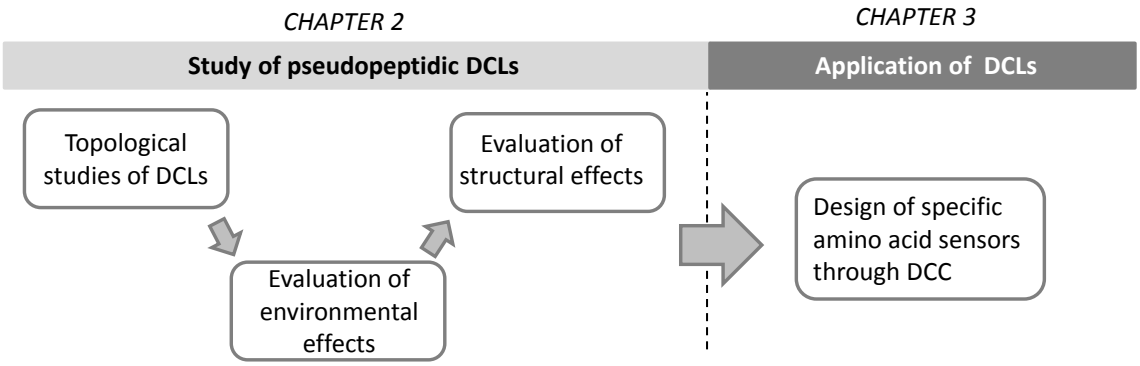
GENERAL OBJECTIVES

The general objective of the present thesis was to advance in the study of disulfide-based dynamic combinatorial libraries formed by pseudopeptidic compounds synthesized previously in our group. The novelty of the concept lies on the formation of chemical libraries including components with different number of functionalities that can form dynamic bonds (monopodals, bipodals, tripodals). These combinations can potentially lead to a large structural diversity of chemical species (linear, cyclic, polycyclic) by suitable combination of the corresponding pseudopeptidic BBs in dynamic processes. Therefore, the first objective proposed was to study the effect of geometry and chemical valence in the composition of dynamic libraries. The generation of molecular diversity opens the possibility of creating highly complex and efficient molecular recognition systems, starting from very simple BBs.

A second objective was to deepen in the adaptive properties of pseudopeptidic dynamic libraries in response to different environmental stimuli (pH, solvent polarity, ionic strength, concentration of the BBs, etc.) and to structural effects arising from the BBs. This last point resulted into another research line where we expanded the study of structural characteristics in pseudopeptidic libraries, which can selectively self-organize into complex structures and according to well-defined structural parameters. This has led to the study of these structural parameters through different techniques. The systems that emerged from the different DCLs that give rise to efficient molecular recognition processes was studied in detail. These studies allowed us to obtain a broader knowledge of the behavior of the corresponding dynamic systems.

Finally, based on the knowledge acquired, we considered modifying the corresponding BBs to perform dynamic systems with a practical application, and specifically we focused on the synthesis and study of specific chemical sensors of amino acids in aqueous media. The aim was to achieve for the first time the preparation of a dynamic system capable of producing a fluorescent response in presence of L-Cysteine (Cys) or L-cystine in biological fluids.

One of the main challenges of this project is to work in aqueous media in order to optimize systems that can have a direct application in biological media. This implies by itself an ambitious goal, considering that water is a highly competitive media for polar noncovalent interactions.



The specific general objectives are:

1) Synthesis and characterization of pseudopeptidic BBs capable of generating dynamic combinatorial libraries in aqueous media.
    a. Design and synthesis of systems with different valence: monopodals, bipodals, tripodals.
    b. Design and synthesis of systems with different functional groups capable of giving rise to different supramolecular interactions.
2) Preparation and study of the corresponding dynamic combinatorial libraries with the previously synthetized BBs.
    a. Study a wide range of different dynamic combinatorial libraries with different valence in aqueous medium.
    b. Specific study of the adaptive behavior of these libraries.
3) Explore the effect of different external stimuli (pH, ionic strength, polarity of the solvent, concentration of the constituent blocks, etc ...) in the composition of DCLs.
4) Study of the structural effects (modifying one or more components of the initial library) in the composition of DCLs and in the proportion of their final products.
5) Design and synthesis of specific amino acid sensors through DCC in aqueous media. Specifically, sensors that respond selectively to the presence of L-Cys and L-cystine through self-recognition processes and give a quantitative response by spectrophotometric measurements (fluorescence).

CHAPTER II

Study of pseudopeptidic DCLs

CHAPTER II: Study of pseudopeptidic DCLs

2.1. Introduction

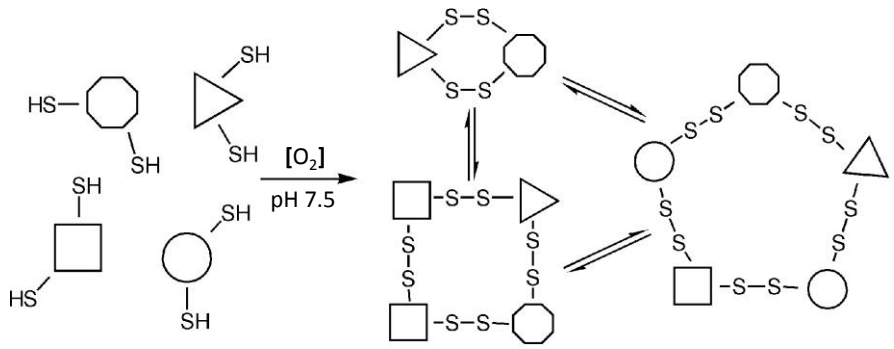
2.1.1. Disulfide based DCLs

Disulfide exchange has become one of the most widely used reactions in DCC,[1] due to the controllability of this exchange reaction and the mild reaction conditions compatible with biomolecules in aqueous media. Disulfide-based DCLs are commonly generated simply by dissolution of thiol BBs in water at pH 7-9 under air.

The essence of the disulfide exchange mechanism involves two processes (explained in the General Introduction): the thiol oxidation and the disulfide exchange. The exchange process is reversible as long as thiolate anions are present in solution, but the oxidation process is irreversible, and the exchange stops after the BBs are fully oxidized. Therefore, when all thiols have been oxidized to disulfides, typically within several days,[2] the process ends. Since exchange requires a deprotonated thiol, the reaction is highly pH-sensitive, allowing the process to control the thiolate concentration and thus the exchange ratio through the pH.

If the library reaches a thermodynamic equilibrium, the concentration of each of the members will be dictated by their relative stabilities. To achieve this purpose, it is necessary that the generation of the DCLs occur in conditions in which the exchange reaction is faster than the oxidation of the thiols, since the final situation is not subjected to additional equilibration processes. In contrast, 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.[3] The almost immediate disulfide formation prevents the disulfide exchange to significantly take place in the process, thus reaching a non-equilibrated final situation, i.e. a kinetic trap.

One of the firsts cases reported using equilibrium mixtures of disulfides was described facing the challenge of defining the two dimensional structure of biological membranes. Specifically, a new approach to study the lateral organization of the lipid mixing in bilayers had been reviewed.[4] Although it was increasingly common to work with simple mixtures of disulfides,[5] it was not until 2000 when it was reported for the first time the successful use of the disulfide exchange reactions for the generation of reversible DCLs. Otto, Furlan and Sanders[2] demonstrated that DCLs of structurally diverse macrocyclic disulfides can be generated in a single step using a variety of simple dithiol BBs under extremely mild conditions and with no external catalyst needed for the exchange process (Scheme 2.1). The authors were inspired by the reported use of macrocyclic disulfides as receptors and as synthetic ionophores[6-8] and started to investigate the possibilities of disulfide-based DCLs for molecules with these properties. From these studies the disulfide chemistry emerged as an important addition to the limited number of reactions that were available for the generation of DCLs until then.



Scheme 2.1. Formation of a DCL of macrocyclic disulfides from dithiol BBs (figure modified from reference [2]).

As the field of DCC advanced, supramolecular chemists further explored the use of reversible disulfide bonds, as stated above, often in water at pH 7–8, to construct and study complex molecular architectures (and combinatorial libraries thereof) whose conventional synthesis would be either very tedious or impossible to achieve. Only through the dynamic nature of DCC impressive examples have been reported in which the disulfide chemistry allows the preparation of macrocycles,[3, 9-13] catalysts,[14-16] capsular molecules,[17] catenanes,[18-21] molecular knots[22-23] and even self-replicating molecules.[24-27]

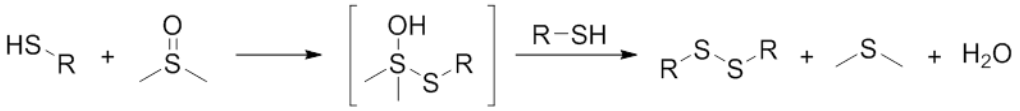
It is also remarkable the fact that in disulfide systems, DCLs commonly display great sensitivity to templation effects, where addition of a new molecule alters their equilibrium position through supramolecular interactions with library constituents. This is because a small change in free energy leads to a logarithmically amplified change in equilibrium position; $\Delta G = -RT \ln K$. Disulfide chemistry offers many possibilities to design systems which lead to complex structural outcomes through the templation sensitivity to chemical interactions and environmental stimuli. In fact, the varied synthetic uses of disulfide chemistry are matched by an equally extensive diversity of observed structures, exhibiting interesting properties and functions.

2.1.2. DMSO as an organic co-solvent

Dimethyl sulfoxide (DMSO) is a polar aprotic solvent that has widely been used as an organic co-solvent because of its large solubilization capacity and its relatively low toxicity. It is one of the most powerful readily available organic solvents, being able to dissolve a large variety of substances, including both polar and non-polar compounds. Water/DMSO mixtures have important industrial applications and aqueous mixtures with small amounts of DMSO are even used in biological applications, such as in the field of drug discovery.[28] Probably, the main objection or limitation associated to the use of DMSO is its high boiling point (189 °C) which makes very difficult to recover the compounds dissolved in this organic solvent.

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 case of precipitation, the composition of the whole library could be shifted towards the formation of the precipitate, obtaining a kinetic trap. The formation of disulfides in aqueous solvent with DMSO helps to ensure the total solubility of the members of the library. The use of DMSO allows faster oxidation rates in contrast with the relatively long reaction times

required for the disulfide formation by air oxidation. In the proposed reaction pathway,[29] a thiol molecule combines with a molecule of DMSO in the rate determining step to form a sulfoxide-thiol adduct (Scheme 2.2). Then, the adduct reacts more rapidly with another molecule of thiol to form the observed products: the corresponding disulfide, dimethyl sulfide and water. Therefore, the reaction can be performed in a wide range of pH, in contrast with the conventional oxidation by air that is applicable only at a narrow basic pH range.[30]



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. In contrast with the previously described oxidation by air, when DMSO is used as an oxidant the presence of deprotonated thiolate anions is not a required condition and the reaction can also proceed at acidic pH. Additionally, different reaction solvents have also been described to have an effect on the disulfide exchange reaction (Scheme 1.1b). A higher polarity of the solvent involves a larger the stabilization of the reactants in relation to the transition state of the disulfide exchange mechanism. In this regard, hydrophobic environments are known to decrease the barrier energy, catalyzing the process.[31] 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 or methanol.[32-33]

A few years ago, in our group, an extensive study of the use of DMSO as a co-solvent for the preparation of disulfide-based DCLs was carried out by Dr. Joan Atcher,[10] opening a wide range of possibilities in the DCC.[d] Very interestingly, this study reported the beneficial effect of the use of DMSO as a co-solvent on the thiol-disulfide DCvC. In this work, the suitable experimental conditions in which the disulfide-based DCLs can reversibly combine were stablished, generating useful dynamic combinatorial libraries under thermodynamic control over a wide range of pHs. Thus, the suitable use of DMSO as a co-solvent to shorten the oxidation times required to prepare the dynamic libraries was demonstrated (Figure 2.1). Interestingly, the reactions performed in DMSO:water mixtures proceeded faster than in either pure DMSO or pure water. Moreover, for different DMSO:water proportions, the higher content of DMSO also induces a slightly faster oxidation. Most importantly, the foundations for all the future studies of our group in disulfide-based DCLs were laid and, therefore, the majority of the studies comprehended in this thesis were performed in aqueous medium with 25 % DMSO to obtain faster oxidation and exchange processes and to ensure the total solubility of all species involved.

[d] To the best of our knowledge, before our group began to carry out these studies, 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.[34] A. V. Gromova, J. M. Ciszewski, B. L. Miller, *Chem. Commun.* **2012**, *48*, 2131-2133. 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.

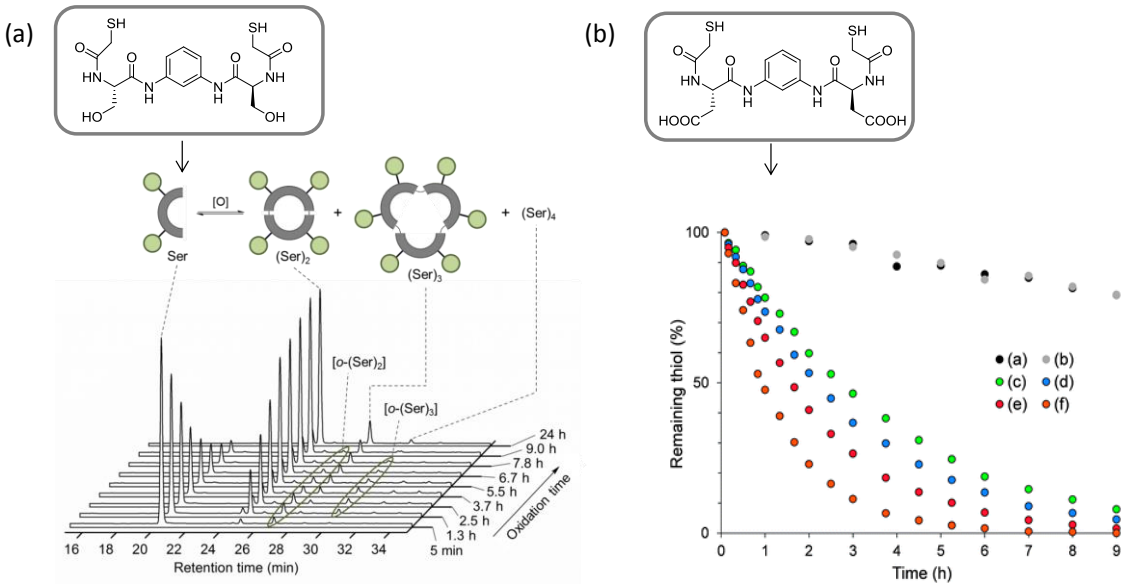


Figure 2.1. (a) Schematic representation and HPLC-UV traces (254 nm) at different reaction times of the oxidation process for 2 mM Ser-derivate pseudopeptide in aqueous phosphate-citrate buffer (pH 7.5) with 25% (v/v) DMSO. (b) Plot of the remaining thiol (%) over the oxidation time *(h)* for 2 mM Asp-derivate at different oxidation conditions: 0% DMSO at pH 7.5 (●); pure DMSO (●); 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 (●) (figure modified from reference [10] and Dr. Joan Atcher Thesis).

To sum up DMSO offers: i) better solubility of the libraries, ii) rapid oxidation, iii) fast disulfide exchange and iv) strong solvation properties being the solvent more similar to the effect of water.

2.1.3. Precedents of topologically diverse networks in DCC

As stated in the general introduction, DCC proposes the creation of dynamic chemical networks formed by simple members inter-connected through reversible chemical processes.[35-37]

Dynamic complex systems are excellent benchmark models for decoding the chemical information stored in the combination of simple BBs towards the emergence of new assembled structures.[25, 38-39] These systems have a huge potential for delivering new applications in areas ranging from materials science to medicine. The complexity, responsiveness and dynamicity of molecular synthetic systems may contribute to developing a better understanding of the complex molecular networks encountered in nature.[40-41] In this context, DCC provides a valuable methodology to generate and study complexity, and disulfide exchange is particularly relevant as it allows the formation of dynamic chemical libraries and molecular networks in aqueous media. Dynamic libraries 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 catenanes,[18-21] rotaxanes,[42-43] pentafoil knots[44] and daisy chains.[45-46]

Despite this impressive unexpected geometries, the tridimensional space covered is relatively small because the libraries are usually formed by difunctional components that are mostly directed towards the generation and interconversion of cyclic oligomers.[47-48] The formation of cage-like organic structures has seldom been described by oxidation of tripodal BBs or by the combination of tripodal and bipodal BBs.[49] In figure 2.2, there are two examples of the few multivalence systems described in DCC, both reported by Sanders group. In the first case (figure 2.2a), the discovery of a new class of organic cages assembled from simple tri- and dithiol BBs using DCC was described.[17] This consisted in the first example of a DCL where the amplification of up to eleven components assembled into a well-defined cage-like structure appeared but the stoichiometry and the presence of a template were critical. These water-soluble disulfide-linked architectures are only generated upon templation by positively charged polyamine guests of the appropriate shape and length such as spermine and spermidine. More recently, the same group described diverse topologies obtained by a combination of tri- and monothiols in water (figure2.2b).[50] This publication provides the first example of a DCL in which a subtle balance between the number of reversible linkages and the structural characteristics of the functional groups lead to a topologically diverse system with at least three distinct types of architectures.

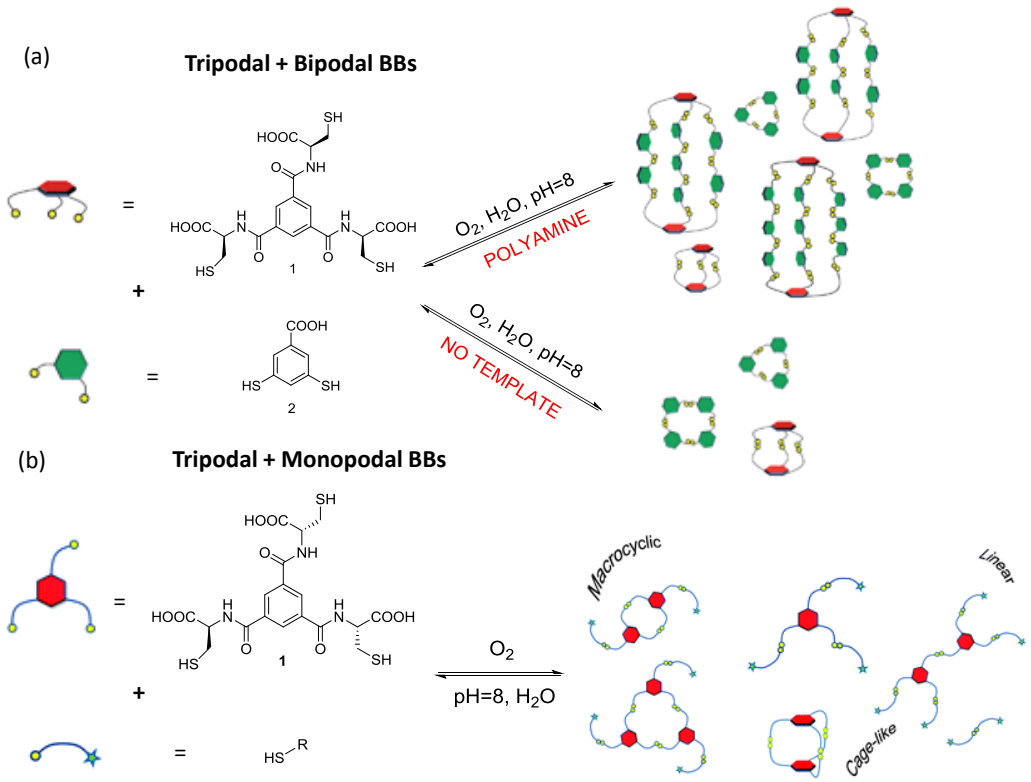


Figure 2.2. Schematic representation of DCLs with multivalence systems, both reported by Sanders group. (a) tripodal and bipodal and (b) tripodal and monopodal BBs in the same DCL (figure extracted from reference [17] and [50]).

A more recent example of topologically diverse DCLs was reported by Ulrich and co-workers.[51] They reported a new class of tripodal and tetrapodal aromatic cage-type compounds that are self-assembled in one-pot using simultaneous dynamic covalent reactions employing acylhydrazone and disulfide bond formation. Implementation of three distinct functional groups (thiols, aldehydes and hydrazides) in the structure of two simple BBs

resulted in their spontaneous and selective self-assembly into aromatic cage-type architectures (figure 2.3). These molecular structures consist of up to ten components connected through up to twelve dynamic covalent bonds. In their publicaton, they showed two examples of structurally distinct cage systems: a tripodal cage architecture and a fluorescent tetrapodal cage; the assembly and disassembly of the latter can be monitored by fluorescence spectroscopy. The exclusive formation (self-selection) of these complex structures from dynamic combinatorial libraries may be the result of intramolecular interactions and/or geometric/connectivity fit which would also explain the narcissistic self-sorting behavior observed. This is the first example of multicomponent cage formation in aqueous media by applying simultaneously two distinct dynamic covalent bonds.

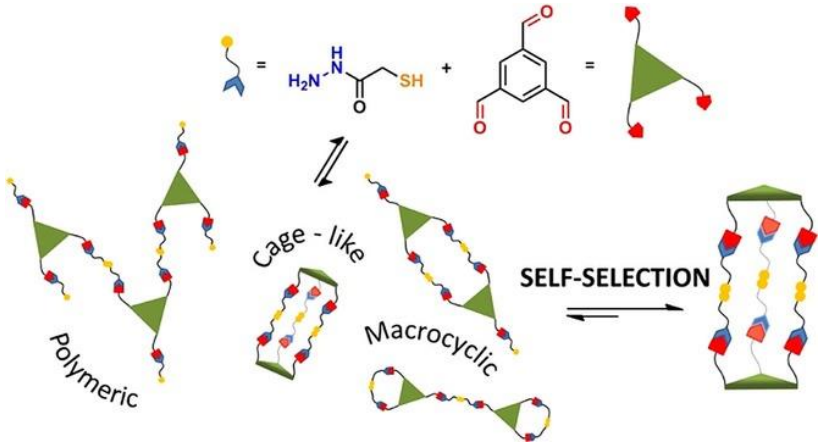


Figure 2.3. Generation of dynamic combinatorial libraries from aromatic aldehyde and thioglycolic hydrazide using acylhydrazone and disulfide reversible covalent reactions, and examples of the various polymeric, macrocyclic and cage constituents that can be expressed in the system. (figure extracted from reference [51]).

Generally in disulfide-based chemistry, relatively simple thiol entities are synthetized to be used as BBs for the generation of DCLs by means of the disulfide formation and methatesis. In order to expand the complexity of the geometries and topologies involved we decided to study libraries where molecules with different number of thiols are present (we could consider therefore combinations of compounds with different valence). In the present Thesis, a combination of BBs carrying different structural and chemical information was envisioned. Thus , the structure of the BBs consists of three differentiated parts (figure 2.4): i) an aromatic moiety (in black) as a linker to join the different constituents of the structure, being a characteristic chromophore, ii) amino acid frameworks (in green) as a source of chemical diversity that can contain the group responsible to generate dynamic bonds (thiol group). Except for the thiol-containing amino acid cysteine an extra step is needed which consists in the addition of the exchange triggering thiol groups incorporated by the addition of mercaptoacetyl frameworks (red part).

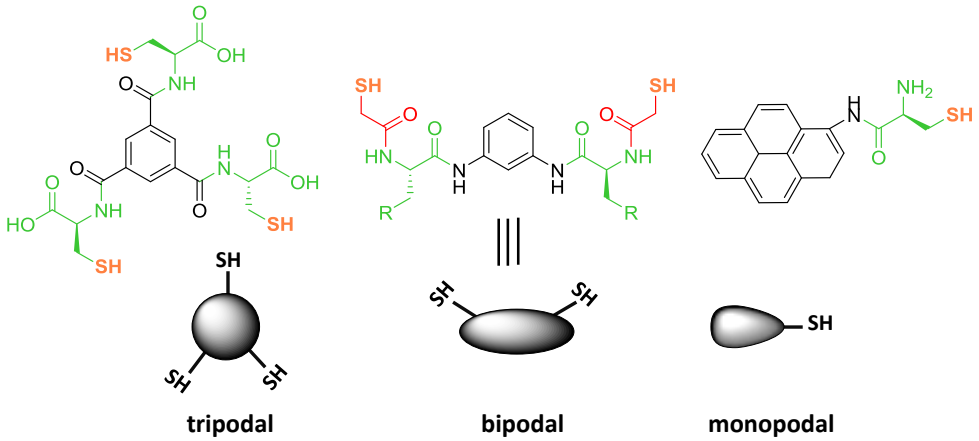


Figure 2.4. Schematic representation of different examples of structurally diverse BBs.

2.1.4. Adaptive behavior of DCLs

As previously described, self-assembling of complex molecular systems is dictated by the chemical and structural information stored in their components. This information can be expressed through an adaptive process that determines the structurally fittest assembly under given environmental conditions. Accordingly, the distribution of products can be altered by the presence of external or intrinsical stimuli. Moreover, because the networks are intrinsically complex and dynamic, new properties may arise from the mixture. The careful study of dynamic combinatorial libraries (DCLs) provides an excellent model to help chemists to understand what factors govern the formation of a given structure.

2.1.4.1. Environmental effects as an external stimulus

In a DCL where the distribution changes upon variation of external conditions, amplification of a particular species within the library is indicative of its greater stability, relative to the other library members, upon 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,[52-55] ionic strength,[19-23, 56-62] temperature,[53, 56-57, 63-64] light,[65-67] gelation,[68-70] mechanical forces[26] and presence of an electric field.[71-72] However, environmental effects such as pH or ionic strength have seldom been explored by scientists working on DCC. In the case of pH, it can probably be explained by the precise conditions that are required to generate libraries, particularly a slightly basic pH is needed to generate DCLs based on disulfide chemistry. The following lines briefly comment on the main reported examples of the use of pH and ionic strength stimulus.

As stated above, there are very few studies about the effects of the pH in dynamic libraries. However, Lehn and co-workers were pioneering in the study of adaptive effects of DCLs induced by chemical and/or physical stimuli and, among others, investigated the effects of increasing the acidity of the medium, moving from the neutral or slightly basic pH normally used, with organic acids like CF_3COOH. Previous studies of Lehn's group showed a synergistic adaptive behavior of DCLs, that is the case of their studies about the selection processes within DCLs. In those libraries, proton concentration was able to rearrange the composition of

equilibrating libraries of imines and Zn^{2+} metal ions interacted with the library constituents, strongly affecting its composition (explained in detail in the next Chapter).[55] Later, Giuseppone and Lehn investigated how the product distribution in a DCL of imines can be directed by the variation of temperature and pH (Figure 2.5a).[53] In this study, they reported a DCL formed by four different components (**1-4**) where depending on the temperature or the acidity of the medium a different distribution of the imine products (**A-D**) was obtained. In function on these parameters, the main product varied among the four possible imines formed (Figure 2.5b). The complexity of the effects does not allow a detailed quantitative explanation in terms of structure and mechanism.[e] However, the results illustrate the response of such a dynamic system to a physical stimulus (temperature) and a chemical effector ($[H^+]$), thus demonstrating the adaptive behavior of the system under the pressure of external factors. They also pointed out the possibility of modulating a given functional property (optical, electronic, ionic) by constitutional recomposition induced by a specific trigger. Their work highlights the perspectives opened by constitutional dynamic chemistry towards the design of smart materials, capable of expressing different latent properties in response to environmental conditions.

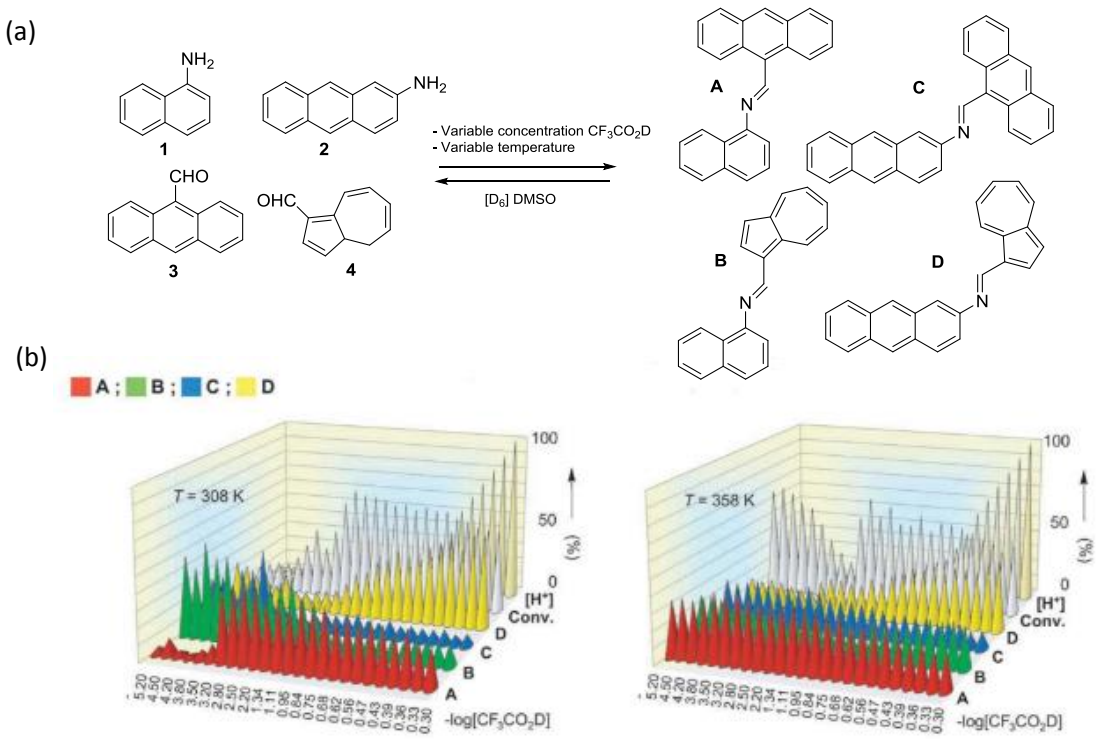


Figure 2.5. (a) Dynamic library of the four components **1–4** and the four constituents **A–D**. (b) Distributions of imines **A–D** at two temperatures as a function of acidity $[H^+]$ (figure modified from reference [53]).

There are relatively few examples where the increase of the ionic strength resulting from the addition of an inorganic salt has shown a remarkable influence on the composition of DCLs generated in aqueous media.

[e] The basicities in DMSO are not known. For comparison, the pKa values in aqueous ethanol are: 3.4, 3.4 and 4.05.

The use of the ionic strength as an external stimulus was reported for the first time in the generation of mechanically interlocked molecular architectures in water. Sanders' group investigated the use of aromatic and hydrophobic interactions between library components made from naphthalene diimide (NDI) acceptor and dialkoxynaphthalene (DN) donor BBs (e.g. Figure 2.6). Surprisingly, the study of these libraries generated by disulfide exchange led to the discovery of donor-acceptor [2]-catenanes,[18, 20-21, 60-62] [3]-catenanes[19, 59] and even giant species (Figure 2.6),[59] showing the utility of DCC for preparing topologically complex molecules.

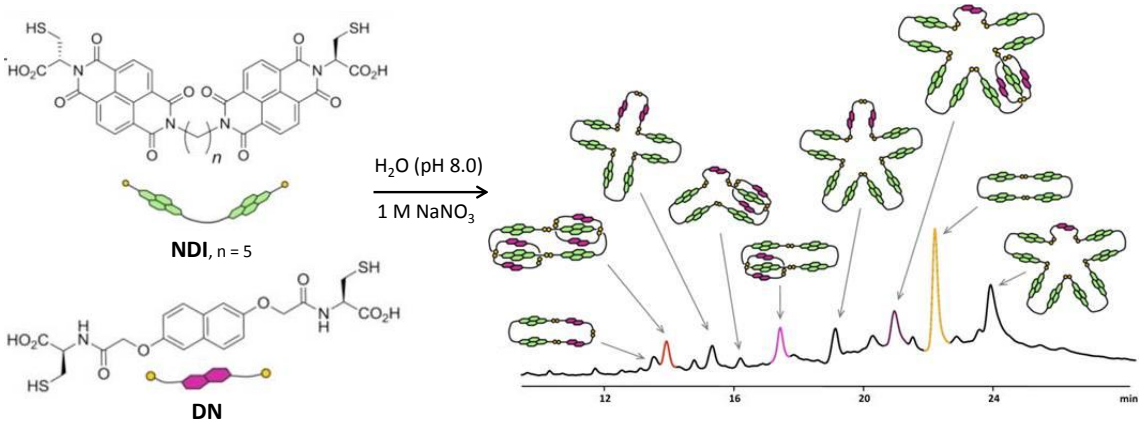


Figure 2.6. HPLC trace (383 nm) of a DCL generated from the mixture of a naphthalene diimide (NDI) acceptor and a dialkoxynaphthalene (DN) donor BBs (2.5 mM each) in aqueous medium (pH 8.0) in the presence of 1 M $NaNO_3$ (figure modified from reference [59]).

The catenane assembly is controlled by a fine balance between kinetics and thermodynamics and for that is essential the control of various structural parameters of the BBs (linker length, flexibility, chirality) as well as some external factors. The ionic strength resulted to be a crucial parameter to achieve such impressive complex molecules; the strength of the hydrophobic interaction is known to be enhanced by increasing the concentration of the electrolytes, i.e. increasing the ionic strength.[73] 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 2.6). 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.

More recently, our group has reported the effects of salt concentration on libraries of macrocyclic pseudopeptides[12] and represents one of the first examples of the use of this external factor as a bioinspired stimulus. These studies expose that in the absence of salt, a dynamic library is highly dependent on the electrostatic interactions showing a composition markedly different from the statistical distribution. The increase in ionic strength shields these interactions and the system approaches the statistically favored proportion (Figure 2.7a). The corresponding binary mixtures of the library contain the irreducible information of all the co-adaptive relationships present within the complex library. The careful analysis of these binary mixtures together with structural studies performed for selected members of the library,

allowed unravelling interesting adaptive trends. Firstly, for the species bearing charges of the same sign, the salt-adaptation was inversely related to the length of the side chain. Secondly, negatively charged homodimers were 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. The deep understanding of the behavior of the system as a network has brought the opportunity to set up dynamic libraries displaying intended relationships. Thus, either competitive or cooperative co-adaptive trends have been reproduced as a response to the salt increase (Figure 2.7b). In summary, the authors have demonstrated that the increase of the ionic strength within simple DCLs of macrocyclic pseudopeptides induces the amplification of the members concentrating a large number of acidic side chains. This behavior has a remarkable resemblance with the natural evolution of the proteins of halophilic organisms to survive in hypersaline media. These artificial molecular ecosystems demonstrate the potential of DCLs to mimic important processes occurring in more elaborate (bio)molecular networks.

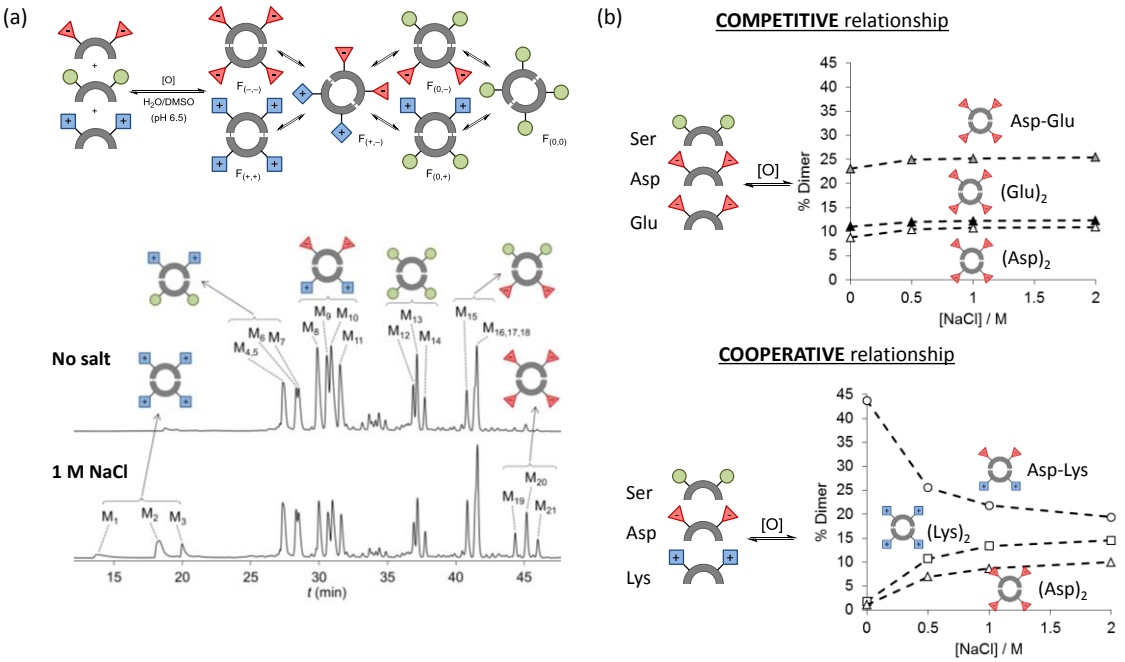


Figure 2.7. (a) HPLC-UV traces (254 nm) of the DCL generated from the mixture of the six bipodal BBs in aqueous Bis-Tris buffer with 25% (v/v) DMSO at pH 6.5 in the absence of salt (upper trace) and in the presence of 1m NaCl (lower trace). (b) Plot of the percentage concentration of rationally designed co-evolutionary systems of three different charge BBs against the concentration of NaCl (M) (figure modified from reference [12] and Dr. Joan Atcher Thesis).

After these findings, we were prompted to study the effects of the media on potentially complex libraries. Moreover, we extended the previous investigations of our group to put the focus on the variables that drive the system to its final composition and wanted to analyze the complex molecular rearrangements that occur during the oxidation and equilibration processes. Our studies comprehended in this Chapter highlight the delicate conditions that produce effective recognition between counter parts.

2.1.4.2. Structural effects as an intrinsic stimulus

Intrinsic factors that arise from the structure of the BBs play an important role on the product distribution in a DCL and their modification can alter completely the final composition, leading to a possible molecular recognition within the BBs. Therefore, subtle structural changes can be decisive in the outcome of a DCL and it is expected that suppression of possible interactions or the establishment of new ones will clearly modify the composition of the library.[74] This effect has been exploited for the dynamic combinatorial discovery of synthetic receptors[9, 75-78] and ligands for biomolecules[79-80] by exposing the system to a corresponding template. The template-induced amplification of specific DCL members has been widely studied.[81-86] Nevertheless, other molecular recognition process that take place in the absence of added templates, occurring between library members have been explored, leading to interlocked structures[20, 22-23] self-replicating molecules[26, 87-91] and self-assembling materials.[24, 92-93] In these systems, noncovalent interactions shift the equilibrium toward molecules that engage most efficiently in noncovalent interactions. In these cases, for example a specific chemical function or a suitable length on the side-chain of the BBs could result in profound changes in the nature of the products formed. These structural effects clearly highlight the importance of having the correct interactions for the selection of one specific geometry or another.

A clear example of the importance of subtle structural changes is the use of the chiral recognition between enantiomeric pairs that leads to self-recognition or self-discrimination,[94-95] depending on whether an enantiomer preferentially recognizes itself or its mirror image, to generate homo- or heterochiral species respectively. In recent years, DCvC has been proposed to enable chiral self-sorting via covalent bond formation. In this framework, our group has worked previously in macrocyclization via DCvC with inherent selectivity for homochiral products in solution.[13] The aim for chiral self-sorting within a DCL of flexible and bio-inspired pseudopeptidic BBs was an extraordinary and ambitious challenge. These studies reported a simple DCL consisting of homo- and heterochiral dimeric pseudopeptides which was successfully used to study the chiral self-sorting phenomenon. The mixture of BBs (SS)-BB and (RR)-BB leads to the formation of a disulfide-based DCL consisting of two homochiral dimers and one heterochiral dimer. 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.

Self-replicating molecules are another example of the important effects of a delicate structural variation on the emergence of self-recognition process. In this case if two or more molecules of a particular product can stabilize one another through noncovalent binding, the equilibrium will shift toward formation of this product at the expense of the other library members. Otto and co-workers have extensive expertise in self-replicating molecules; previously they proposed that it should be possible to use dynamic combinatorial libraries to develop molecules capable of promoting their own formation, while forming extended assemblies at the same time.[37] In a remarkable example, they reported two self-replicating

peptide-derived macrocycles that emerge from a small DCL and compete for a common feedstock.[25-26] A dynamic network was generated by the oxidation of the dithiol BB shown in Figure 2.8a, rendering a library of oligomers initially dominated by the corresponding trimeric and tetrameric macrocyclic species (Figure 2.8b). Very interestingly, after some days the hexameric and heptameric species suddenly emerged, growing exponentially to quickly dominate the library (Figure 2.8c). These two species were found to aggregate into thin micron-long fibers (Figure 2.8d). 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. When they investigated the nature of the interactions holding the fibers together, the circular dichroism (CD) spectra of solutions containing mainly hexamer or heptamer show most of the features typical for β - sheet formation, whereas the spectrum of a solution containing mostly trimer and tetramer shows features typical for random-coil peptides. The authors were able to give a satisfactory explanation for such an unexpected behavior, and additionally examined the effect of different peptide chains[24] and solvent compositions.[96]

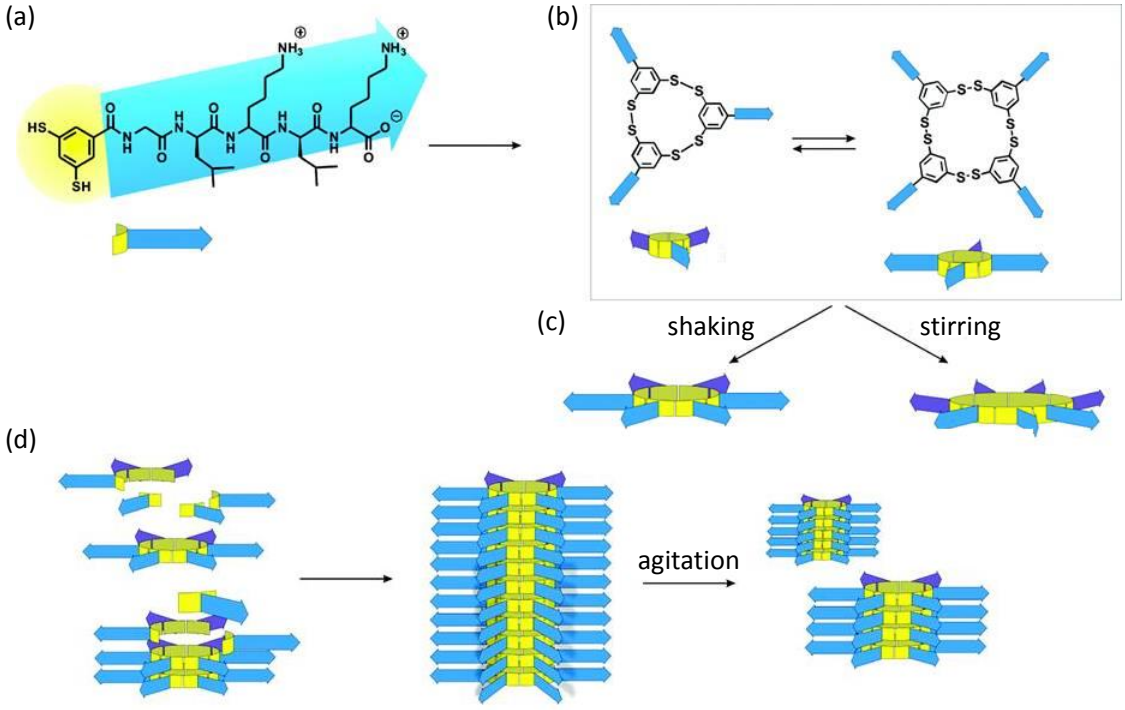


Figure 2.8. (a) Peptide-functionalized dithiol BB, that initially gives rise to (b) a mixture dominated by cyclic trimer and tetramer. (c) Upon agitation a conversion to larger macrocycles takes place. (d) Schematic representation of the stacks of the hexamer, held together by β-sheets formed by the peptide chains. Breaking the fibers by agitation increases the number of ends and promotes fiber growth (figure extracted from reference [25]).

The systematic study of the equilibrium composition of a DCL depending on intrinsic factors rendered important information about noncovalent intramolecular interactions. However, in contrast to the well-established theory of template-induced amplification, the theoretical understanding on how the selection of different modes of self-assembly relates to covalent selection remains underdeveloped.

If we understand the behavior of the interactions that are responsible of the formation of different products and learn what factors are leading to their formation, we will decode the key to direct the thermodynamic equilibrium towards a desired formation of the products. The supramolecular complexes open a highly promising way to analyze noncovalent interactions in detail. Systematically planned variations of the BB structures and of the experimental conditions can provide a basis for a better understanding of supramolecular association in natural and synthetic systems and for the theoretical prediction of noncovalent interactions. Additionally, there are novel studies that correlate binding data with empirically or computationally derived factors, that can help to characterize the nature of the essential noncovalent interactions,[97] and result into an excellent tool to predict the formation of compounds and to contribute to the understanding of biologically important associations.

This fact further confirms DCC as an extremely useful tool for systems chemistry research. Furthermore, structural insights regarding noncovalent interactions and folding properties of peptide-like molecules in highly competitive media can be extracted from carefully designed DCC assays. The further use of this knowledge should lead to the discovery of new materials for which structural selection will translate into functional selection.

2.1.5. References

[1] S. P. Black, J. K. M. Sanders, A. R. Stefankiewicz, *Chem. Soc. Rev.* **2014**, *43*, 1861-1872.
[2] S. Otto, R. L. E. Furlan, J. K. M. Sanders, *J. Am. Chem. Soc.* **2000**, *122*, 12063-12064.
[3] F. Ulatowski, A. Sadowska-Kuziola, J. Jurczak, *J. Org. Chem.* **2014**, *79*, 9762-9770.
[4] S. M. K. Davidson, S. L. Regen, *Chem. Rev.* **1997**, *97*, 1269-1280.
[5] H. Hioki, W. C. Still, *J. Org. Chem.* **1998**, *63*, 904-905.
[6] D. Ranganathan, V. Haridas, I. L. Karle, *J. Am. Chem. Soc.* **1998**, *120*, 2695-2702.
[7] D. Ranganathan, V. Haridas, K. P. Madhusudanan, R. Roy, R. Nagaraj, G. B. John, M. B. Sukhaswami, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1105-1107.
[8] D. Ranganathan, V. Haridas, S. Kurur, R. Nagaraj, E. Bikshapathy, A. C. Kunwar, A. V. S. Sarma, M. Vairamani, *J. Org. Chem.* **2000**, *65*, 365-374.
[9] S. Otto, R. L. E. Furlan, J. K. M. Sanders, *Science* **2002**, *297*, 590-593.
[10] J. Atcher, I. Alfonso, *RSC Advances* **2013**, *3*, 25605-25608.
[11] J. Atcher, A. Moure, I. Alfonso, *Chem. Commun.* **2013**, *49*, 487-489.
[12] J. Atcher, A. Moure, J. Bujons, I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 6869-6878.
[13] J. Atcher, J. Bujons, I. Alfonso, *Chem. Commun.* **2017**, *53*, 4274-4277.
[14] H. Fanlo-Virgós, A. N. R. Alba, S. Hamieh, M. Colomb-Delsuc, O. Sijbren, *Angew. Chem. Int. Ed.* **2014**, *53*, 11346-11350.
[15] L. Vial, J. K. M. Sanders, S. Otto, *New J. Chem.* **2005**, *29*, 1001-1003.
[16] B. Brisig, J. K. M. Sanders, S. Otto, *Angew. Chem. Int. Ed.* **2003**, *42*, 1270-1273.
[17] A. R. Stefankiewicz, M. R. Sambrook, J. K. M. Sanders, *Chem. Sci.* **2012**, *3*, 2326-2329.
[18] R. T. S. Lam, A. Belenguer, S. L. Roberts, C. Naumann, T. Jarrosson, S. Otto, J. K. M. Sanders, *Science* **2005**, *308*, 667-669.
[19] F. B. L. Cougnon, N. A. Jenkins, G. D. Pantoş, J. K. M. Sanders, *Angew. Chem. Int. Ed.* **2012**, *51*, 1443-1447.
[20] F. B. L. Cougnon, H. Y. Au-Yeung, G. D. Pantoş, J. K. M. Sanders, *J. Am. Chem. Soc.* **2011**, *133*, 3198-3207.
[21] H. Y. Au-Yeung, G. D. Pantoş, J. K. M. Sanders, *J. Org. Chem.* **2011**, *76*, 1257-1268.
[22] N. Ponnuswamy, F. B. L. Cougnon, G. D. Pantoş, J. K. M. Sanders, *J. Am. Chem. Soc.* **2014**, *136*, 8243-8251.
[23] N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş, J. K. M. Sanders, *Science* **2012**, *338*, 783-785.
[24] M. Malakoutikhah, J. J. P. Peyralans, M. Colomb-Delsuc, H. Fanlo-Virgós, M. C. A. Stuart, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 18406-18417.
[25] S. Otto, *Acc. Chem. Res.* **2012**, *45*, 2200-2210.
[26] J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans, S. Otto, *Science* **2010**, *327*, 1502-1506.
[27] Y. Ura, J. M. Beierle, L. J. Leman, L. E. Orgel, M. R. Ghadiri, *Science* **2009**, *325*, 73-77.
[28] A. Chaudhary, D. U. Nagaich, N. Gulati, V. K. Sharma, R. L. Khosa, *JAPER* **2012**, *2*, 32-67.
[29] J. T. Snow, J. W. Finley, M. Friedman, *Biochem. Biophys. Res. Commun.* **1975**, *64*, 441-447.
[30] J. P. Tam, C. R. Wu, W. Liu, J. W. Zhang, *J. Am. Chem. Soc.* **1991**, *113*, 6657-6662.
[31] P. A. Fernandes, M. J. Ramos, *Chem. Eur. J.* **2004**, *10*, 257-266.
[32] R. Singh, G. M. Whitesides, *J. Am. Chem. Soc.* **1990**, *112*, 6304-6309.
[33] R. Singh, G. M. Whitesides, *J. Am. Chem. Soc.* **1990**, *112*, 1190-1197.
[34] A. V. Gromova, J. M. Ciszewski, B. L. Miller, *Chem. Commun.* **2012**, *48*, 2131-2133.
[35] J. Li, P. Nowak, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 9222-9239.
[36] F. B. L. Cougnon, J. K. M. Sanders, *Acc. Chem. Res.* **2012**, *45*, 2211-2221.
[37] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652-3711.
[38] E. Moulin, G. Cormos, N. Giuseppone, *Chem. Soc. Rev.* **2012**, *41*, 1031-1049.
[39] S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders, J. F. Stoddart, *Angew. Chem. Int. Ed.* **2002**, *41*, 898-952.
[40] J. M. Lehn, *Angew. Chem. Int. Ed.* **2013**, *52*, 2836-2850.
[41] G. M. Whitesides, B. Grzybowski, *Science* **2002**, *295*, 2418-2421.
[42] M. E. Belowich, C. Valente, R. A. Smaldone, D. C. Friedman, J. Thiel, L. Cronin, J. F. Stoddart, *J. Am. Chem. Soc.* **2012**, *134*, 5243-5261.

[43] M. E. Belowich, C. Valente, J. F. Stoddart, *Angew. Chem. Int. Ed.* **2010**, *49*, 7208-7212.
[44] J.-F. Ayme, J. E. Beves, D. A. Leigh, R. T. McBurney, K. Rissanen, D. Schultz, *Nat. Chem.* **2011**, *4*, 15.
[45] G. Du, E. Moulin, N. Jouault, E. Buhler, N. Giuseppone, *Angew. Chem. Int. Ed.* **2012**, *51*, 12504-12508.
[46] O. A. Bozdemir, G. Barin, M. E. Belowich, A. N. Basuray, F. Beuerle, J. F. Stoddart, *Chem. Commun.* **2012**, *48*, 10401-10403.
[47] M. Rauschenberg, S. Bomke, U. Karst, B. J. Ravoo, *Angew. Chem. Int. Ed.* **2010**, *49*, 7340-7345.
[48] M. Eisenberg, I. Shumacher, R. Cohen-Luria, G. Ashkenasy, *Bioorg. Med. Chem.* **2013**, *21*, 3450-3457.
[49] K. R. West, K. D. Bake, S. Otto, *Org. Lett.* **2005**, *7*, 2615-2618.
[50] A. R. Stefankiewicz, J. K. M. Sanders, *Chem. Commun.* **2013**, *49*, 5820-5822.
[51] D. Wojciech, B. Camille, K. Clément, R. Sébastien, B. Mihail, C. Sébastien, S. A. R., U. Sébastien, *Chem. Eur. J.* **2017**, *23*, 18010-18018.
[52] E. Liang, H. Zhou, X. Ding, Z. Zheng, Y. Peng, *Chem. Commun.* **2013**, *49*, 5384-5386.
[53] N. Giuseppone, J. M. Lehn, *Chem. Eur. J.* **2006**, *12*, 1715-1722.
[54] N. Giuseppone, G. Fuks, J. M. Lehn, *Chem. Eur. J.* **2006**, *12*, 1723-1735.
[55] N. Giuseppone, J.-M. Lehn, *J. Am. Chem. Soc.* **2004**, *126*, 11448-11449.
[56] N. Hafezi, J.-M. Lehn, *J. Am. Chem. Soc.* **2012**, *134*, 12861-12868.
[57] H. Y. Au-Yeung, G. D. Pantoş, J. K. M. Sanders, *Angew. Chem. Int. Ed.* **2010**, *49*, 5331-5334.
[58] K. B. Jadhav, R. J. Lichtenecker, A. Bullach, B. Mandal, H. D. Arndt, *Chem. Eur. J.* **2015**, *21*, 5898-5908.
[59] F. B. L. Cougnon, N. Ponnuswamy, N. A. Jenkins, G. D. Pantoş, J. K. M. Sanders, *J. Am. Chem. Soc.* **2012**, *134*, 19129-19135.
[60] H. Y. Au-Yeung, F. B. L. Cougnon, S. Otto, G. D. Pantos, J. K. M. Sanders, *Chem. Sci.* **2010**, *1*, 567-574.
[61] H. Y. Au-Yeung, G. D. Pantoş, J. K. M. Sanders, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10466-10470.
[62] H. Y. Au-Yeung, G. Dan Pantoş, J. K. M. Sanders, *J. Am. Chem. Soc.* **2009**, *131*, 16030-16032.
[63] P. Kovaříček, J.-M. Lehn, *J. Am. Chem. Soc.* **2012**, *134*, 9446-9455.
[64] P. Reutenauer, P. J. Boul, J. M. Lehn, *Eur. J. Org. Chem.* **2009**, *2009*, 1691-1697.
[65] G. Vantomme, S. Jiang, J.-M. Lehn, *J. Am. Chem. Soc.* **2014**, *136*, 9509-9518.
[66] G. Vantomme, J. M. Lehn, *Angew. Chem. Int. Ed.* **2013**, *52*, 3940-3943.
[67] M. N. Chaur, D. Collado, J. M. Lehn, *Chem. Eur. J.* **2011**, *17*, 248-258.
[68] Y. Maeda, N. Javid, K. Duncan, L. Birchall, K. F. Gibson, D. Cannon, Y. Kanetsuki, C. Knapp, T. Tuttle, R. V. Ulijn, H. Matsui, *J. Am. Chem. Soc.* **2014**, *136*, 15893-15896.
[69] S. K. M. Nalluri, R. V. Ulijn, *Chem. Sci.* **2013**, *4*, 3699-3705.
[70] E. Buhler, N. Sreenivasachary, S.-J. Candau, J.-M. Lehn, *J. Am. Chem. Soc.* **2007**, *129*, 10058-10059.
[71] A. Herrmann, N. Giuseppone, J. M. Lehn, *Chem. Eur. J.* **2009**, *15*, 117-124.
[72] N. Giuseppone, J. M. Lehn, *Angew. Chem. Int. Ed.* **2006**, *45*, 4619-4624.
[73] R. Zangi, M. Hagen, B. J. Berne, *J. Am. Chem. Soc.* **2007**, *129*, 4678-4686.
[74] R. L. E. Furlan, S. Otto, J. K. M. Sanders, *Proc. Natl. Acad. Sci.* **2002**, *99*, 4801-4804.
[75] C. S. Mahon, D. A. Fulton, *Chem. Sci.* **2013**, *4*, 3661-3666.
[76] S. Hamieh, V. Saggiomo, P. Nowak, E. Mattia, R. F. Ludlow, S. Otto, *Angew. Chem. Int. Ed.* **2013**, *52*, 12368-12372.
[77] L. Ratjen, G. Vantomme, J. M. Lehn, *Chem. Eur. J.* **2015**, *21*, 10070-10081.
[78] R.-C. Brachvogel, F. Hampel, M. von Delius, *Nat. Commun.* **2015**, *6*, 7129.
[79] D. J. Hansen, I. Manuguerra, M. B. Kjelstrup, K. V. Gothelf, *Angew. Chem. Int. Ed.* **2014**, *53*, 14415-14418.
[80] M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, *44*, 2455-2488.
[81] Z. Grote, R. Scopelliti, K. Severin, *Angew. Chem. Int. Ed.* **2003**, *42*, 3821-3825.
[82] K. Severin, *Chem. Eur. J.* **2004**, *10*, 2565-2580.
[83] I. Saur, K. Severin, *Chem. Commun.* **2005**, 1471-1473.
[84] P. T. Corbett, S. Otto, J. K. M. Sanders, *Chem. Eur. J.* **2004**, *10*, 3139-3143.
[85] P. T. Corbett, J. K. M. Sanders, S. Otto, *J. Am. Chem. Soc.* **2005**, *127*, 9390-9392.
[86] P. T. Corbett, J. K. M. Sanders, S. Otto, *Chem. Eur. J.* **2008**, *14*, 2153-2166.

[87] J. W. Sadownik, D. Philp, *Angew. Chem. Int. Ed.* **2008**, *47*, 9965-9970.
[88] B. Rubinov, N. Wagner, H. Rapaport, G. Ashkenasy, *Angew. Chem. Int. Ed.* **2009**, *48*, 6683-6686.
[89] Z. Dadon, M. Samiappan, A. Shahar, R. Zarivach, G. Ashkenasy, *Angew. Chem. Int. Ed.* **2013**, *52*, 9944-9947.
[90] M. Colomb-Delsuc, E. Mattia, J. W. Sadownik, S. Otto, *Nat. Commun.* **2015**, *6*, 7427.
[91] P. Nowak, M. Colomb-Delsuc, S. Otto, J. Li, *J. Am. Chem. Soc.* **2015**, *137*, 10965-10969.
[92] R. Nguyen, L. Allouche, E. Buhler, N. Giuseppone, *Angew. Chem. Int. Ed.* **2009**, *48*, 1093-1096.
[93] J. Li, J. M. A. Carnall, M. C. A. Stuart, S. Otto, *Angew. Chem. Int. Ed.* **2011**, *50*, 8384-8386.
[94] M. M. Safont-Sempere, G. Fernández, F. Würthner, *Chem. Rev.* **2011**, *111*, 5784-5814.
[95] S. W. Sisco, J. S. Moore, *Chem. Sci.* **2014**, *5*, 81-85.
[96] G. Leonetti, S. Otto, *J. Am. Chem. Soc.* **2015**, *137*, 2067-2072.
[97] F. Biedermann, H.-J. Schneider, *Chem. Rev.* **2016**, *116*, 5216-5300.

2.2. Specific objectives and hypothesis

The main objective of the present Chapter is to explore the effect of geometry and structurally diverse on the composition of dynamic libraries through the combination of building blocks with different valence. Furthermore, the study of the adaptive properties of these DCLs in response to different environmental stimuli and structural effects arising from the BBs. With this purpose, we hypothesized that the understanding of the behavior of dynamic libraries in response to environmental and structural effects could decode the key to direct the thermodynamic equilibrium towards a desired formation of the products.

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

1) To synthetize and characterize structurally diverse pseudopeptidic BBs capable of generate dynamic combinatorial libraries in aqueous media.
2) To design and prepare the corresponding dynamic combinatorial libraries with the previously synthetized BBs. To study a wide range of different dynamic combinatorial libraries with different valence in aqueous medium.
3) To explore the adaptive behavior of DCLs to different external stimuli (pH, ionic strength, polarity of the solvent, concentration of the constituent blocks, etc ...).
4) To study the structural effects (modifying one or more components of the initial library) in the composition and proportion of final products of DCLs.

2.3. PUBLICATION A:

Constitutional self-selection from dynamic combinatorial libraries in aqueous solution through supramolecular interactions.

Chem. Commun. **2014**, *50*, 4564-4566

Publiation Date (online): 5 March 2014

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Constitutional self-selection from dynamic combinatorial libraries in aqueous solution through supramolecular interactions†

Cite this: *Chem. Commun.*, 2014, 50, 4564

Received 11th January 2014,
Accepted 27th February 2014

DOI: 10.1039/c4cc00245h

www.rsc.org/chemcomm

Jordi Solà,* Maria Lafuente, Joan Atcher and Ignacio Alfonso*

We describe the predominant formation of a specific constitution arising from the combination of building blocks with different topologies through disulphide chemistry in a Dynamic Combinatorial Library (DCL). The supramolecular interactions established by a zwitterionic cysteine moiety are responsible for the self-selection of one product from all the virtual members of a large library.

Self-assembling and self-organization are inherent processes in Nature that occur by the spontaneous folding and assembling of relatively simple chemical entities.[1] For that, the delicate inter- and intramolecular interactions are established and regulated by proof-reading and self-correction.[2] Accordingly, the chemical information stored in a given chemical structure is decoded to render beautifully complex and fully functional biomolecular systems. For decades, supramolecular chemists have been fascinated and inspired by this biomolecular machinery. More recently, Constitutional Dynamic Chemistry (CDC) and Dynamic Combinatorial Chemistry (DCC) have proposed the implementation of complex dynamic systems formed by members able to exchange under thermodynamic control.[3] These dynamic systems (Dynamic Combinatorial Libraries, DCLs) are excellent benchmark models for decoding the chemical information stored in the combination of simple building blocks towards the emergence of new assembled structures.[4]

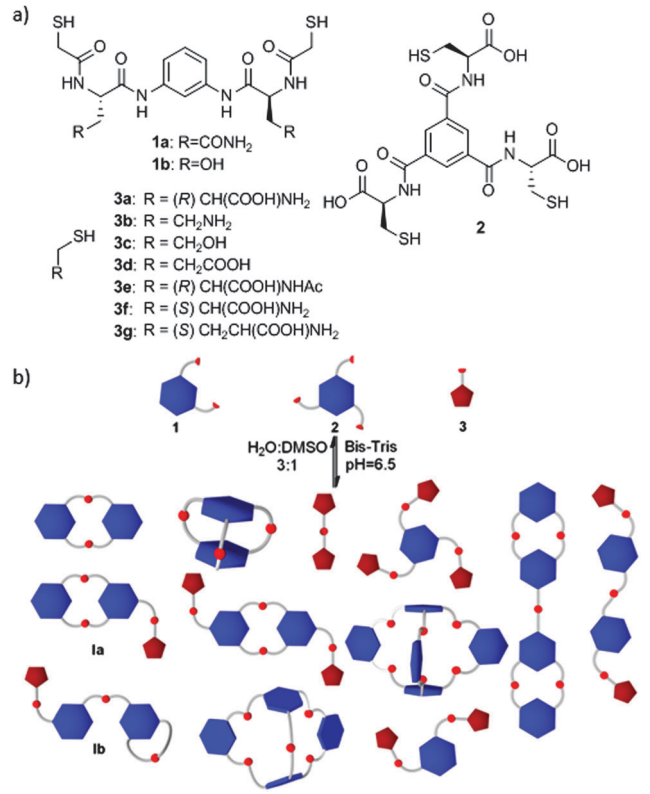
Disulphide chemistry has been exploited widely for the generation of DCLs in aqueous media.[5] However, most of the systems described are based on molecules with two reacting sites (bipodal). Thus, although some impressive (and somehow unexpected) geometries have been described such as catenanes and knots,[6] the tridimensional space covered is relatively small, and the libraries are mostly directed towards the generation and interconversion of cyclic oligomers.[7] The formation of cage-like organic structures has seldom been described by oxidation of tripodal building blocks or by the combination of tripodal and bipodal building blocks.[8] The stoichiometry and the presence of a template were critical for the latter. More recently Sanders has described the diverse topologies obtained by a combination of tri- and monothiols in water.[9] In order to expand the structural and topological diversity of supramolecular assemblies by DCC we decided to study the effect of combining mono-, di- and trisulphides on the same DCL. In principle, by combining a series of building blocks of different topology, a broad structural variety is expected. We chose to combine pseudopeptidic C_2-symmetric dithiols **1a–b** described earlier by our group[10] with the tripodal building block **2**[8] and a series of monothiols **3a–g**. Thus, one may expect the formation of several compounds with different architectures some of which are schematically represented in Scheme 1b.

We started by mixing trithiol **2** with neutral dithiols **1a** or **1b** at pH 6.5 in an aqueous solution containing 25% DMSO. The use of DMSO allows faster oxidation and exchange within a few hours even at a slightly acidic pH.[11] The mixtures were analysed after 24 h and 48 h with no significant changes. Several compounds can be detected in accordance with the results published in the literature.[8–10] Our DCLs, however, presented less variety of cyclic products arising from oxidation of dithiols **1a–1b** (Fig. 1a, blue trace) as they prefer to form mixed products with **2** (presumably cage-like compounds of formula 1_22_2 and 1_32_2).

To our surprise, when the same experiment was repeated for **1a** and **2** in the presence of 2.5 mM of cysteine (enough to saturate all free thiols) a product containing each of the three building blocks (**1a23a**) was predominantly formed (Fig. 1a, red trace). The other minor detectable species are a combination of **2** with 3 cysteines and a structure of the formula **1a$_2$23a**. Since Ser-derivative **1b** presented similar behaviour (Fig. 1b) we focused our studies on Asn-derivative **1a**. The distribution of products was mostly unchanged upon varying the concentration of Cys from 0.5 mM (equimolar with other components) up to 100 mM, showing the stability of the product formed.[12] In addition, we combined building blocks **1a** and **2** in different proportions (1 : 2 and 2 : 1) in the presence of 1 eq. of Cys per thiol. In these experiments almost all the limiting reagent was consumed in the same product of formula **1a23a** whereas the reagent in excess formed either adducts with Cys or the corresponding

Department of Biological Chemistry and Molecular Modelling, IQAC-CSIC, Jordi Girona 18-26, Barcelona, Spain. E-mail: jordi.sola@iqac.csic.es, ignacio.alfonso@iqac.csic.es; Fax: +34 932 045 904; Tel: +34 934 006 100

† Electronic supplementary information (ESI): Experimental details, synthetic procedures, additional HPLC traces and NMR spectra are available. See DOI: 10.1039/c4cc00245h



Scheme 1 (a) Thiol building blocks used in this study. (b) Schematic representation of some of the possible architectures generated by the combination of topologically different building blocks.

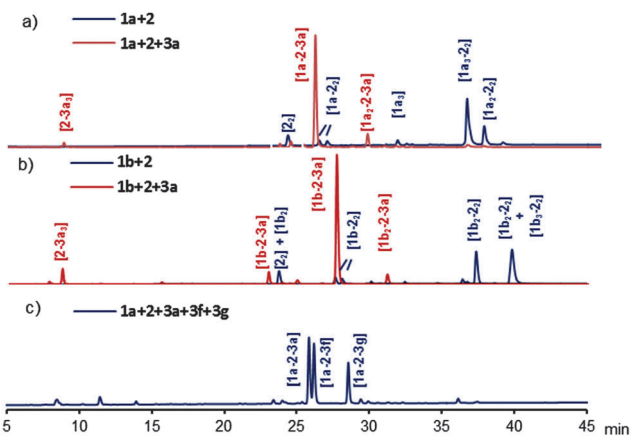


Fig. 1 (a and b) DLCs formed by **1a**,**b** and **2** (0.5 mM each) in aqueous solution (25% DMSO, pH 6.5) in the presence (red) or absence (blue) of cysteine (2.5 mM). (c) DLC formed by **1a** (0.5 mM) and **2** (0.5 mM) in the presence of L-Cys, D-Cys and homo-Cys (1.0 mM each).

homodimer (ESI†). We attribute the selection of the major product formed in the mixture to some sort of self-recognition within the building blocks during the assembly process. Thus, stabilization through supramolecular interactions would lead to the formation of a major, more stable compound. Reversibility tests confirmed the reaching of the thermodynamic equilibrium: addition of **3a** after 24 h to a sample containing **1a** and **2** gave essentially the same mixture as the one formed by **1a**+**2**+**3a** from the beginning (ESI†).

In order to ascertain what structural elements were crucial for the selection, other monothiols (**3b–e**) were tested in the libraries. All of the thiols performed notably worse than Cys in selecting the corresponding major product. These experiments showed that both the ammonium and the carboxylate groups are needed in order to successfully select a particular structure, with the ammonium group being more critical, as a hydrogen bond donor in that position clearly enhances recognition. In fact, the experiments done with thiols **3a–e** showed the trend NH_3^+ > OH > NHAc (ESI†) to effectively select a predominant constitution from the mixture.

We turned then to investigate the structure of the multi-component molecule to see what possible interactions would lead to the formation of the major product. In principle two architectures are possible for a molecule of the formula **1a23a**: the symmetrical structure **Ia** and the less symmetrical **Ib** (Scheme 1). We considered **Ia** more likely as it would present a more compact structure in solution.

To confirm this possible arrangement we reproduced the synthesis of **1a23a** on a preparative scale. The NMR spectra of the purified product in buffered water at pH 6.5 with 15% DMSO-d_6 were acquired using water signal suppression by excitation sculpting (Fig. 2a). All the signals were successfully assigned by a combination of ^{1}H-NMR and 2D-TOCSY experiments. All expected protons are visible in the ^{1}H-NMR spectra apart from the three protons corresponding to the alpha carbons of the amino acids from bipodal and tripodal fragments, which fall in the region of water suppression. However, these protons are visible in the TOCSY experiment. Three doublets corresponding to the three amide bonds coupled to the chiral centre of the amino acids are clearly visible in a 2 : 2 : 1 ratio (H_A, H_B, H_E). In addition, protons H_C and H_{18} from the phenylenediamine fragment, and H_4 and H_2 from the trimesic acid appear as singlets. Overall, this data is consistent only with the formation of the more symmetrical product **Ia**.

Some other structural considerations can be made based on the ^{1}H-NMR spectra. H_E displays a relatively low chemical shift compared to H_A, which could be indicative of a twisting out of the aromatic ring plane. This also explains the large chemical shift difference between H_2 and H_4, both from the tripodal aromatic ring. On the other hand, H_C and H_B appear in the downfield amide region, which is consistent with some kind of hydrogen bonding phenomena, most likely with any of the carboxylate anions present in the molecule. To shine some light on the possible tri-dimensional structure of the compound in solution we performed a conformational search using MMFFaq force field. Molecular modelling rendered a folded structure with several stabilizing interactions as shown in Fig. 2b.

The most stable conformer shows a structure where the pendant Cys folds over the macrocyclic cavity. The ammonium cation resides in the centre of the macrocycle and is stabilized by the three carboxylates that arise from the trimesic acid moiety. The carboxylate group of the cysteine residue would as well be stabilized by hydrogen bonding interactions with the macrocycle amide protons (H_C/H_B). Other hydrogen bonding interactions between the macrocyclic peptidic amides and the carboxylates would further stabilise the structure. Interestingly a *cis* amide bond is observed for the tripodal amide that is not part of the macrocycle. We can also observe different dispositions (in/out) for the carbonyls of the

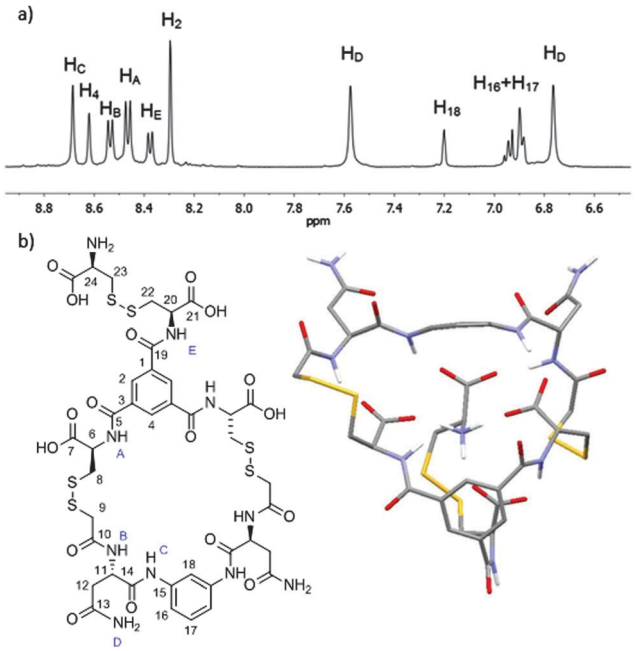


Fig. 2 (a) Partial ^{1}H-NMR spectrum (500 MHz, buffered H_20 with 15% DMSO-d_6, 298 K) showing the aromatic and amide region for molecule **1a23a**. (b) Proposed structure for the major product **1a23a** (**Ia**) with atom labelling (left) and structure obtained by molecular modelling (right).

amides both in the bipodal and tripodal moieties. This data is in agreement with what we observed by 2D-NOESY experiments. NOE correlations were visible for proton H_A with both H_4 and H_2 which would indicate a dynamic rotation of the amide bonds. Similar behaviour is observed for H_C and H_{18}, H_{16} showing that some degree of conformational freedom is present.

Unluckily, we could not assign unambiguous through-space correlations that supported the proposed folding, since proton H_{23} (from the pendant cysteine) overlaps protons H_9 (within the macrocycle backbone). Besides, as stated before, protons H_6, H_{11} and H_{20} fall in the water suppression region and thus, no NOE correlations could be assigned for these protons either. Nevertheless, further evidence of the folded structure was obtained when the sample was acidified with TFA (10 eq.), because some significant shifts of the NMR signals were observed (ESI†). Thus, H_E and H_A moved downfield and to a similar chemical shift while H_B and H_4 moved upfield. This is consistent with an unfolding of the structure upon protonation, which would leave all the amide or aromatic protons of the tripodal moiety in a more similar chemical environment.

Finally, to prove that small structural changes had a decisive effect on the product distribution we tested the unnatural enantiomer of cysteine (D-Cys, **3f**) and structurally related homocysteine (**3g**). D-Cys formed a major product containing all the building blocks with no major differences with L-Cys. Homocysteine, however, formed a major product **1a23g**, but other compounds were also readily detected. Strikingly, in a competition experiment L-Cys was slightly preferred over D-Cys (51 : 49, ESI†) and in a competition experiment between L-, D-, and homocysteine the tendency showed that the preference is L > D > homo (39 : 37 : 24) demonstrating that subtle structural differences have an impact on the formation of the major species (see Fig. 1c). This provides further evidence of a folded structure in solution as homocysteine would have a looser fit.

We show the importance of the synergic action of subtle recognition events within the components of a DCL. A combination of building blocks of different topology could result in a plethora of possible structures. However, cooperative supramolecular interactions between the building blocks result in the self-recognition and the selective amplification of a singular constitution even when it is statistically disfavoured. The structural importance of these interactions is proved by the difference observed in the chromatographic profile of the DCL when delicate changes are introduced. We expect that our findings will improve the comprehension of the behaviour of complex dynamic libraries of compounds.

This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO, CTQ2012-38543-C03-03) and EU (FP7-PEOPLE-2012-CIG-321659). Personal financial support for J.S. (MINECO, Ramón y Cajal contract), M.L. (MINECO, FPI fellowship) and J.A. (CSIC and European Social Fund, JAE-predoc fellowship) are gratefully acknowledged. We thank Dr Yolanda Pérez for helpful assistance with the NMR experiments.

Notes and references

1 G. M. Whitesides and B. Grzybowski, *Science*, 2002, **295**, 2418–2421; J.-M. Lehn, *Angew. Chem., Int. Ed.*, 2013, **52**, 2836–2850.
2 J.-M. Lehn, *Chem. Soc. Rev.*, 2007, **36**, 151–160; J.-M. Lehn, *Angew. Chem., Int. Ed. Engl.*, 1990, **29**, 1304–1319.
3 R. F. Ludlow and S. Otto, *Chem. Soc. Rev.*, 2008, **37**, 101–108; J. Li, P. Nowak and S. Otto, *J. Am. Chem. Soc.*, 2013, **135**, 9222–9239; J. R. Nitschke, *Nature*, 2009, **462**, 736–738; G. M. Whitesides and R. F. Ismagilov, *Science*, 1999, **284**, 89–92; F. B. L. Cougnon and J. K. M. Sanders, *Acc. Chem. Res.*, 2011, **45**, 2211–2221; P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652–3711.
4 S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders and J. F. Stoddart, *Angew. Chem., Int. Ed.*, 2002, **41**, 899–952; S. Otto, *Curr. Opin. Drug Discovery Dev.*, 2003, **6**, 509–520; I. Huc and J.-M. Lehn, *Actual. Chim.*, 2000, 51–54; R. Custelcean, *Top. Curr. Chem.*, **322**, 193–216; E. Moulin, G. Cormos and N. Giuseppone, *Chem. Soc. Rev.*, 2012, **41**, 1031–1049; N. Giuseppone, *Acc. Chem. Res.*, 2012, **45**, 2178–2188; S. Otto, *Acc. Chem. Res.*, 2012, **45**, 2200–2210.
5 S. P. Black, J. K. M. Sanders and A. R. Stefankiewicz, *Chem. Soc. Rev.*, 2014, **43**, 1861–1872.
6 M.-K. Chung, S. J. Lee, M. L. Waters and M. R. Gagne, *J. Am. Chem. Soc.*, 2012, **134**, 11430–11443; F. B. L. Cougnon, N. Ponnuswamy, N. A. Jenkins, G. D. Pantos and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2012, **134**, 19129–19135; N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş and J. K. M. Sanders, *Science*, 2012, **338**, 783–785.
7 M. Rauschenberg, S. Bomke, U. Karst and B. J. Ravoo, *Angew. Chem., Int. Ed.*, 2010, **49**, 7340–7345; M. Eisenberg, I. Shumacher, R. Cohen-Luria and G. Ashkenasy, *Bioorg. Med. Chem.*, 2013, **21**, 3450–3457.
8 K. R. West, K. D. Bake and S. Otto, *Org. Lett.*, 2005, **7**, 2615–2618; A. R. Stefankiewicz, M. R. Sambrook and J. K. M. Sanders, *Chem. Sci.*, 2012, **3**, 2326–2329.
9 A. R. Stefankiewicz and J. K. M. Sanders, *Chem. Commun.*, 2013, **49**, 5820–5822.
10 J. Atcher, A. Moure and I. Alfonso, *Chem. Commun.*, 2013, **49**, 487–489.
11 J. Atcher and I. Alfonso, *RSC Adv.*, 2013, **3**, 25605–25608.
12 For concentrations of cysteine higher than 10 mM a precipitate appears. However HPLC analysis of the mixture has little change in absorbance showing that the major part of the precipitate is cystine, which is invisible at 254 nm.

2.3.1. Electronic Supplementary Information (ESI) for:

PUBLICATION A

Constitutional self-selection from dynamic combinatorial libraries in aqueous solution through supramolecular interactions

Jordi Solà, Maria Lafuente, Joan Atcher and Ignacio Alfonso**

Department of Biological Chemistry and Molecular Modelling.
Institute of Advanced Chemistry of Catalonia, IQAC-CSIC
Jordi Girona 18-26, 08034, Barcelona, Spain. Fax: (+34)932045904
E-mail: ignacio.alfonso@iqac.csic.es
jordi.sola@iqac.csic.es

Table of Contents:

GENERAL CHARACTERISTICS

General: Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka or Merck) and were used without further purification. Flash chromatographic purifications and preparative reversed-phase purifications were performed on a Biotage® Isolera Prime™ equipment. TLCs were performed using 6x3 cm SiO_2 pre-coated aluminium plates (ALUGRAM® SIL G/UV_{254}).

Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC) analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using:

i) For the characterization of [**1a,b**]: a reversed-phase X-Terra C_{18} (15 x 0.46 cm, 5 µm) column. (CH_3CN + 0.07% TFA and H_2O + 0.1% TFA) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelengths were set at 220 and 254 nm.

ii) For the analysis of the DCLs: a reversed-phase kromaphase C_{18} (25 x 0.46 cm, 5µm) column. (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelength was set at 254 nm.

Nuclear Magnetic Resonance (NMR) spectroscopic experiments were carried out on a Varian INOVA 500 spectrometer (500 MHz for ^1H and 125 MHz for ^{13}C) and a Varian Mercury 400 instrument (400 MHz for ^1H and 101 MHz for ^{13}C). The chemical shifts are reported in ppm relative to trimethylsilane (TMS), and coupling constants (J) are reported in Hertz (Hz).

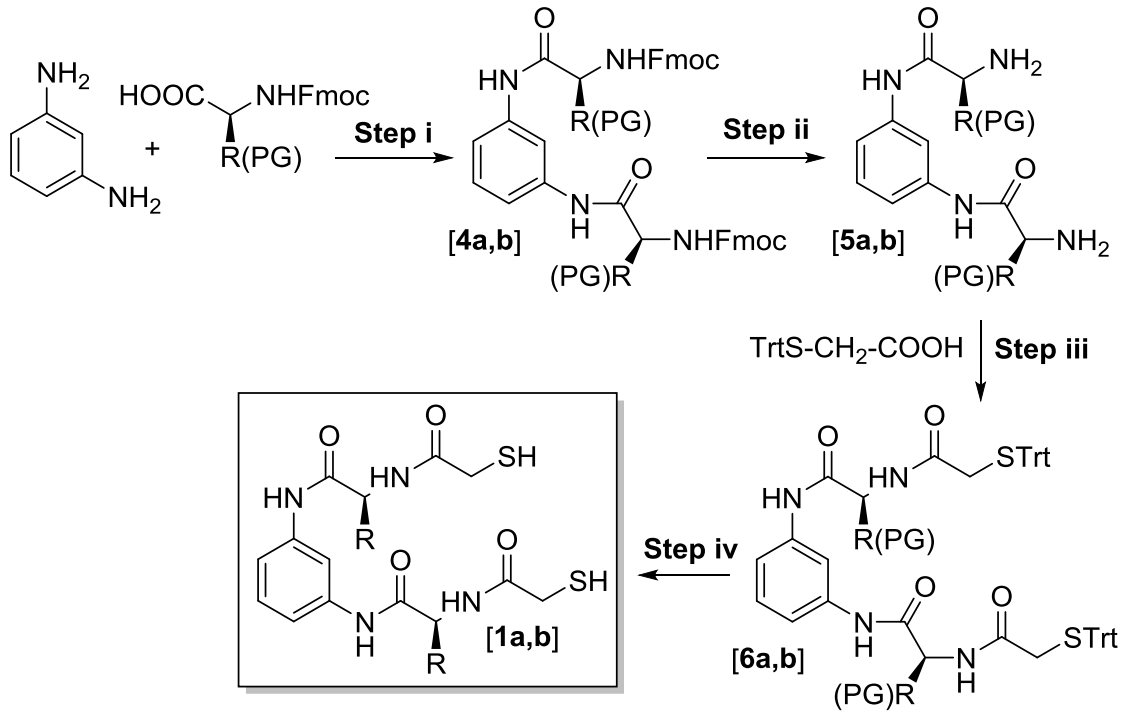
pH measurements were performed at room temperature on a Crison GLP21 pH-meter with the electrodes Crison 50 14T (≥10 mL samples) and PHR-146 Micro (<10 mL samples).

High Resolution Mass Spectrometry (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.

SYNTHESIS OF THE BUILDING BLOCKS

The tritylsulfanyl acetic acid was prepared as previously described.[1] Also the compound [**2**] was synthesized as previously reported.[2] The compounds [**3a-g**] were purchased from commercial suppliers (Sigma-Aldrich, Iris Biotech).

Synthetic scheme of [**1a,b**]

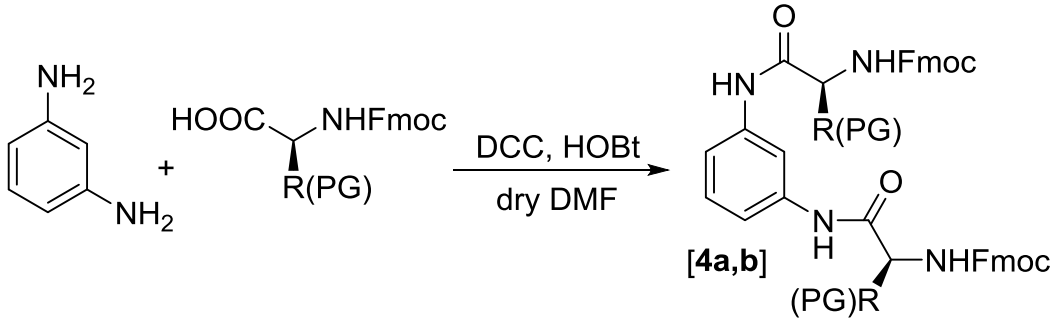


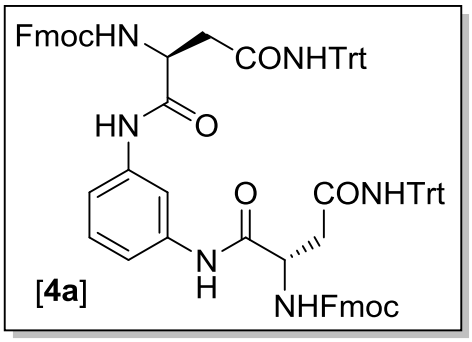
Related references:

1. A. P. Kozikowski, Y. Chen, A. Gaysin, B. Chen, M. A. D'Annibale, C. M. Suto and B. C. Langley, *J. Med. Chem.*, **2007**, *50*, 3054-3061.

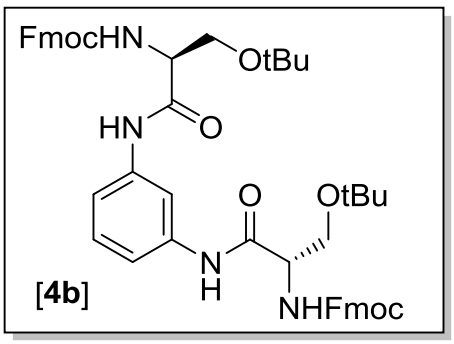
2. K. R. West, K. D. Bake and S. Otto, *Org. Lett.*, **2005**, *7*, 2615-2618.

Step i: Experimental procedure for the synthesis of [**4a,b**]





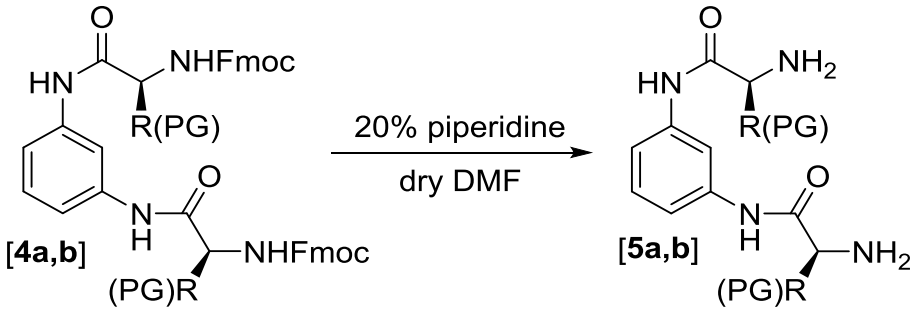
[**4a**]: Fmoc-Asn(Trt)-OH (4.24 g, 7.11 mmol) was dissolved in dry DMF (16 mL) and both dicyclohexylcarbodiimide (DCC, 2.231 g, 10.81 mmol) and 1-hydroxybenzotriazole (HOBt, 1.252 g, 9.27 mmol) were added over the solution. The reaction mixture was cooled to 0ºC. A solution of *m*-phenylenediamine (334 mg, 3.09 mmol) in dry DMF (10 mL) was added over the mixture through a cannula. The solution 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 hexane: AcOEt as eluent (from 30% to 50% AcOEt) to give 2.047 g of [**4a**] (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 (brs, 2H, NH), 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, NH), 6.54 (brs, 2H, NH), 4.68 (brs, 2H, C*H), 4.51 – 4.32 (m, 4H, CH_2), 4.20 (t, *J* = 7.0 Hz, 2H, CH), 3.16 (d, *J* = 15.7 Hz, 2H, C$\underline{H}_2$C*H), 2.66 (dd, *J* = 15.7, 6.9 Hz, 2H, C$\underline{H}_2$C*H) ppm. ^{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 (2 x C_{Ar}), 143.8 (2 x C_{Ar}), 141.4 (2 x C_{Ar}), 141.4 (2 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 CH_2), 52.2 (2 x C*H), 47.2 (2 x CH), 38.9 (2 x $\underline{C}H_2$C*H) ppm.

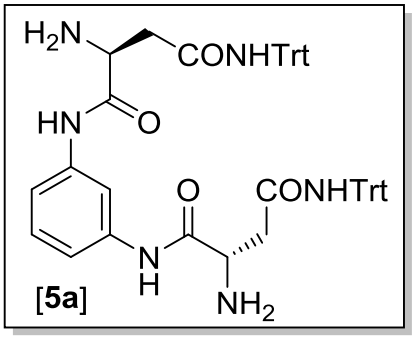


[**4b**]: this compound was obtained as described above starting from the Fmoc-Ser(*t*Bu)-OH. The residue was purified by flash chromatography using hexane: AcOEt as eluent (from 25% to 40% AcOEt, Rf AcOEt/Hexane, 3:2 (v:v): 0.83) to give 1.05 g of [**4b**] (43% yield) as a white solid. HRMS (ESI+) calcd. 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 (brs, 2H, NH), 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 (brs, 2H, NH), 4.44 (d, *J* = 7.0 Hz, 4H, CH_2), 4.35 (brs,

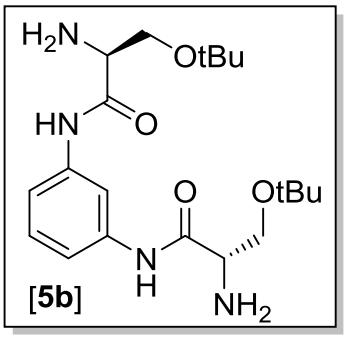
2H, C<u>H</u>$_2$C*H), 4.25 (t, *J* = 6.9 Hz, 2H, CH), 3.92 (brs, 2H, C<u>H</u>$_2$C*H), 3.45 (t, *J* = 8.7 Hz, 2H: C*H), 1.28 (s, 18H, CH_3) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ = 168.5 (2 x CO), 156.2 (2 x CO), 143.9 (4 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.9 (1 x CH_{Ar}), 127.9 (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 CH_2), 61.9 (2 x <u>C</u>H_2C*H), 54.8 (2 x C*H), 47.3 (2 x CH), 27.6 (6 x CH_3) ppm.

<u>Step ii: Experimental procedure for the synthesis of [**5a,b**]</u>



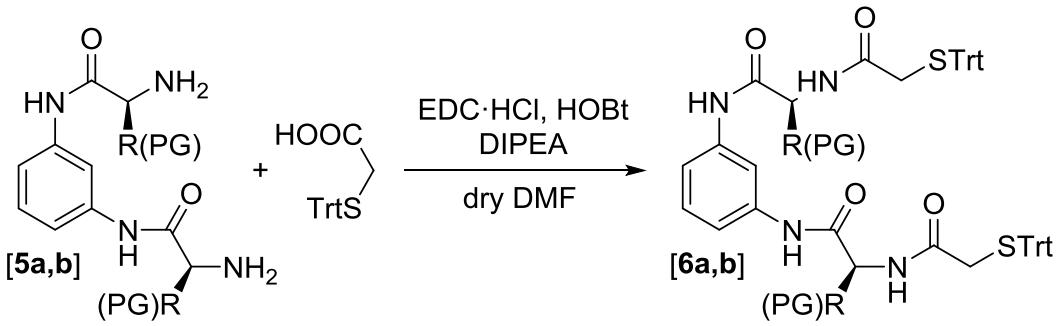


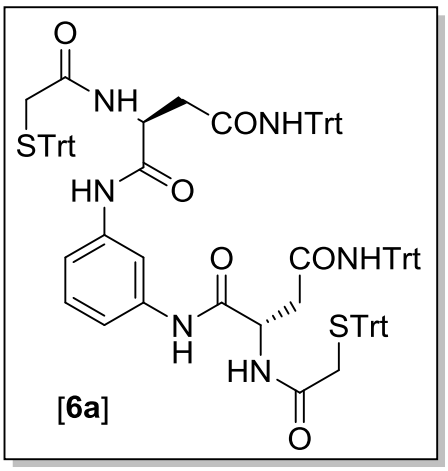
[**5a**]: [**4a**] (600 mg, 0.47 mmol) was dissolved in 4.0 mL of 20% piperidine in dry DMF. After several minutes the product precipitated as a white solid but the mixture was allowed to react at room temperature for 4 hours until complete conversion of starting material. Diethyl ether was added over the reaction mixture and the product was filtered off and washed with diethyl ether. 293 mg of [**5a**] (75% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{52}H_{48}N_6O_4$ $[M+H]^+$ (m/z): 821.3810, found: 821.3712. ^{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) ppm. ^{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) ppm.



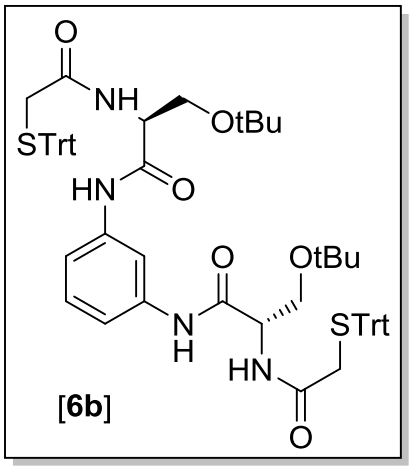
[**5b**]: 522 mg of [**5b**] (99% yield) were obtained as described above starting from the [**4b**]. HRMS (ESI+) calcd. for $C_{20}H_{34}N_4O_4$ $[M+H]^+$ (m/z): 395.2653, found: 395.2672. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.54 (s, 2H, NH), 7.91 (t, *J* = 2.1 Hz, 1H, CH_{Ar}), 7.39–7.35 (m, 2H, CH_{Ar}), 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 (brs, 4H, NH_2), 1.21 (s, 18H, CH_3) ppm. ^{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) ppm.

Step iii: Experimental procedure for the synthesis of [**6a,b**]





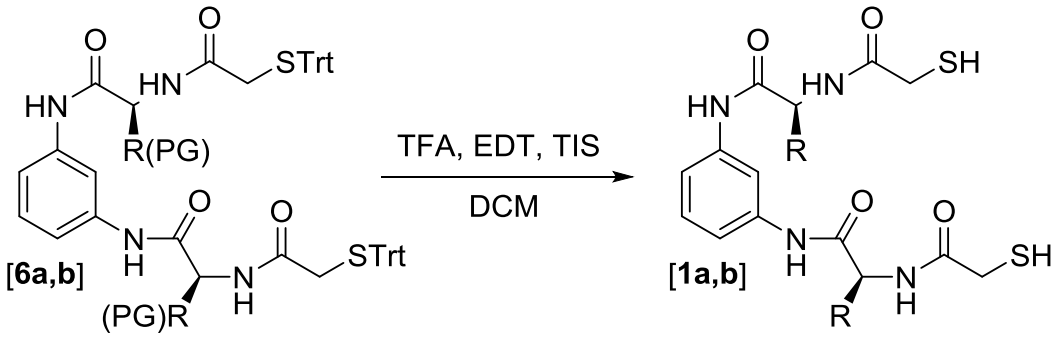
[**6a**]: tritylsulfanyl acetic acid (229 mg, 0.69 mmol) was dissolved in dry DMF (10 mL) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC·HCl, 144 mg, 0.75 mmol), HOBt (101 mg, 0.75 mmol) and *N*,*N*-diisopropylethylamine (DIPEA, 240 µL, 1.38 mmol) were added over the solution. The reaction mixture was cooled to 0ºC and then, [**5a**] (267 mg, 0.33 mmol) was added over the mixture. The solution was allowed to stir 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): 0.46). The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and dried under reduced pressure. The residue was purified by flash chromatography using hexane: AcOEt as eluent (from 40% to 60% AcOEt) to give 343 mg of [**6a**] (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, NH), 7.60 (t, *J* = 2.0 Hz, 1H, CH_{Ar}), 7.54 – 7.00 (m, 65H, 2H x NH + 63H x CH_{Ar}), 6.91 (s, 2H, NH), 4.49 (td, *J* = 7.5, 3.0 Hz, 2H, C*H), 3.05 (ABq, δ_A = 3.02, δ_B = 3.08, *J* = 15.7Hz, 4H, C$\underline{H}_2$STrt), 2.96 – 2.83 (m, 2H, C$\underline{H}_2$C*H), 2.39 (dd, *J* = 15.7, 7.8 Hz, 2H, C$\underline{H}_2$C*H) ppm. ^{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 $\underline{C}H_2$C*H), 36.2 (2 x $\underline{C}H_2$STrt) ppm.

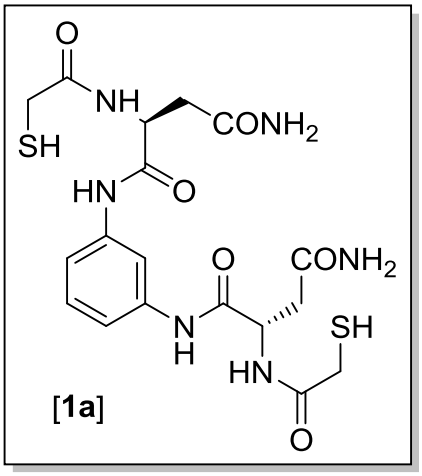


[**6b**]: this compound was obtained as described above starting from [**5b**]. The residue was purified by flash chromatography using hexane: AcOEt as eluent (from 35% to 45% AcOEt, Rf AcOEt/Hexane, 2:3 (v:v): 0.27) to give 612 mg of [**6b**] (51% yield) as a white solid. HRMS (ESI+) calcd. for $C_{62}H_{66}N_4O_6S_2$ $[M+H]^+$ (m/z): 1027.4497, found: 1027.4492. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.68 (s, 2H, NH), 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, NH), 4.24–4.18 (m, 2H, C*H), 3.71 (dd, *J* = 8.6, 4.3 Hz, 2H, C$\underline{H}_2$O^{t}Bu), 3.20–3.08 (m, 6H, 2H x C$\underline{H}_2$O^{t}Bu + 4H x

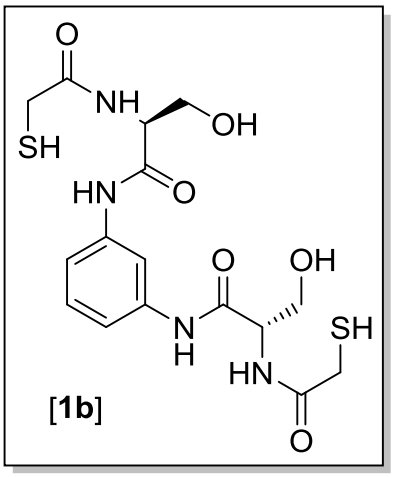
C$\underline{H_2}$STrt), 1.22 (s, 18H, CH_3) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ = 168.7 (2 x CO), 168.2 (2 x CO), 144.1 (6 x C_{Ar}), 138.4 (2 x C_{Ar}), 129.8 (1 x CH_{Ar}), 129.7 (12 x CH_{Ar}), 128.3 (12 x CH_{Ar}), 127.1 (6 x CH_{Ar}), 115.5 (2 x CH_{Ar}), 110.9 (1 x CH_{Ar}), 75.0 (2 x C), 68.0 (2 x C), 61.0 (2 x $\underline{C}H_2$C*H), 53.5 (2 x C*H), 36.2 (2 x CH_2), 27.6 (6 x CH_3) ppm.

Step iv: Experimental procedure for the synthesis of [**1a,b**]





[**1a**]: [**6a**] was dissolved in DCM (1 mL) and 5.5 mL of trifluoroacetic acid (TFA), 313 µL of triisobutylsilane (TIS) and 152 µL of 1,2-ethanedithiol (EDT) were added rapidly and under stirring. The reaction mixture was allowed to stir at room temperature for 40 min, 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. The product was purified using reversed-phase flash chromatography (gradient: from 5% to 30% CH_3CN in H_2O) and 37.8 mg of [**1a**] (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: 485.1271. ^{1}H NMR (400 MHz, DMSO-d_6): δ = 9.97 (s, 2H, N$\underline{H}$COC*H), 8.34 (d, *J* = 7.7 Hz, 2H, C*HN$\underline{H}$CO), 7.93 (t, *J* = 2.0 Hz, 1H, CH_{Ar}), 7.36 (s, 2H, NH_2), 7.29 (dd, *J* = 7.6, 2.0 Hz, 2H, CH_{Ar}), 7.24 – 7.15 (m, 1H, CH_{Ar}), 6.91 (s, 2H, NH_2), 4.67 (app q, *J* = 7.1 Hz, 2H, C*H), 3.17 (d, *J* = 7.9 Hz, 4H, C$\underline{H_2}$SH), 2.73 (t, *J* = 7.9 Hz, 2H, SH), 2.62 – 2.41 (m, 4H, C$\underline{H_2}$C*H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6): δ = 171.1 (2 x $CONH_2$), 169.5 (2 x $\underline{C}$OC*H), 169.4 (2 x $\underline{C}OCH_2$), 139.1 (2 x C_{Ar}), 128.6 (1 x C_{Ar}), 114.6 (2 x C_{Ar}), 110.9 (1 x C_{Ar}), 50.9 (2 x C*H), 37.1 (2 x $\underline{C}H_2$C*H), 27.0 (2 x CH_2SH) ppm.



[**1b**]: 76.4 mg of [**1b**] (90% yield) were obtained as described above starting from [**6b**]. RP-HPLC (gradient: from 5% to 30% CH_3CN in H_2O). HRMS (ESI+) calcd. for $C_{16}H_{22}N_4O_6S_2$ $[M+H]^+$ (m/z): 431.1059, found: 431.1042. ^1H NMR (400 MHz, MeOD-*d*$_4$): δ = 7.92 (t, *J* = 1.8 Hz, 1H, CH_{Ar}), 7.36–7.31 (m, 2H, CH_{Ar}), 7.28–7.23 (m, 1H, CH_{Ar}), 4.56 (t, *J* = 5.3 Hz, 2H, C*H), 3.92–3.82 (m, 4H, C$\underline{H}_2$OH), 3.28 (s, 4H, C$\underline{H}_2$SH) ppm. ^{13}C NMR (101 MHz, MeOD-*d*$_4$): δ = 173.0 (2 x $\underline{C}OCH_2$), 170.2 (2 x $\underline{C}$OC*H), 139.6 (2 x C_{Ar}), 129.9 (1 x CH_{Ar}), 117.3 (2 x CH_{Ar}), 113.4 (1 x CH_{Ar}), 62.8 (2 x CH_2OH), 57.2 (2 x C*H), 27.9 (2 x CH_2SH) ppm.

NMR spectra, HRMS (ESI+) spectra and HPLC traces of [**1a,b**]

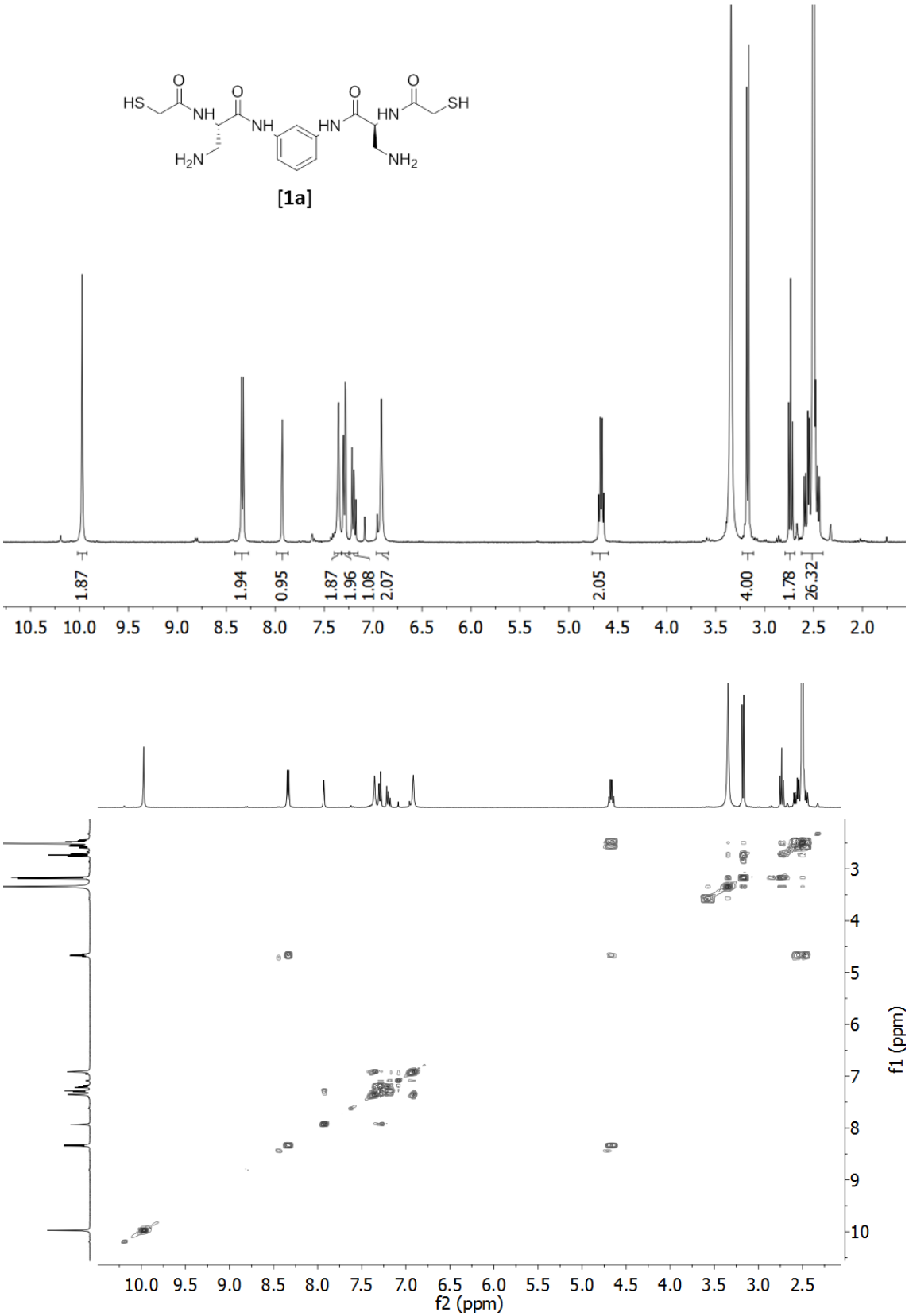


Figure S1: ^{1}H (400 MHz, 298 K in DMSO-d_6) and gCOSY (400 MHz, 298 K in DMSO-d_6) spectra of [**1a**].

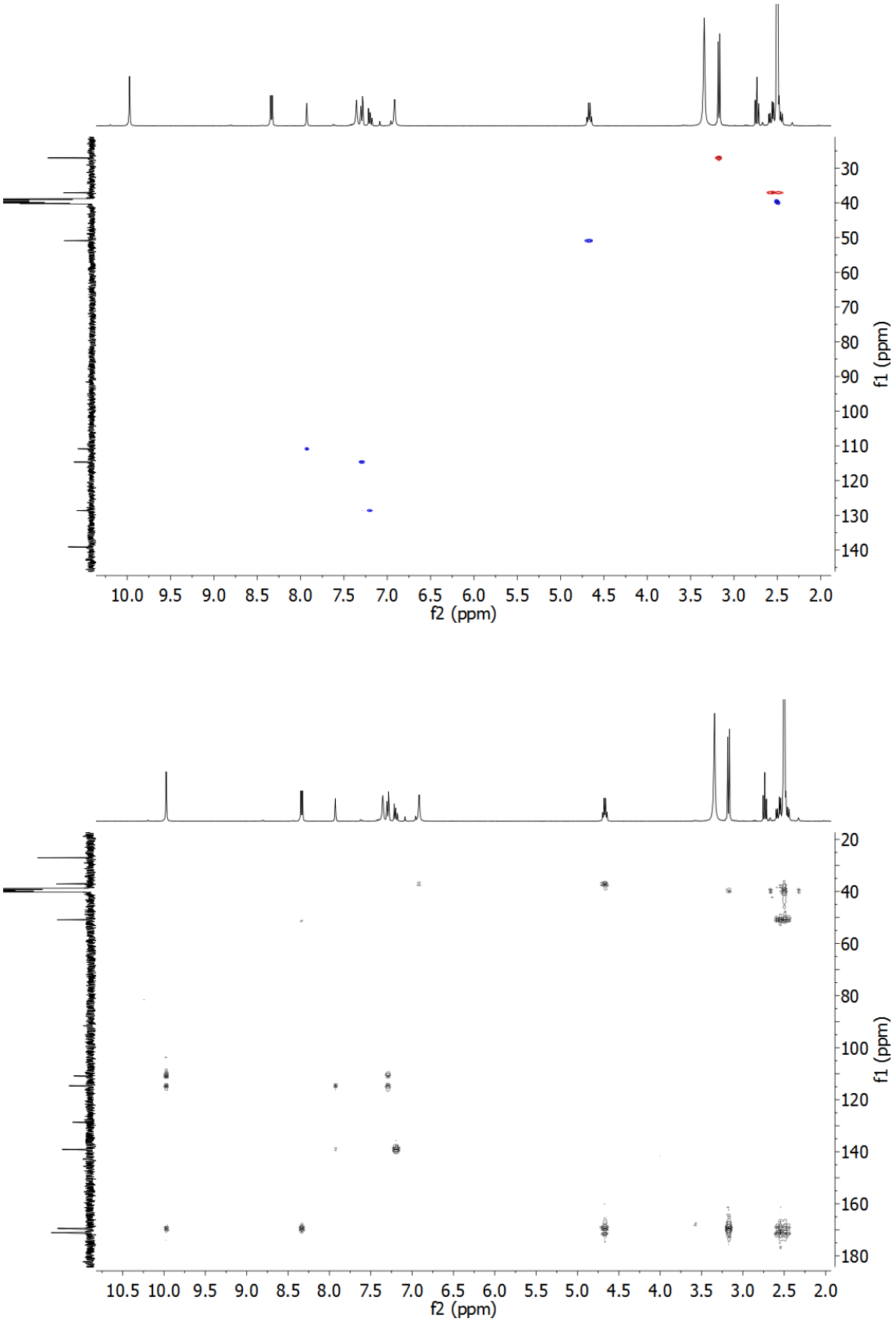


Figure S2: $^{1}H/^{13}C$ gHSQC (400 MHz, 298 K in DMSO-d_6) and $^{1}H/^{13}C$ gHMBC (400 MHz, 298 K in DMSO-d_6) spectra of [**1a**].

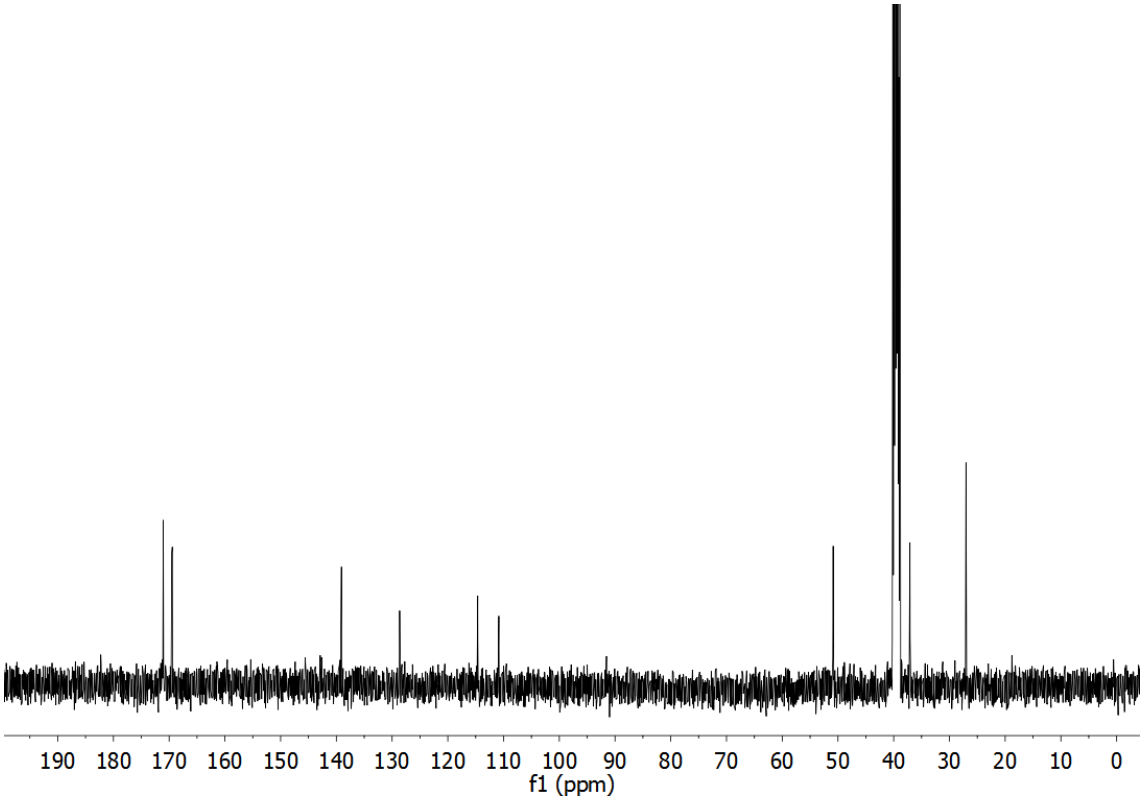


Figure S3: ^{13}C (101 MHz, 298 K in DMSO-d_6) spectrum of [**1a**].

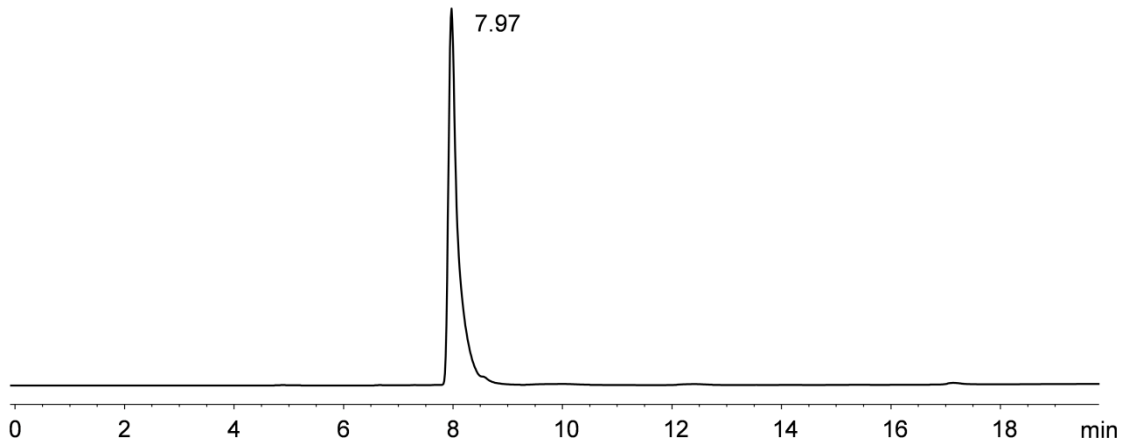


Figure S4: HPLC of [**1a**] (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

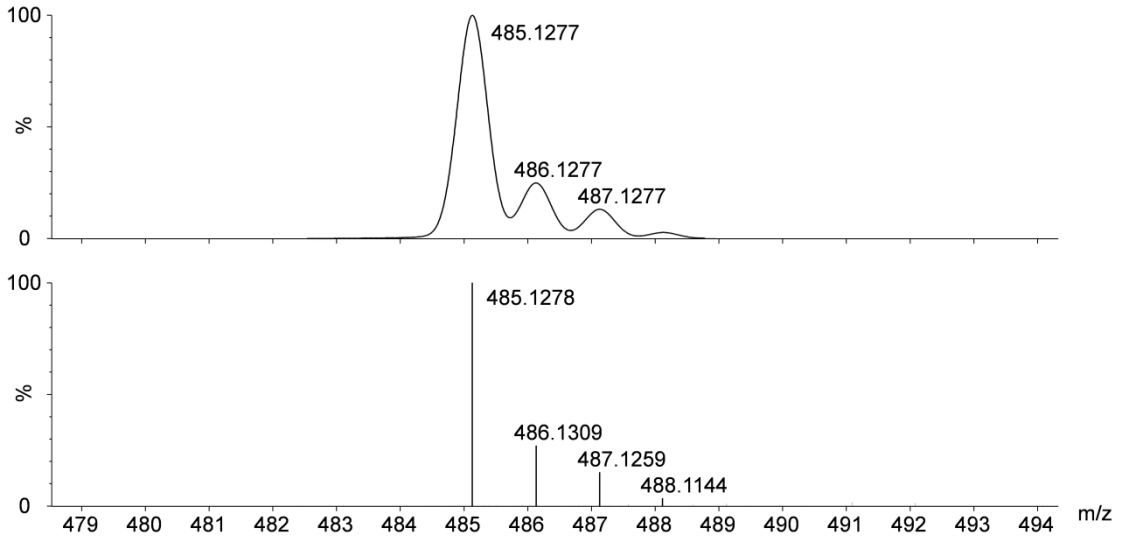


Figure S5: Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1a**].

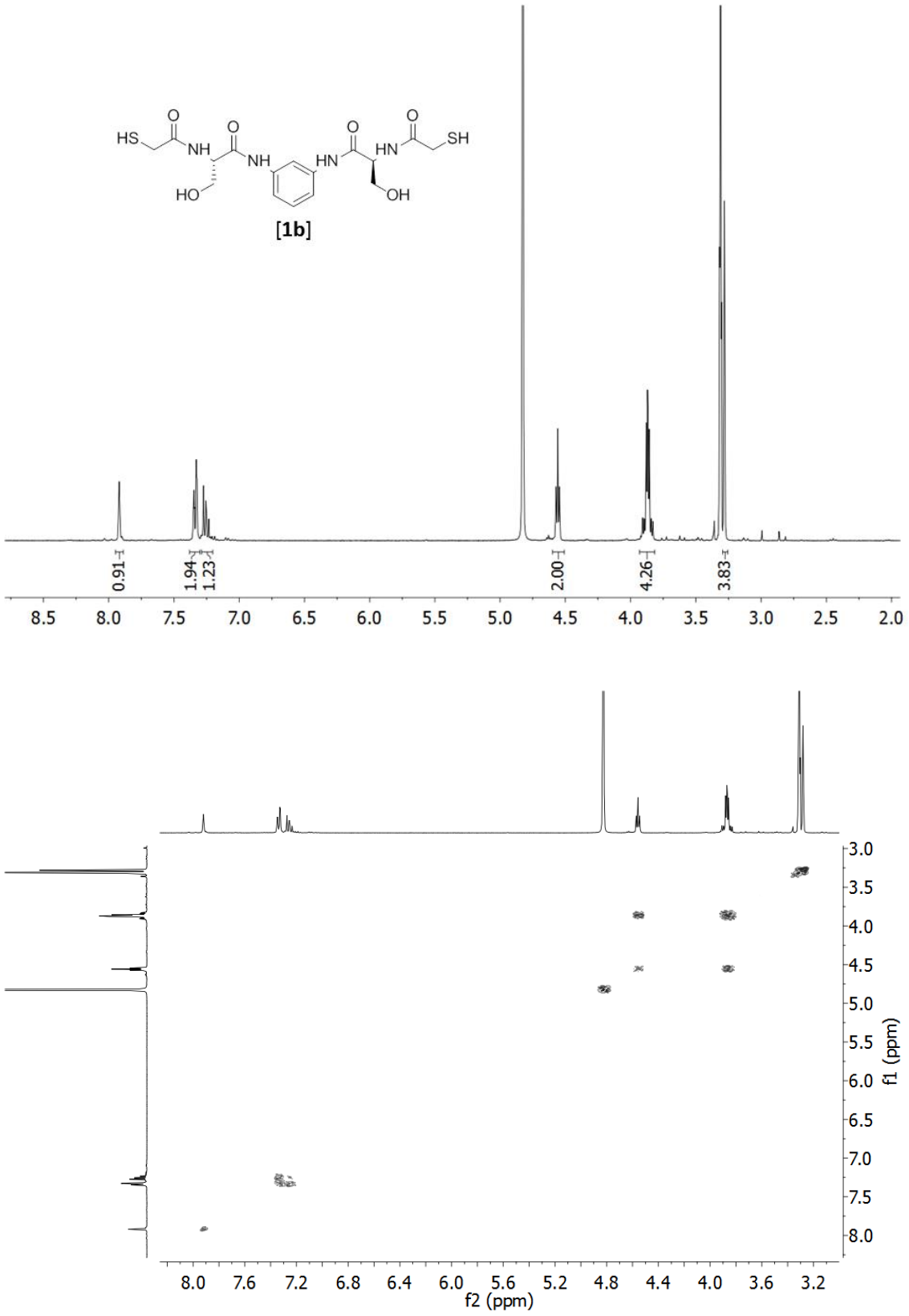


Figure S6: ^{1}H (400 MHz, 298 K in MeOD-d_4) and gCOSY (400 MHz, 298 K in MeOD-d_4) spectra of [**1b**].

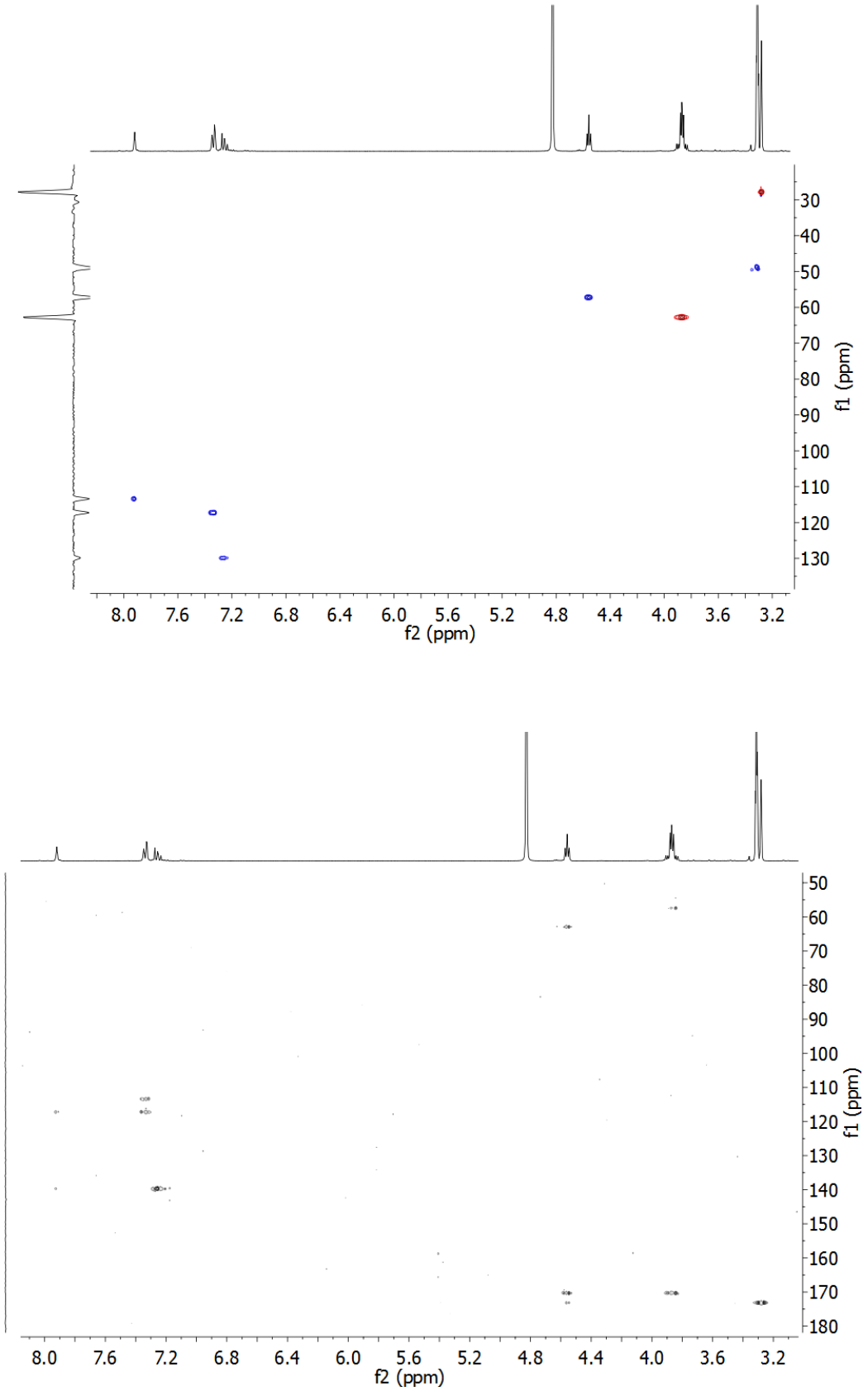


Figure S7: $^{1}H/^{13}C$ gHSQC (400 MHz, 298 K in MeOD-d_4) and $^{1}H/^{13}C$ gHMBC (400 MHz, 298 K in MeOD-d_4) spectra of [**1b**].

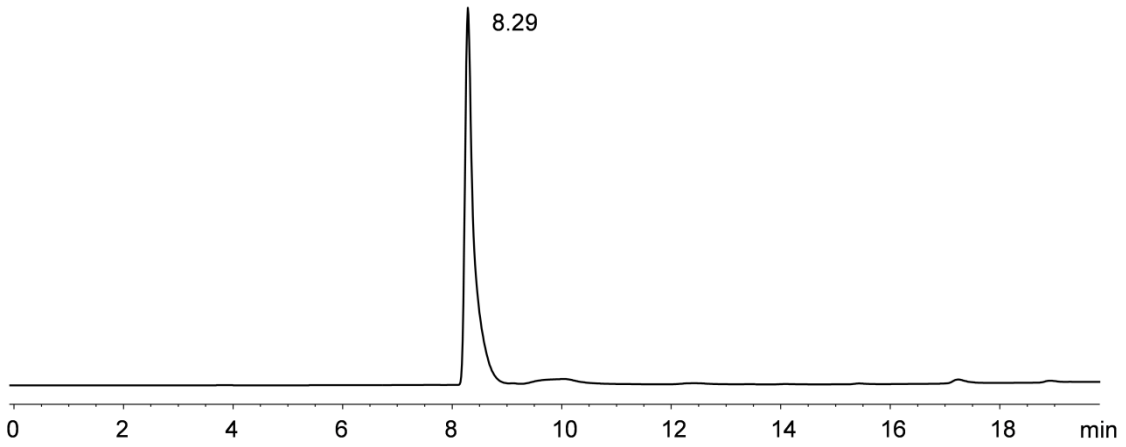


Figure S8: HPLC of [**1b**] (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

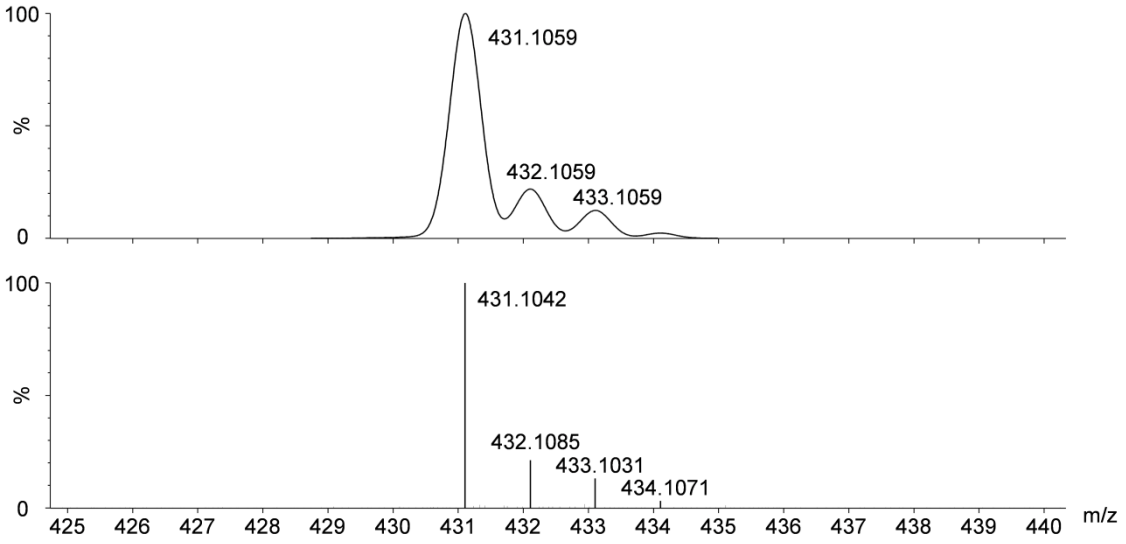


Figure S9: Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1b**].

DYNAMIC COMBINATORIAL LIBRARIES

General procedure for the preparation and HPLC analysis of the DCLs

A 66.7 mM BIS-Tris methane buffer was prepared by dissolving 1.39 g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq).

The reaction mixtures were prepared by dilution of individual stocks of the building blocks (BBs) **1a**, **1b**, **2**, **3a-g**. For those experiments to be compared, the reaction mixtures were prepared by dilution of a stock mixture of the BBs, ensuring no differences in concentration between the reaction mixtures of the same batch. Unless otherwise specified, the conditions for the generation of the DCLs were: 0.5 mM of the di- and tripodal BBs **1a**, **1b** and **2** in a 50 mM BIS-Tris methane buffer (pH 6.5) with 25% DMSO. The concentrations of the monopodal BBs **3a-g** are specified for each of the experiments.

After complete oxidation of the free thiols (24 hours at room temperature) the reaction mixtures were analysed by HPLC. The HPLC samples were prepared by adding 40 μL of the corresponding reaction mixture to 65 μL of a solution of 89% H_2O, 10% CH_3CN and 1% TFA. Eluent used: 2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 40% CH_3CN over 48 min.

Mixture of **1a**+**2** and **1a**+**2**+**3a**

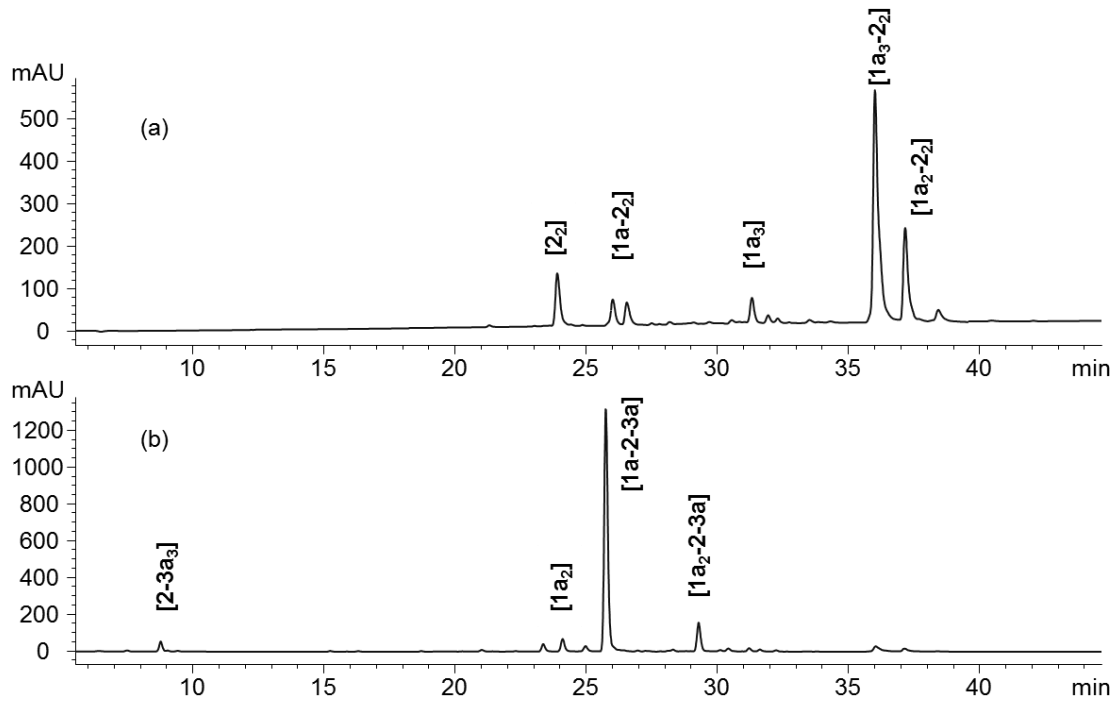


Figure S10: HPLC traces of the mixture **1a**+**2** in the absence of any monopodal BB (a), and in the presence of 2.5 mM of **3a** (b).

Mixture of **1b**+**2** and **1b**+**2**+**3a**

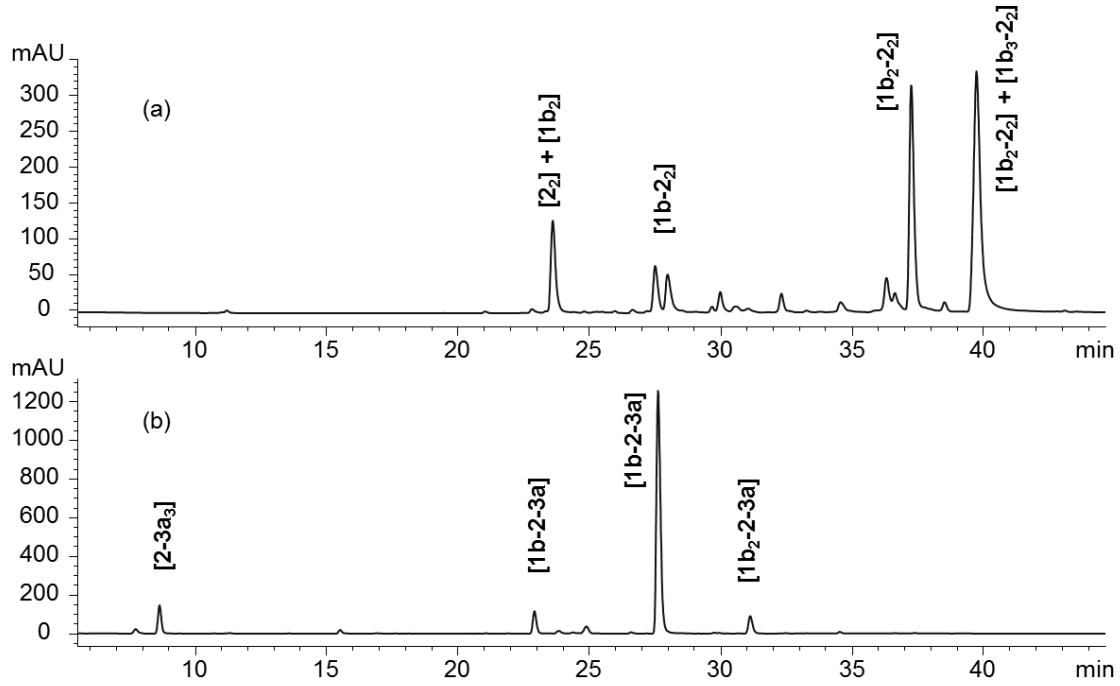


Figure S11: HPLC traces of the mixture **1b**+**2** in the absence of any monopodal BB (a), and in the presence of 2.5 mM of **3a** (b).

Mixture of **1a**+**2** with different concentrations of **3a**

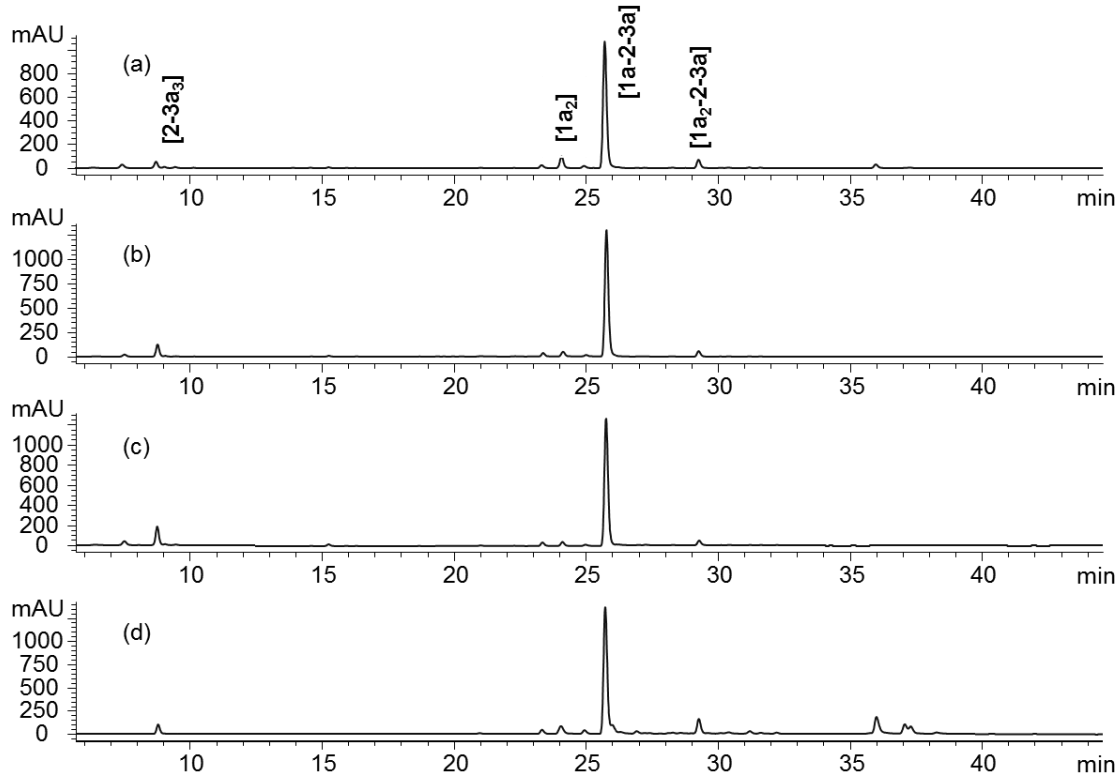


Figure S12: HPLC traces of the mixture of **1a**+**2** in the presence of 100 mM (a), 20 mM (b), 5 mM (c) and 0.5 mM (d) of **3a.**

Mixture of **1a**+**2**+**3b-g**

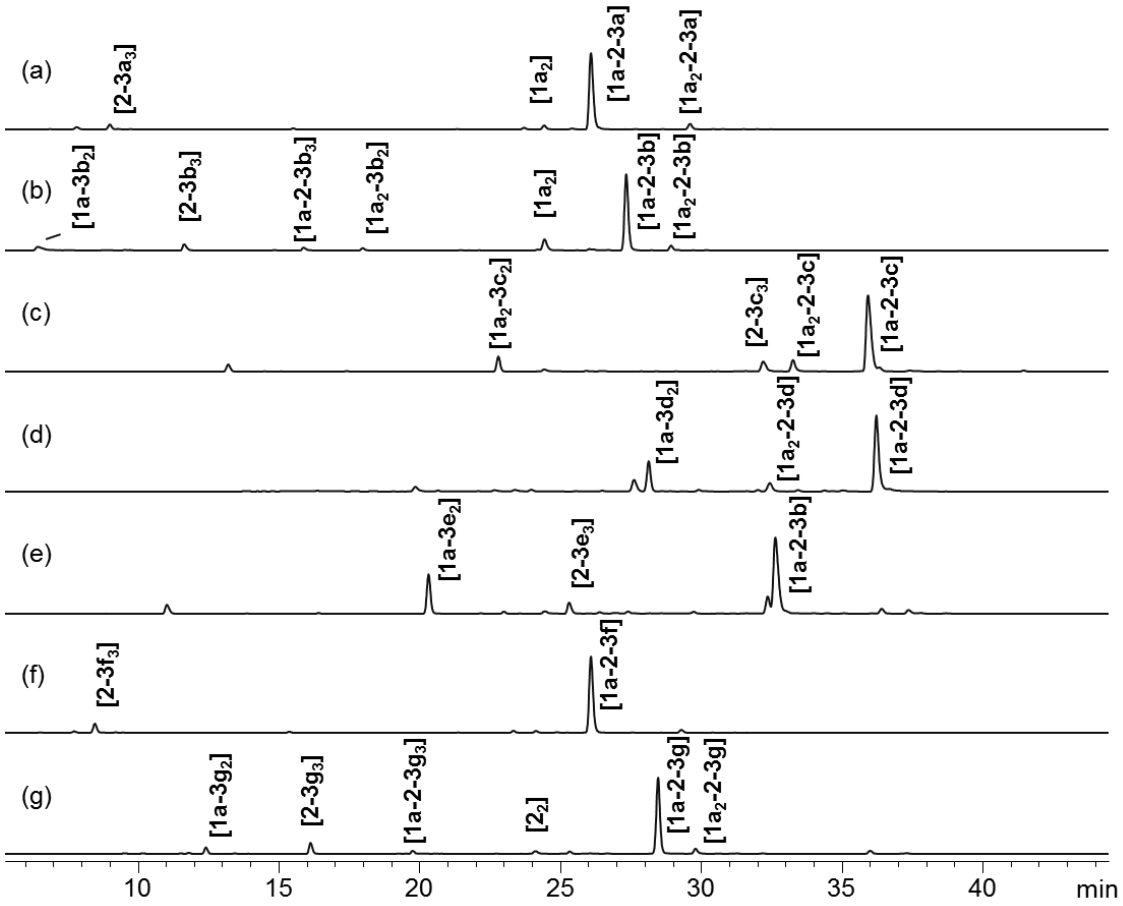


Figure S13: HPLC traces of the mixture of **1a**+**2** in the presence of 2.5 mM of **3a** (a), **3b** (b), **3c** (c), **3d** (d), **3e** (e), **3f** (f) and **3g** (g).

Mixture of **1a**+**2**+**3a**+**3f**+**3g**

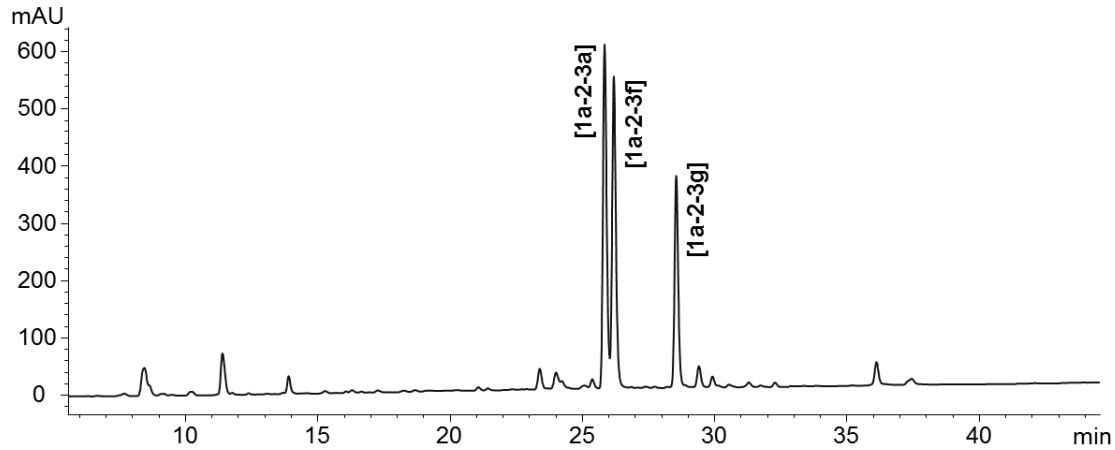


Figure S14: HPLC traces of the mixture of **1a**+**2**+**3a**+**3f**+**3g** (each monopodal BB at 1.0 mM).

Mixture of **1a**+**2**+**3a** at non-equimolar proportions of **1a** and **2**

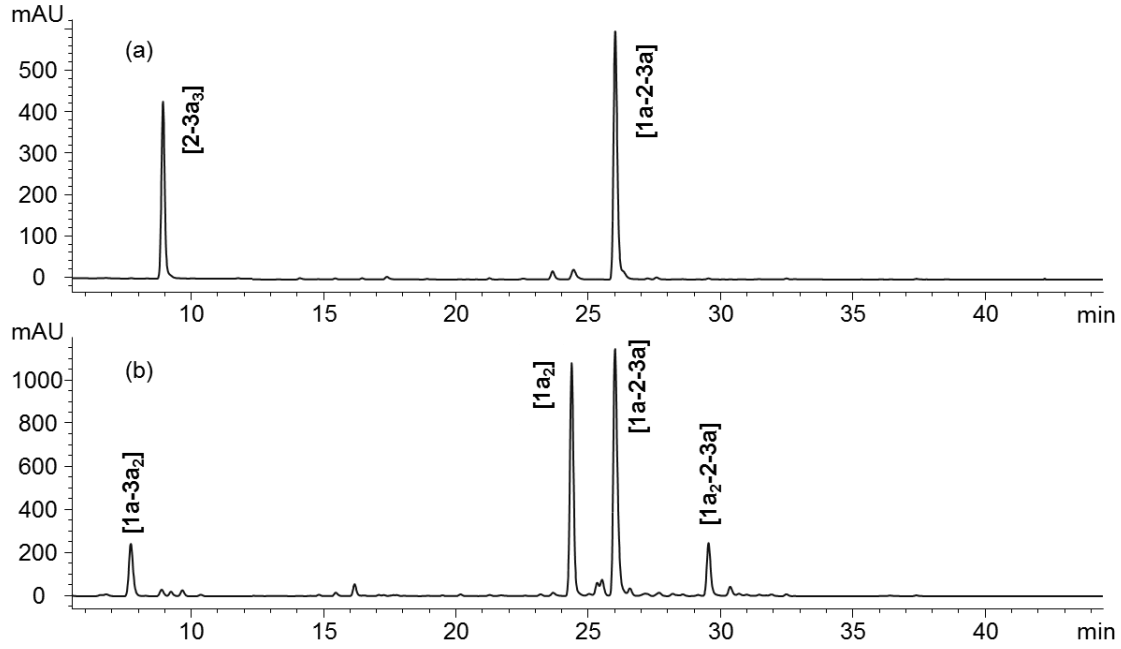


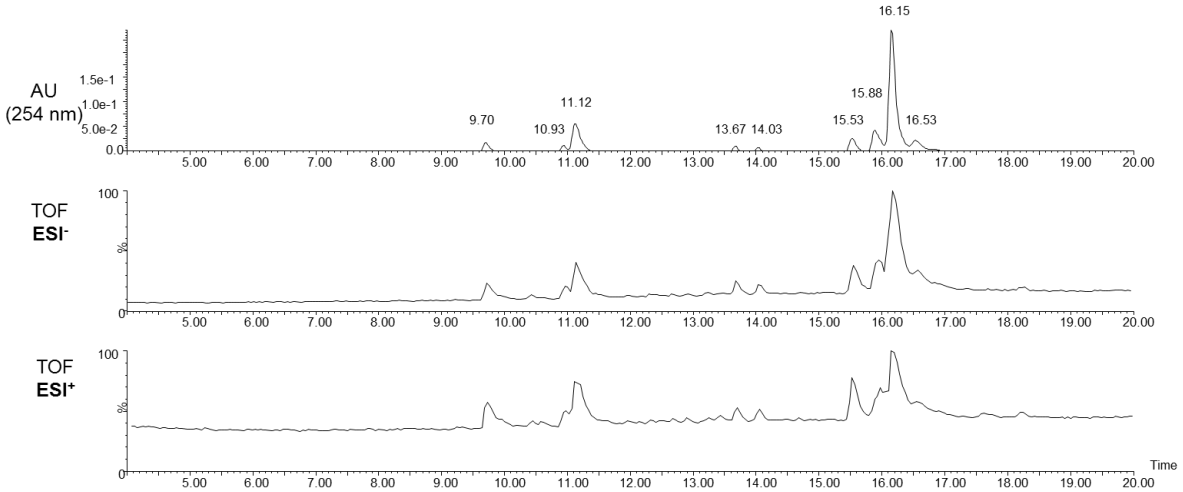
Figure S15: HPLC traces of the mixture of **1a**+**2**+**3a** at 0.5 mM, 1.0 mM and 4.0 mM respectively (a), and at 1.0 mM, 0.5 mM and 3.5 mM respectively (b).

MASS SPECTROMETRY

General procedure for the analysis of the DCLs by HRMS

The HRMS (UPLC-ESI-TOF) samples were prepared by adding 20 µL of the corresponding reaction mixture to 40 µL of a solution of 89% H_2O, 10% MeCN and 1% TFA. Eluent used: 2.5 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 50% CH_3CN over 27.5 min.

Mixture of **1a**+**2** (0.5 mM each, pH.6.5)

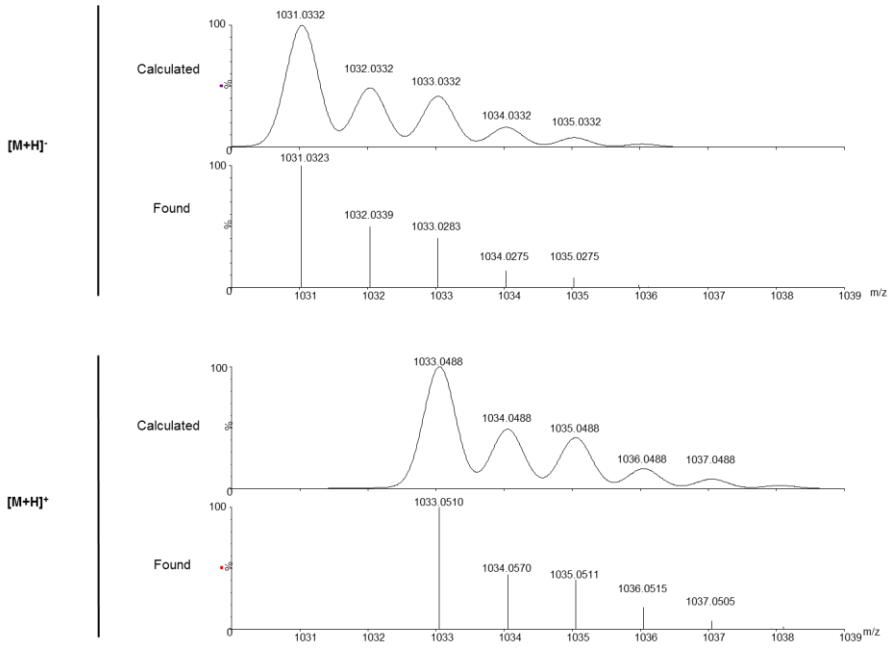


Identification of the products:

[$\mathbf{2_2}$]
Retention time: 9.70 min.

Chemical Formula:
$C_{36}H_{36}N_6O_{18}S_6$

Exact Mass:
1032.0410

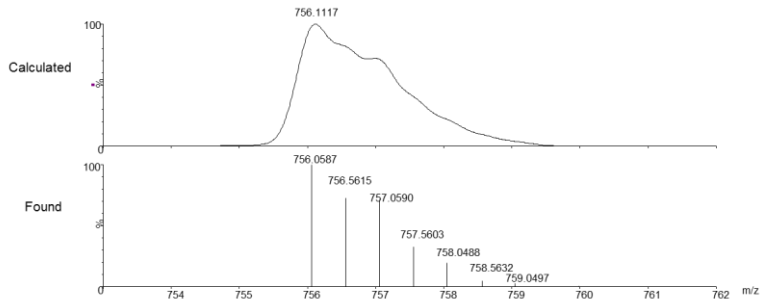


[**1a-2$_2$**]
Retention time: 11.12 min.

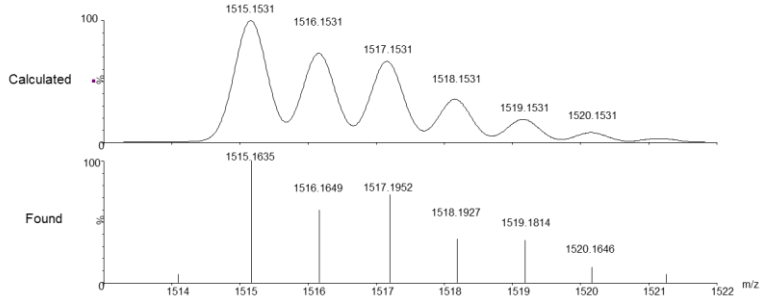
Chemical Formula:
$C_{54}H_{58}N_{12}O_{24}S_8$

Exact Mass:
1514.1453

[M+H]$^{2-}$



[M+H]$^{+}$

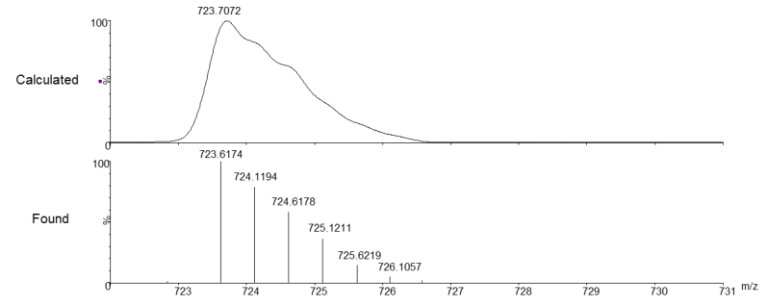


[**1a$_3$**]
Retention time: 13.67 min.

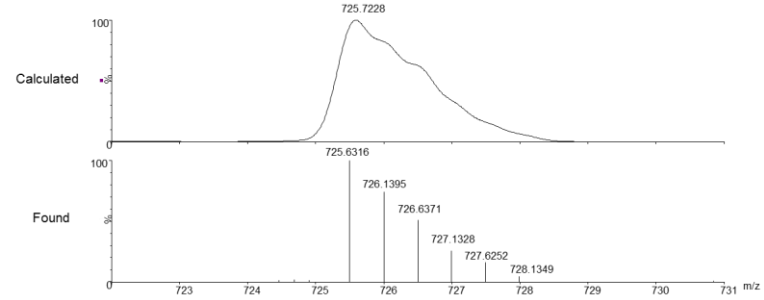
Chemical Formula:
$C_{54}H_{66}N_{18}O_{18}S_6$

Exact Mass:
1446.3127

[M+H]$^{2-}$



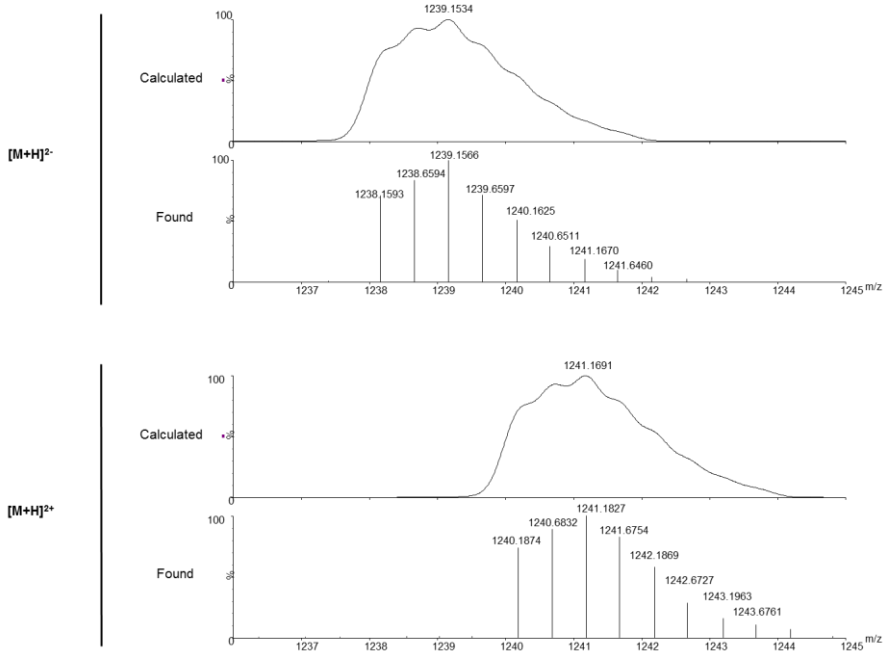
[M+H]$^{2+}$



[**1a$_3$-2$_2$**]
Retention time: 15.88 min.

Chemical Formula:
$C_{90}H_{102}N_{24}O_{36}S_{12}$

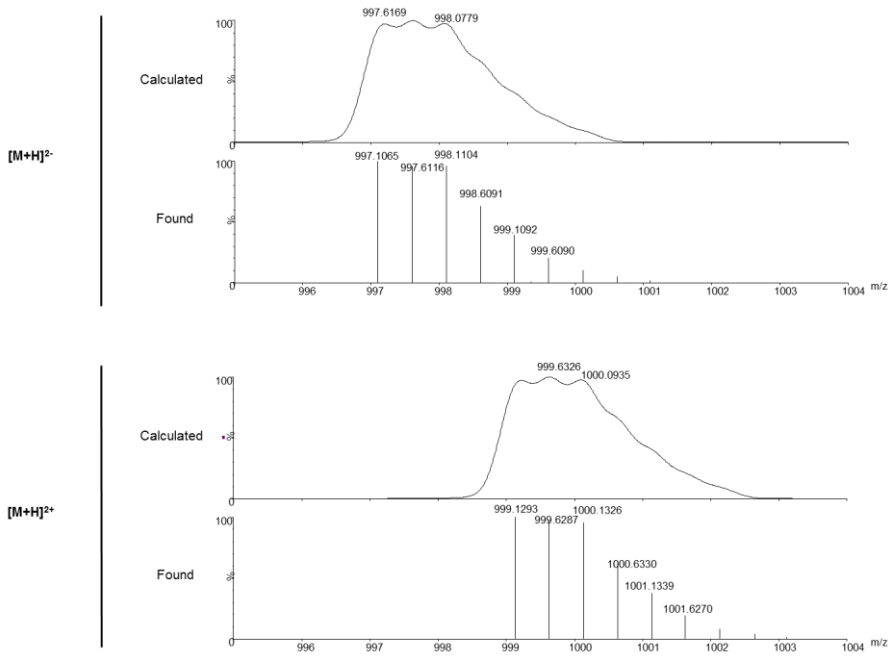
Exact Mass:
2478.3537



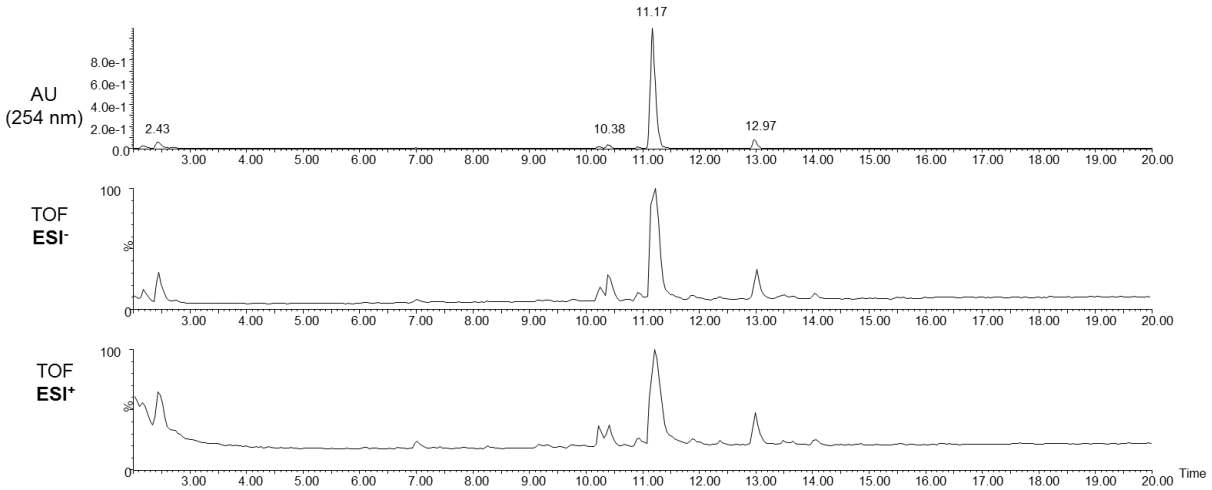
[**1a$_2$-2$_2$**]
Retention time: 16.17 min.

Chemical Formula:
$C_{72}H_{80}N_{18}O_{30}S_{10}$

Exact Mass:
1996.2495



Mixture of **1a**+**2**+**3a** (0.5, 0.5 and 2.5 mM respectively, pH.6.5)

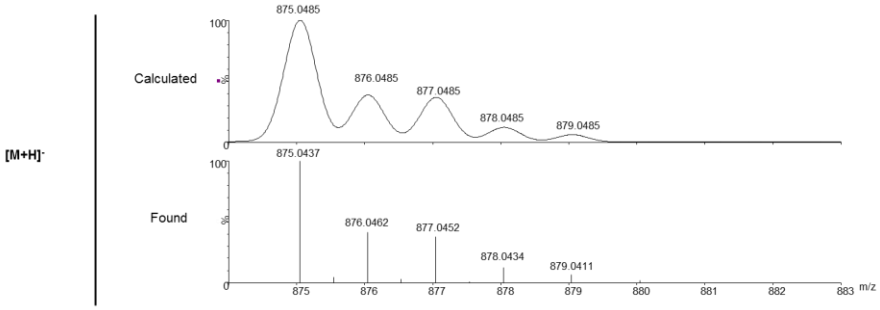


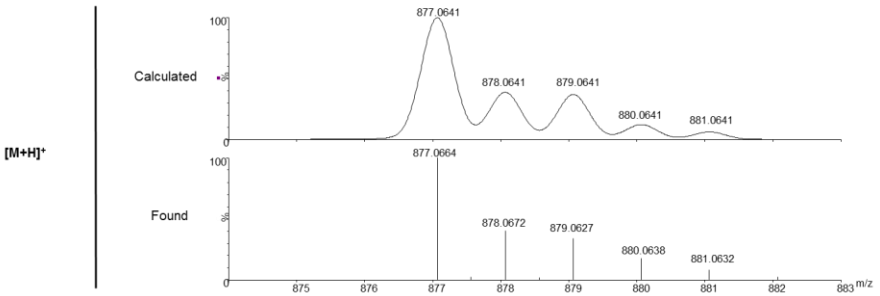
Identification of the products:

[**2-3a$_3$**]
Retention time: 2.43 min.

Chemical Formula:
$C_{27}H_{36}N_6O_{15}S_6$

Exact Mass:
876.0563



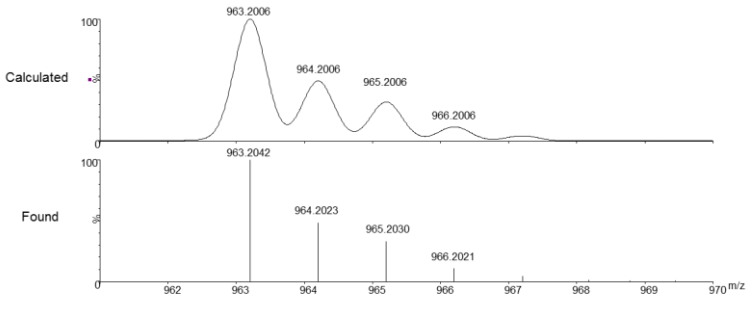


[**1a$_2$**]
Retention time: 10.38 min.

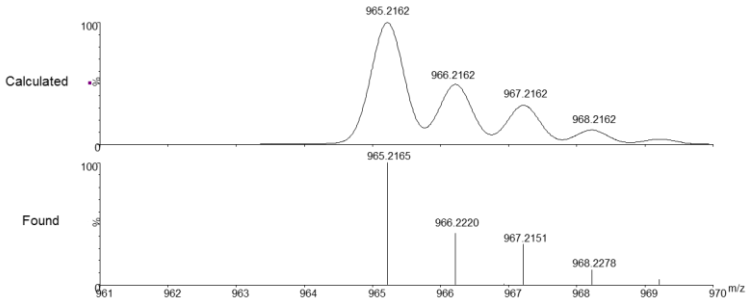
Chemical Formula:
$C_{36}H_{44}N_{12}O_{12}S_4$

Exact Mass:
964.2084

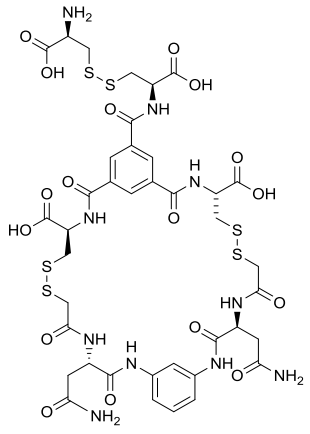
[M+H]$^-$



[M+H]$^+$



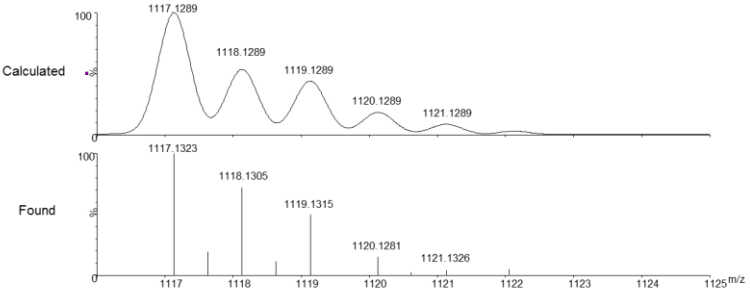
[**1a-2-3a**]
Retention time: 11.17 min.



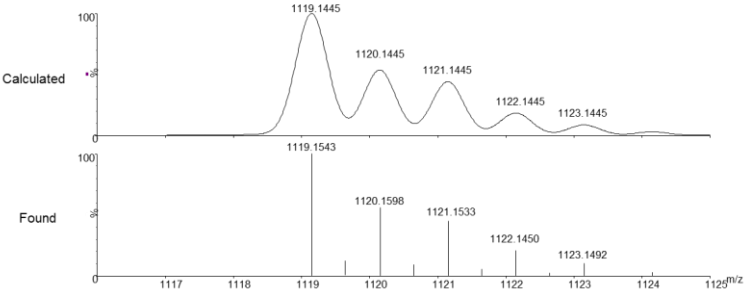
Chemical Formula:
$C_{39}H_{46}N_{10}O_{17}S_6$

Exact Mass:
1118.1367

[M+H]$^-$



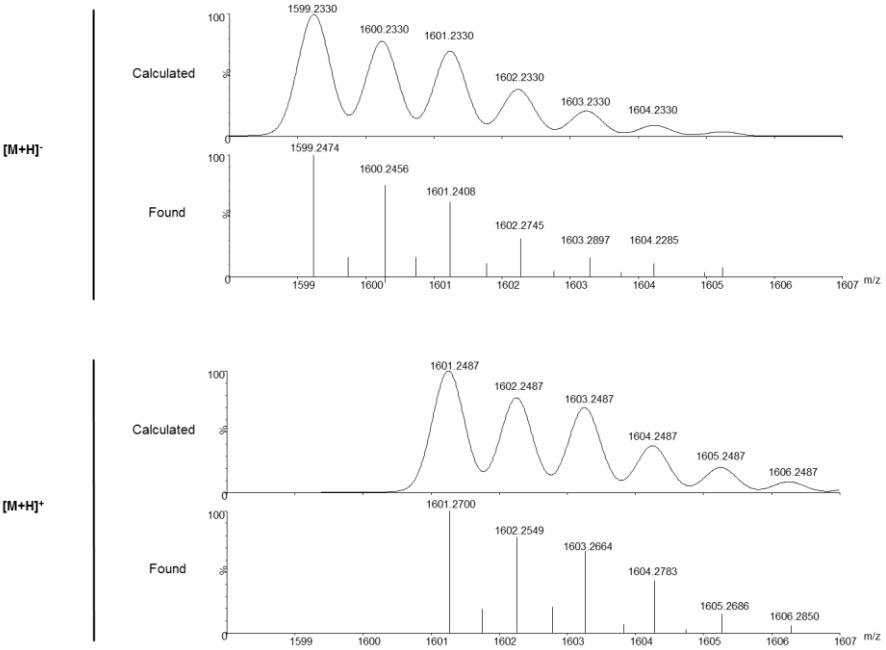
[M+H]$^+$



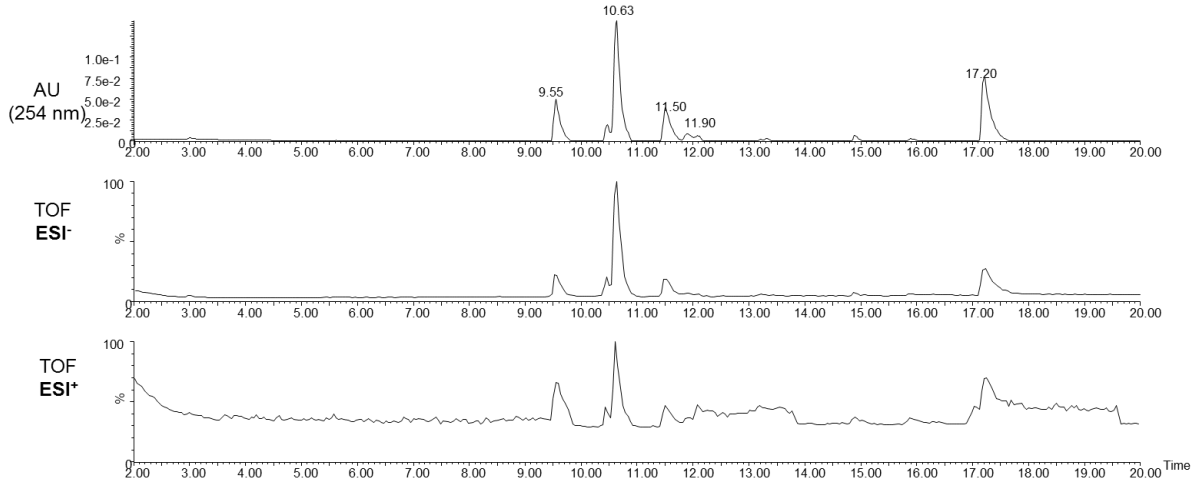
[**1a$_2$-2-3a**]
Retention time: 12.37 min.

Chemical Formula:
$C_{57}H_{68}N_{16}O_{23}S_8$

Exact Mass:
1600.2409



Mixture of **1b**+**2** (0.5 mM each, pH.6.5)



Identification of the products:

[**2$_2$**] (previously identified)
Retention time: 9.55 min.

Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$

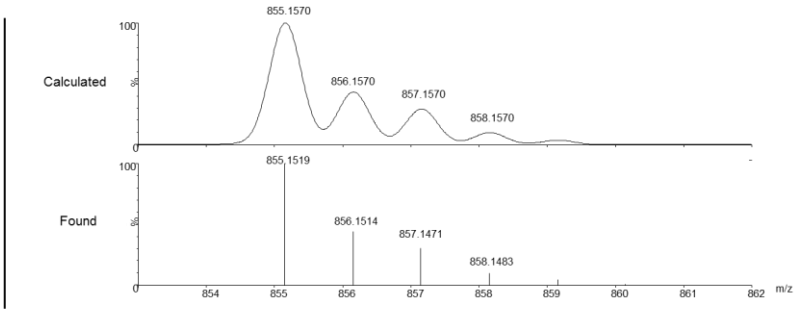
Exact Mass:
1032.0410

[**1b**$_2$]
Retention time: 10.63 min

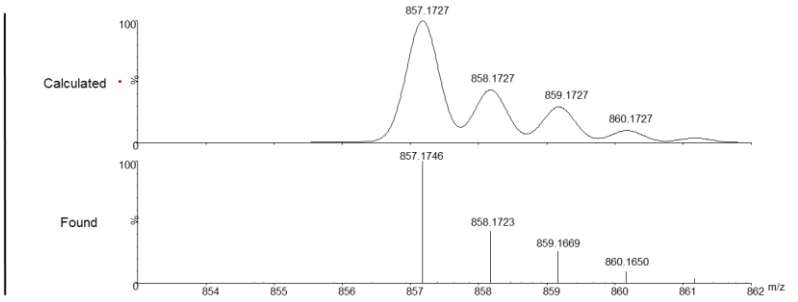
Chemical Formula:
$C_{32}H_{40}N_8O_{12}S_4$

Exact Mass:
856.1649

[M+H]$^-$



[M+H]$^+$

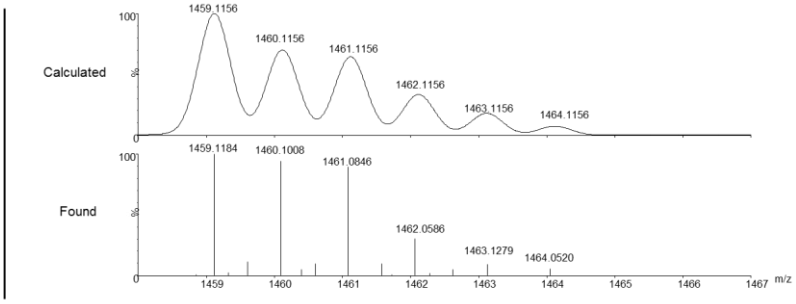


[**1b-2**$_2$]
Retention time: 11.50 and 11.90 min

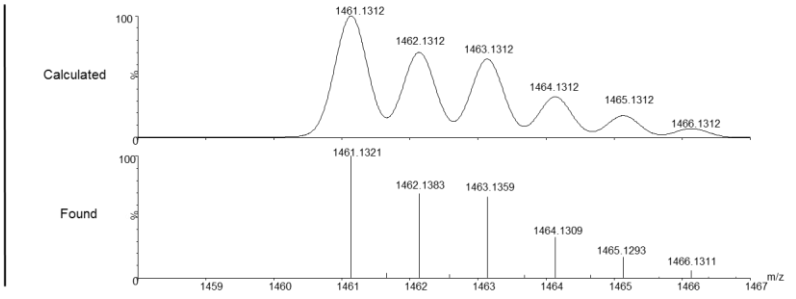
Chemical Formula:
$C_{52}H_{56}N_{10}O_{24}S_8$

Exact Mass:
1460.1235

[M+H]$^-$



[M+H]$^+$

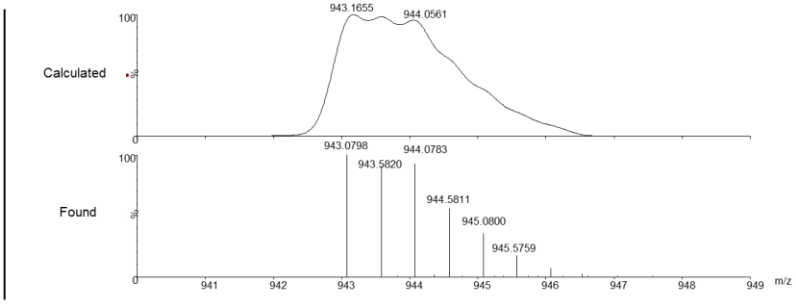


[**1b$_2$-2$_2$**]
Retention time: 15.90 and 17.22 min

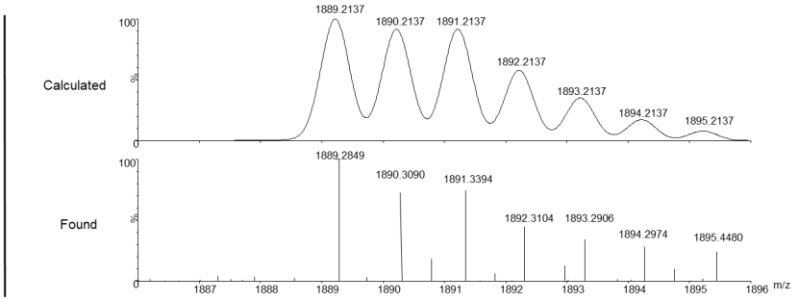
Chemical Formula:
$C_{68}H_{76}N_{14}O_{30}S_{10}$

Exact Mass:
1888.2059

[M+H]$^{2-}$



[M+H]$^{+}$

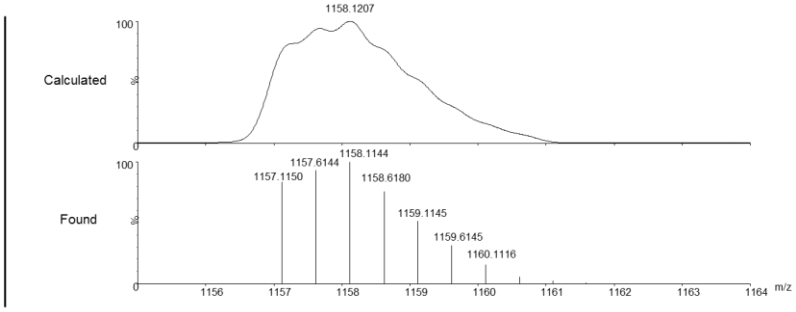


[**1b$_3$-2$_2$**]
Retention time: 17.22 min

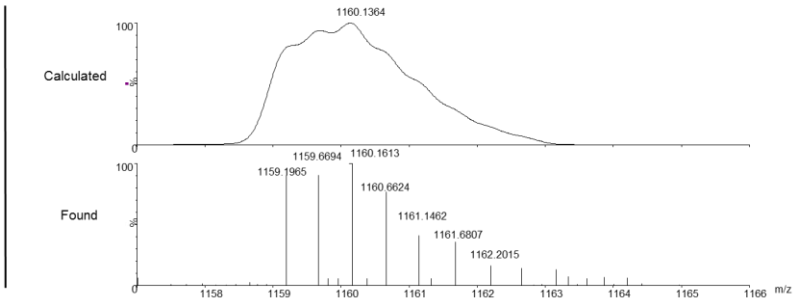
Chemical Formula:
$C_{84}H_{96}N_{18}O_{36}S_{12}$

Exact Mass:
2316.2883

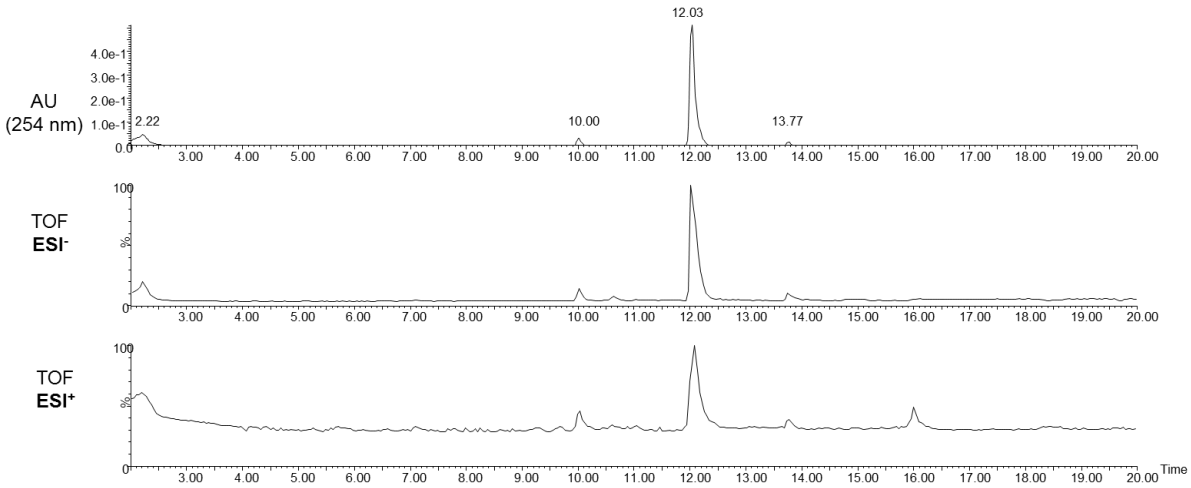
[M+H]$^{2-}$



[M+H]$^{2+}$



Mixture of **1b**+**2**+**3a** (0.5, 0.5 and 2.5 mM respectively, pH.6.5)



Identification of the products:

[**2-3a$_3$**] (previously identified)
Retention time: 2.22 min.

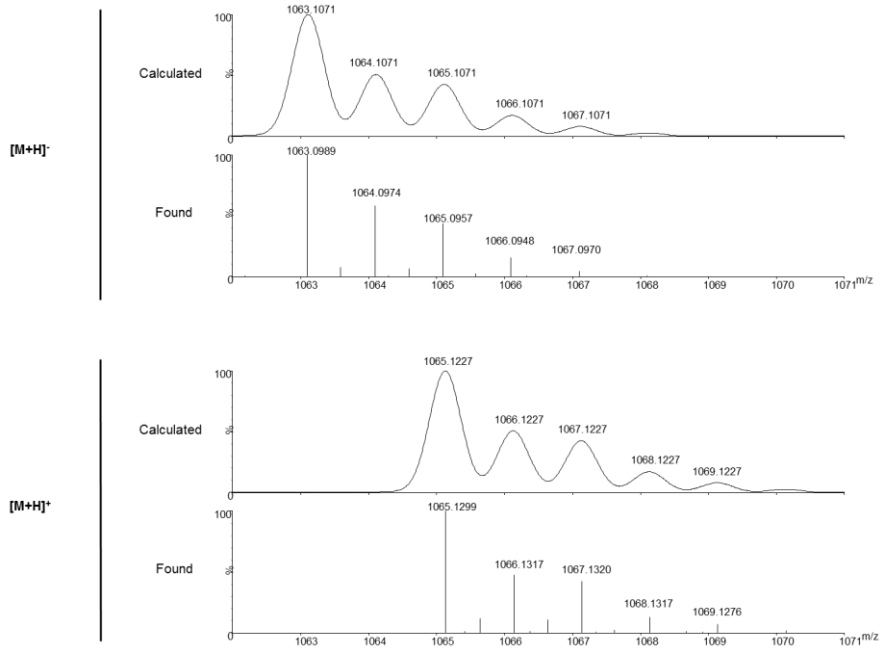
Chemical Formula:
$C_{27}H_{36}N_6O_{15}S_6$

Exact Mass:
876.0563

[**1b-2-3a**]
Retention time: 10.00 and12.03 min.

Chemical Formula:
$C_{37}H_{44}N_8O_{17}S_6$

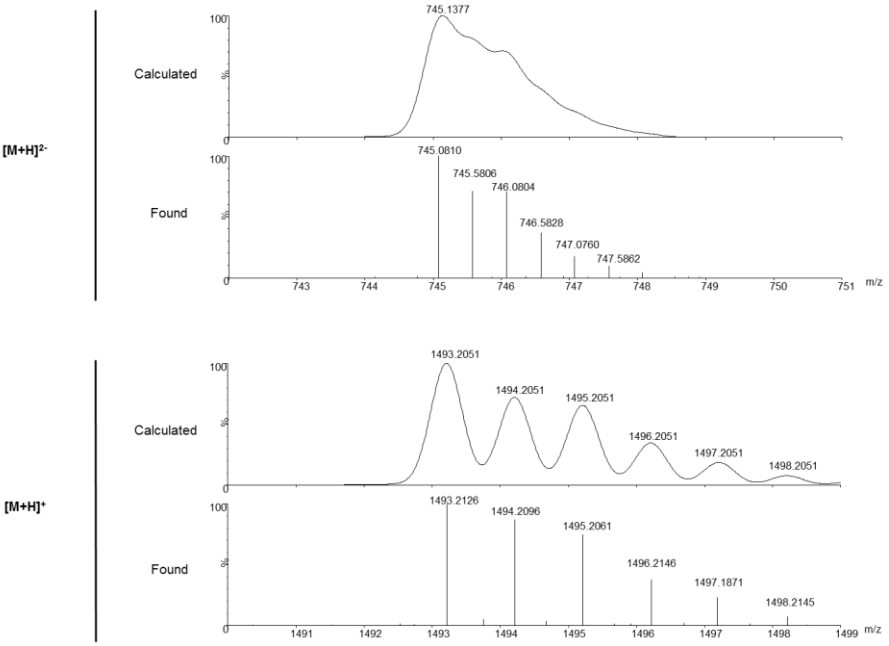
Exact Mass:
1064.1149



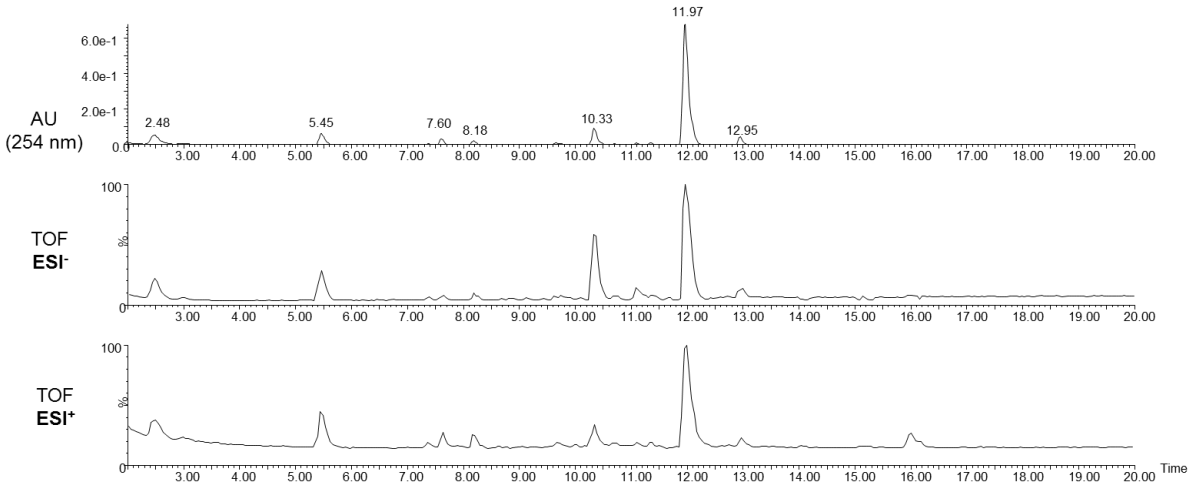
[**1b$_2$-2-3a**]
Retention time: 13.77 min.

Chemical Formula:
$C_{53}H_{64}N_{12}O_{23}S_8$

Exact Mass:
1492.1973



Mixture of **1a**+**2**+**3b** (0.5, 0.5 and 2.5 mM, pH.6.5)



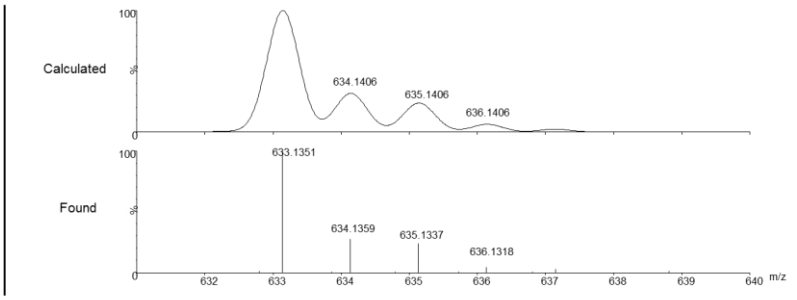
Identification of the products:

[**1a-3b$_2$**]
Retention time: 2.48 min.

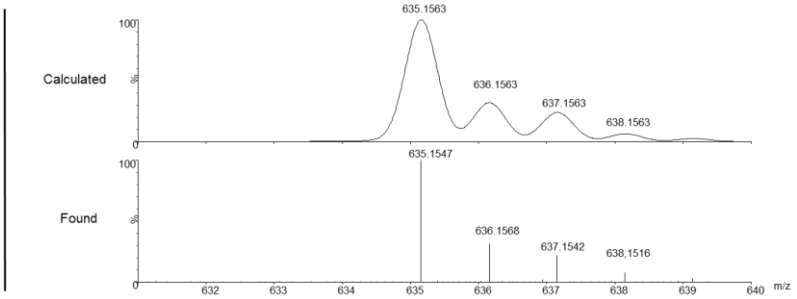
Chemical Formula:
$C_{22}H_{34}N_8O_6S_4$

Exact Mass:
634.1484

[M+H]$^-$



[M+H]$^+$

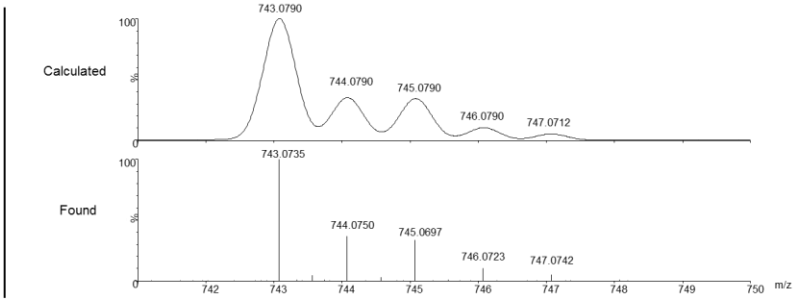


[**2-3b$_3$**]
Retention time: 5.45 min.

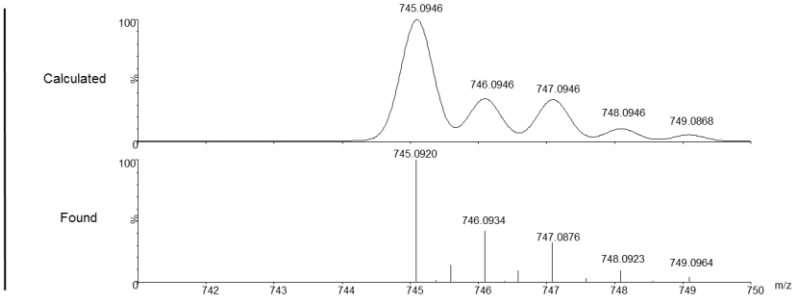
Chemical Formula:
$C_{24}H_{36}N_6O_9S_6$

Exact Mass:
744.0868

[M+H]$^-$



[M+H]$^+$

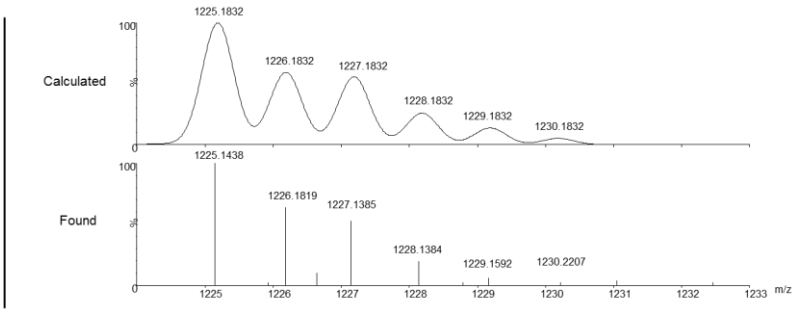


[**1a-2-3b$_3$**]
Retention time: 7.60 min.

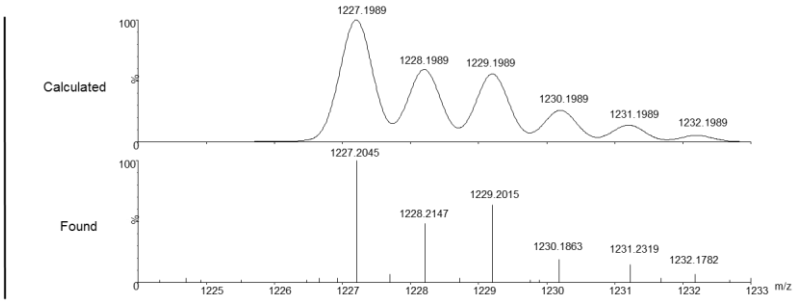
Chemical Formula:
$C_{42}H_{58}N_{12}O_{15}S_8$

Exact Mass:
1226.1910

[M+H]$^-$



[M+H]$^+$

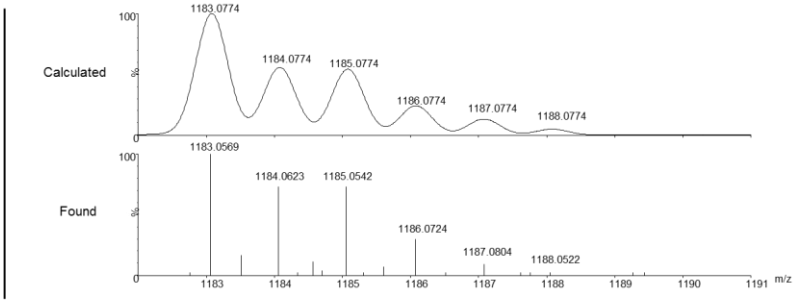


[**2$_2$-3b$_2$**]
Retention time: 8.18 min.

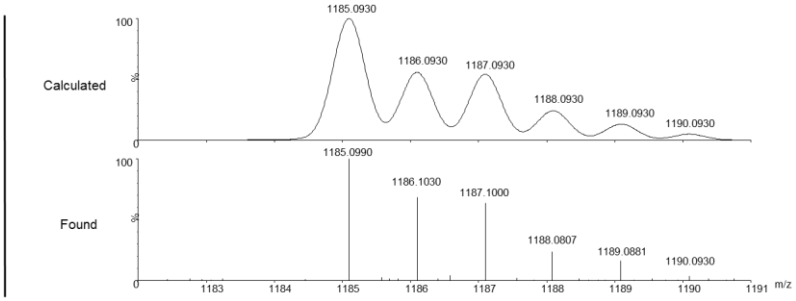
Chemical Formula:
$C_{40}H_{48}N_8O_{18}S_8$

Exact Mass:
1184.0852

[M+H]$^-$



[M+H]$^+$



[**1a$_2$**] (previously identified)
Retention time: 10.32 min.

Chemical Formula:
$C_{36}H_{44}N_{12}O_{12}S_4$

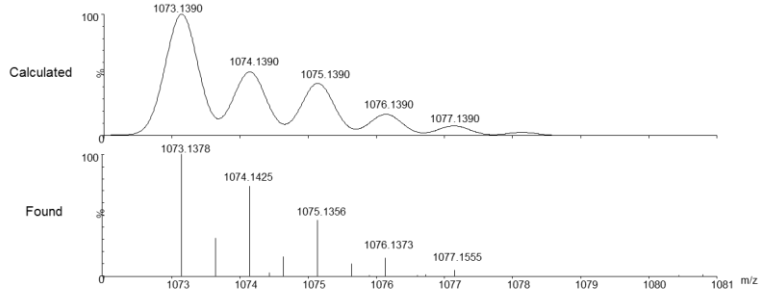
Exact Mass:
964.2084

[**1a-2-3b**]
Retention time: 11.97 min.

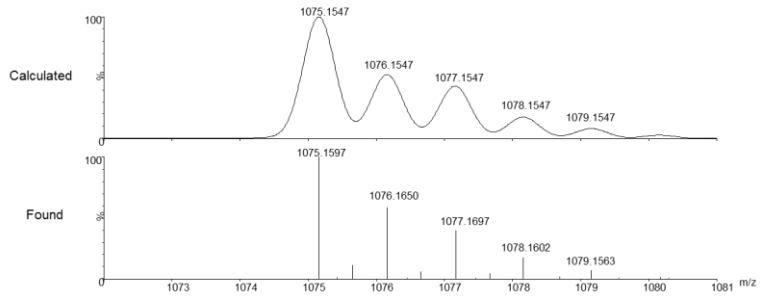
Chemical Formula:
$C_{38}H_{46}N_{10}O_{15}S_6$

Exact Mass:
1074.1468

[M+H]$^-$



[M+H]$^+$

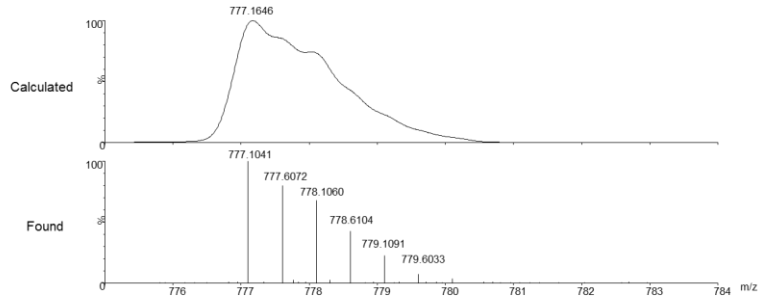


[**1a$_2$-2-3b**]
Retention time: 12.95 min.

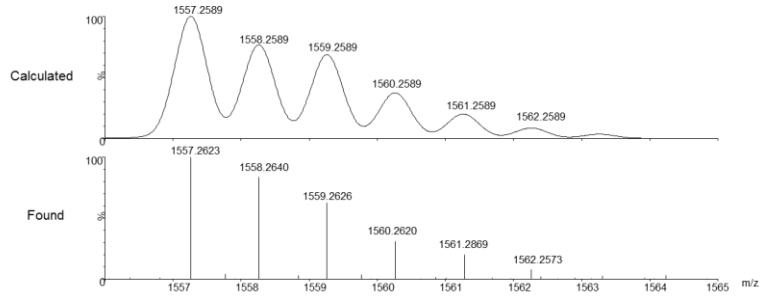
Chemical Formula:
$C_{56}H_{68}N_{16}O_{21}S_8$

Exact Mass:
1556.2511

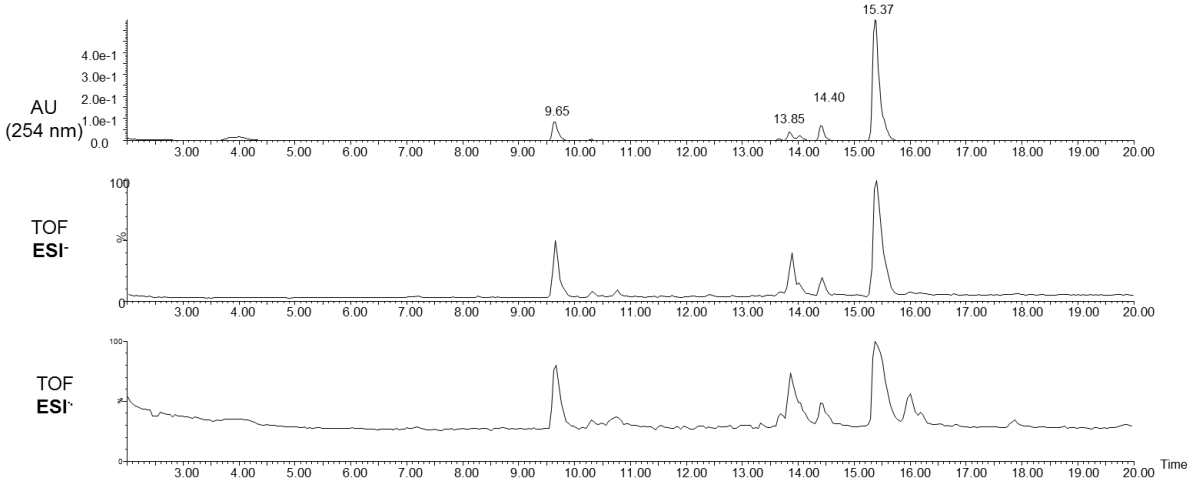
[M+H]$^{2-}$



[M+H]$^+$



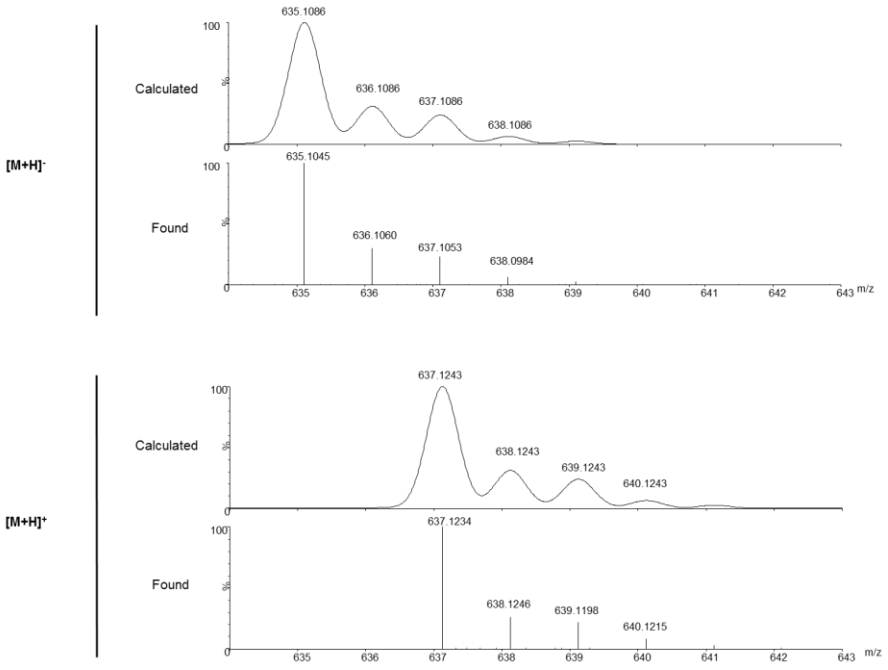
Mixture of **1a**+**2**+**3c** (0.5, 0.5 and 2.5 mM respectively, pH.6.5)



Identification of the products:

[**1a-3c$_2$**]
Retention time: 9.65 min.

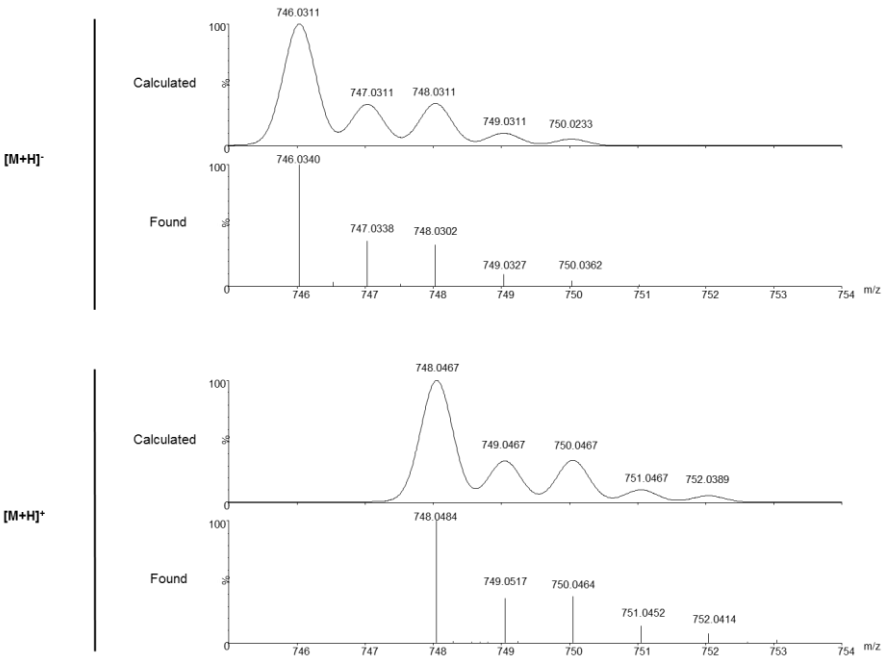
Chemical Formula:
$C_{22}H_{32}N_6O_8S_4$
Exact Mass:
636.1164



[**2-3c$_3$**]
Retention time: 13.85 min.

Chemical Formula:
$C_{24}H_{33}N_3O_{12}S_6$

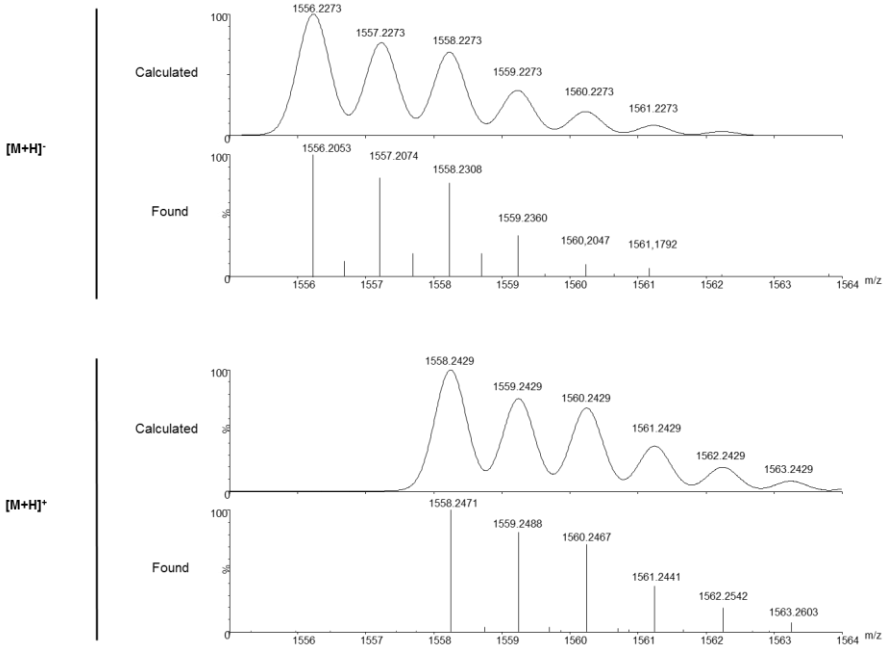
Exact Mass:
747.0388



[**1a$_2$-2-3c$_3$**]
Retention time: 14.40 min.

Chemical Formula:
$C_{56}H_{67}N_{15}O_{22}S_8$

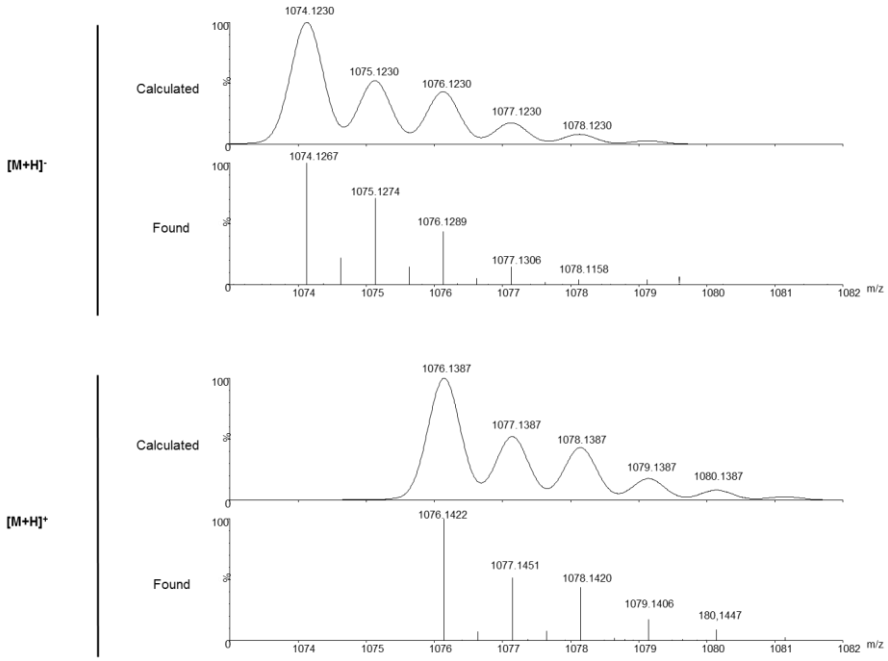
Exact Mass:
1557.2351



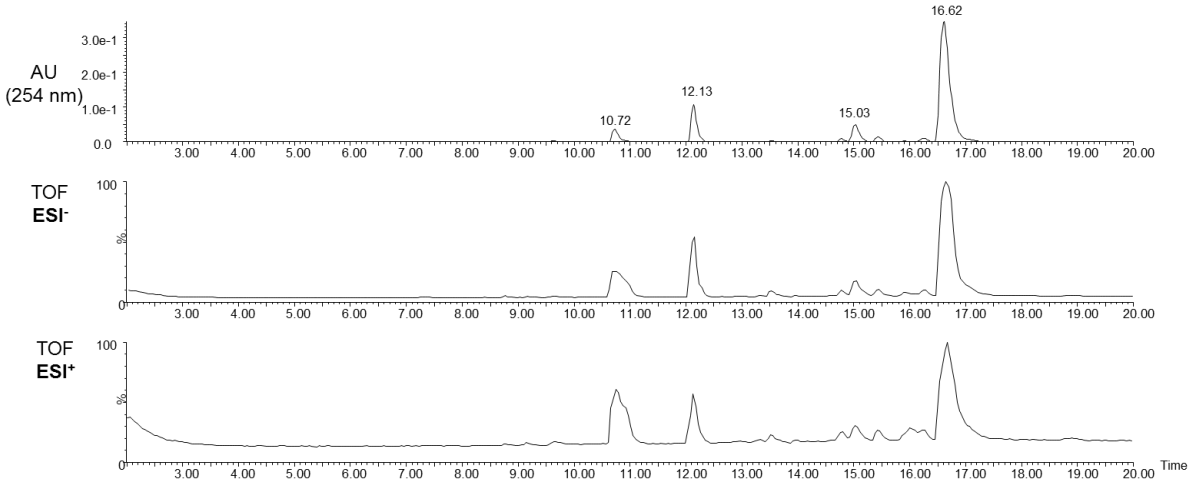
[**1a**$_2$-**2**-**3c**$_3$]
Retention time: 15.37 min.

Chemical Formula:
$C_{38}H_{45}N_9O_{16}S_6$

Exact Mass:
1075.1309



Mixture of **1a**+**2**+**3d** (0.5, 0.5 and 2.5 mM respectively, pH.6.5)



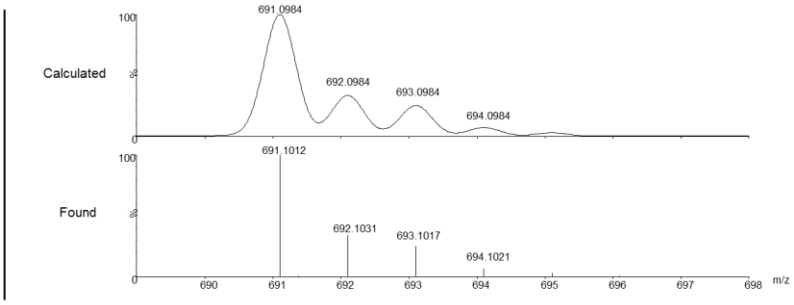
Identification of the products:

[**1a-3d$_2$**]
Retention time: 12.13 min.

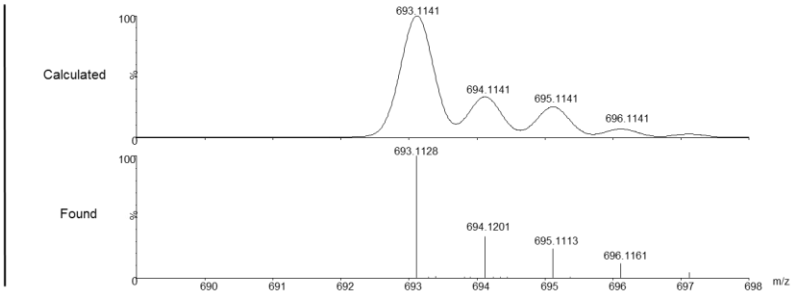
Chemical Formula:
$C_{24}H_{32}N_6O_{10}S_4$

Exact Mass:
692.1063

[M+H]$^-$



[M+H]$^+$

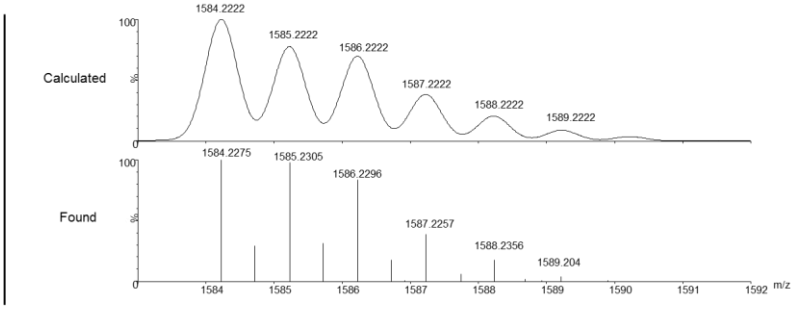


[**1a$_2$-2-3d**]
Retention time: 16.62 min.

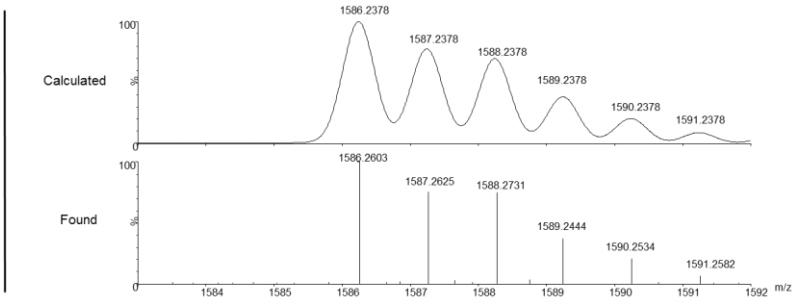
Chemical Formula:
$C_{57}H_{67}N_{15}O_{23}S_8$

Exact Mass:
1585.2300

[M+H]$^-$



[M+H]$^+$

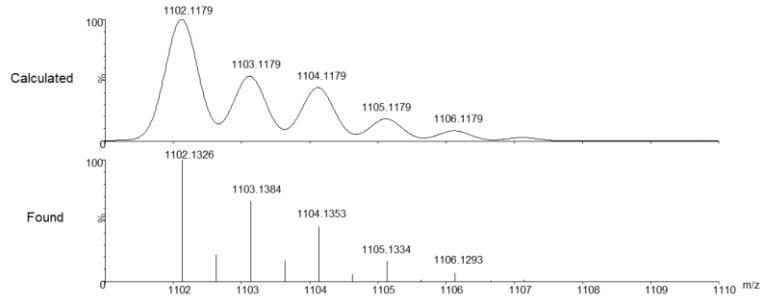


[**1a-2-3d**]
Retention time: 15.03 min.

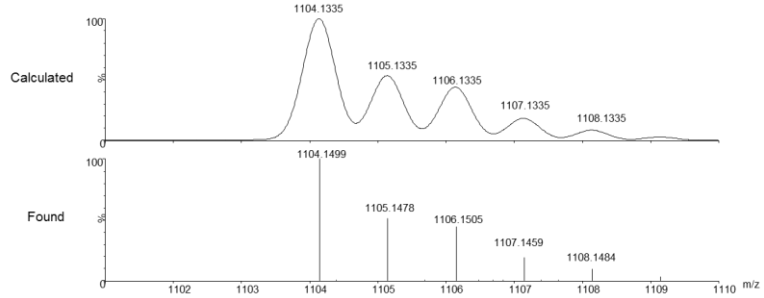
Chemical Formula:
$C_{39}H_{45}N_9O_{17}S_6$

Exact Mass:
1103.1258

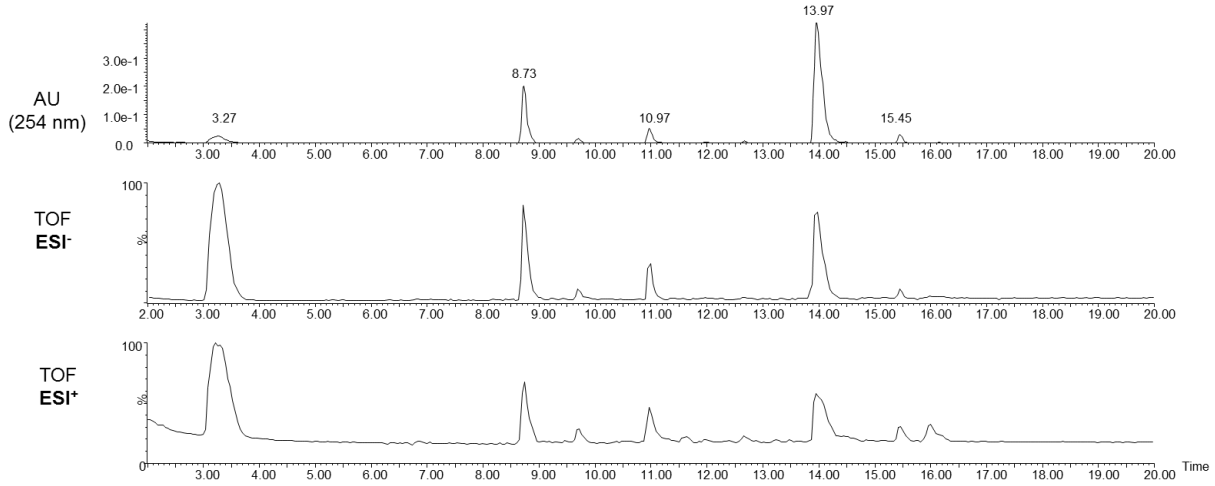
[M+H]⁻



[M+H]⁺



Mixture of **1a**+**2**+**3e** (0.5, 0.5 and 2.5 mM respectively, pH.6.5)



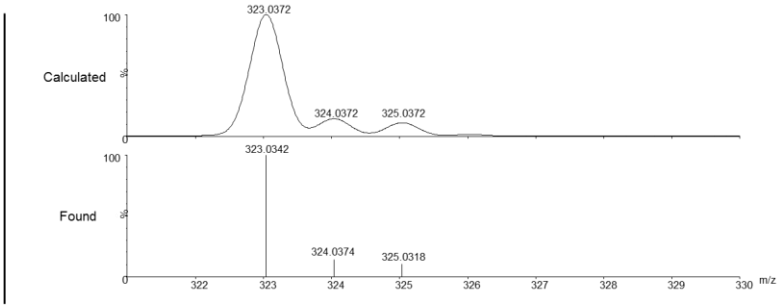
Identification of the products:

[**3e$_2$**]
Retention time: 3.27 min.

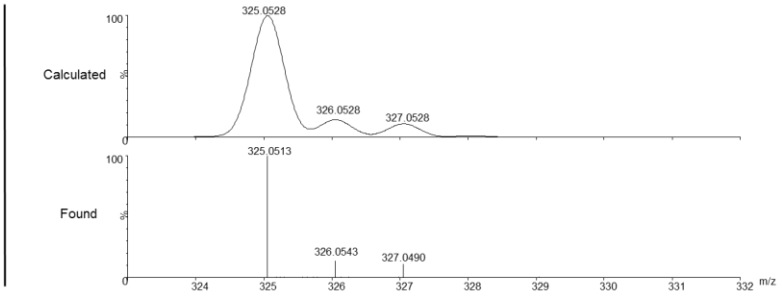
Chemical Formula:
$C_{10}H_{16}N_2O_6S_2$

Exact Mass:
324.0450

[M+H]$^-$



[M+H]$^+$

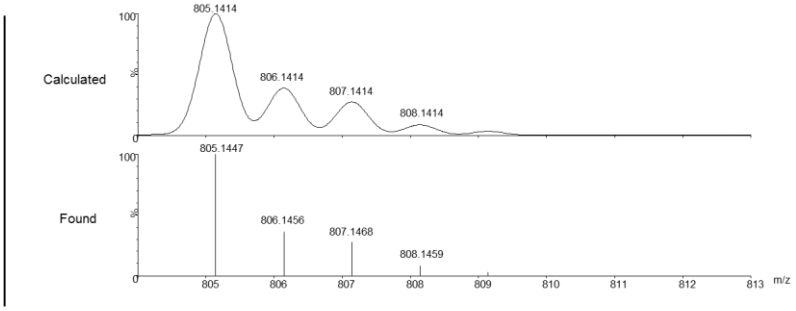


[**1a-3e$_2$**]
Retention time: 8.73 min.

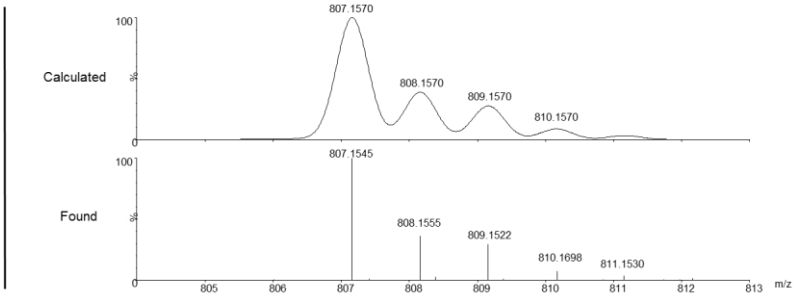
Chemical Formula:
$C_{28}H_{38}N_8O_{12}S_4$

Exact Mass:
806.1492

[M+H]$^-$



[M+H]$^+$

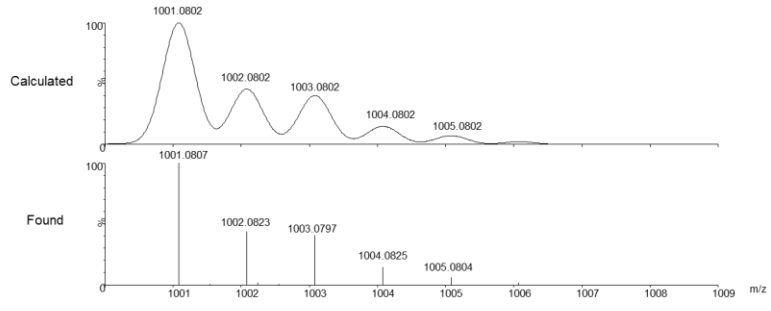


[**2-3e$_3$**]
Retention time: 10.97 min.

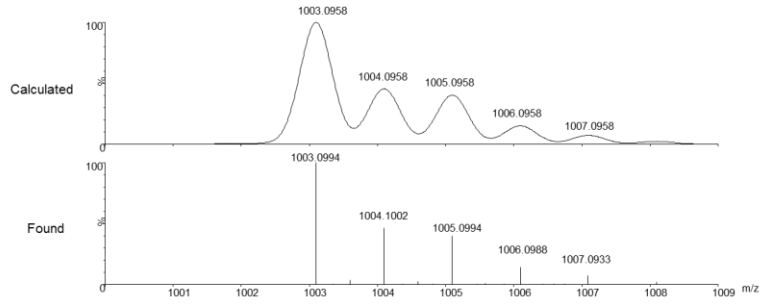
Chemical Formula:
$C_{33}H_{42}N_6O_{18}S_6$

Exact Mass:
1002.0880

[M+H]$^-$



[M+H]$^+$

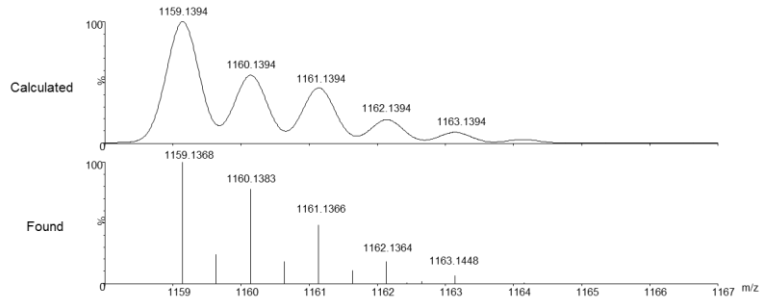


[**1a-2-3e$_3$**]
Retention time: 13.97 min.

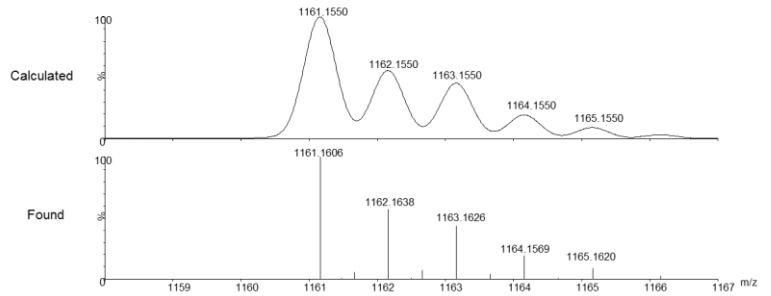
Chemical Formula:
$C_{41}H_{48}N_{10}O_{18}S_6$

Exact Mass:
1160.1472

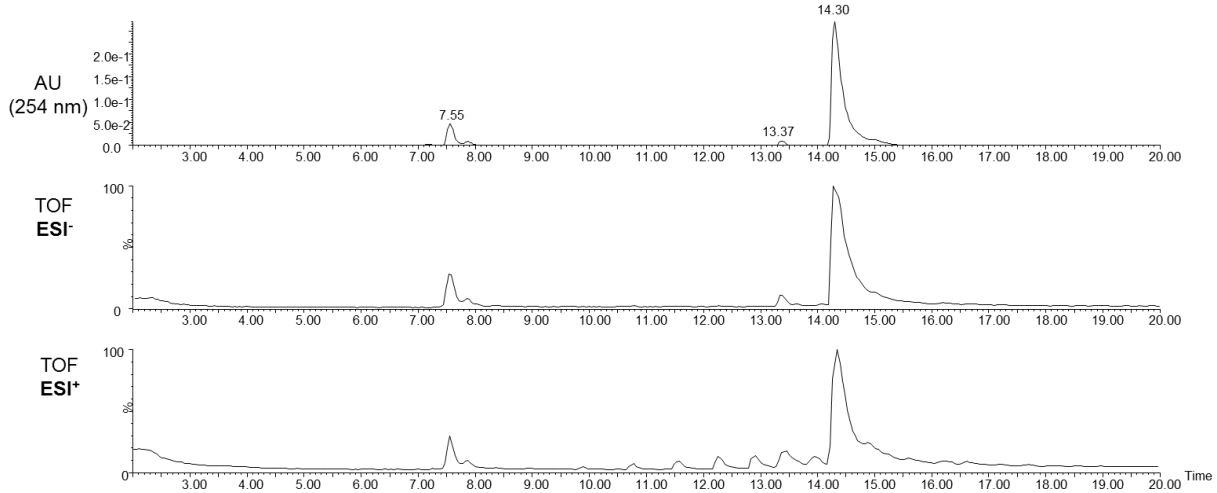
[M+H]$^-$



[M+H]$^+$



Mixture of **1a**+**2**+**3f** (0.5, 0.5 and 2.5 mM respectively, pH.6.5)

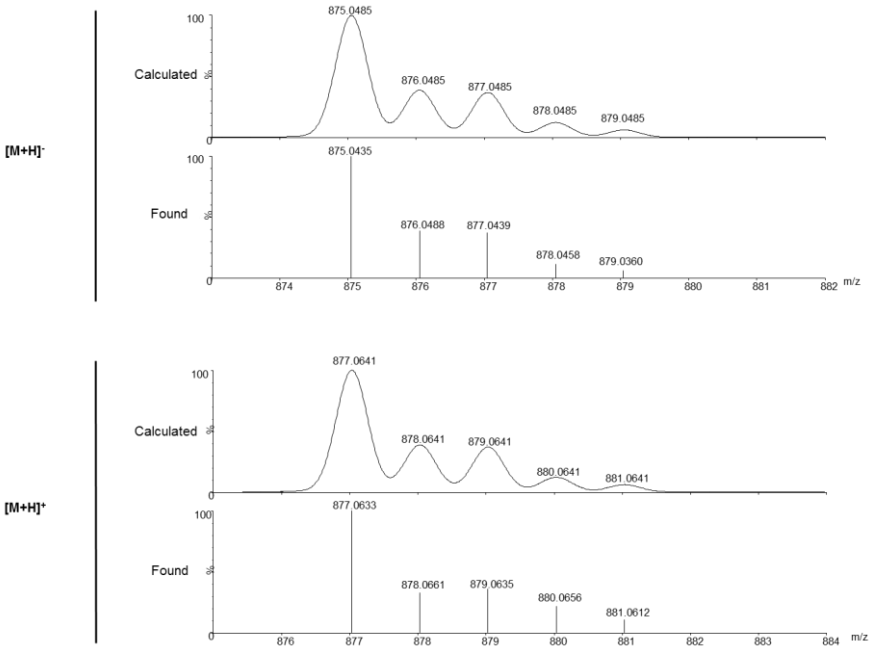


Identification of the products:

[**2-3f$_3$**]
Retention time: 7.55 min.

Chemical Formula:
$C_{27}H_{36}N_6O_{15}S_6$

Exact Mass:
876.0563



[**1a$_2$**] (previously identified)
Retention time: 13.37 min.

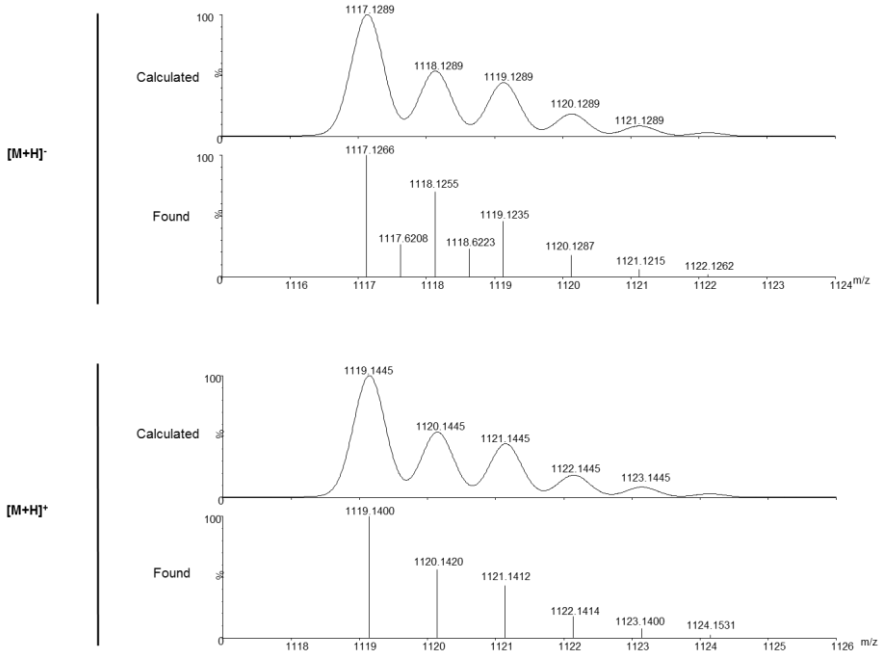
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$

Exact Mass:
964.2084

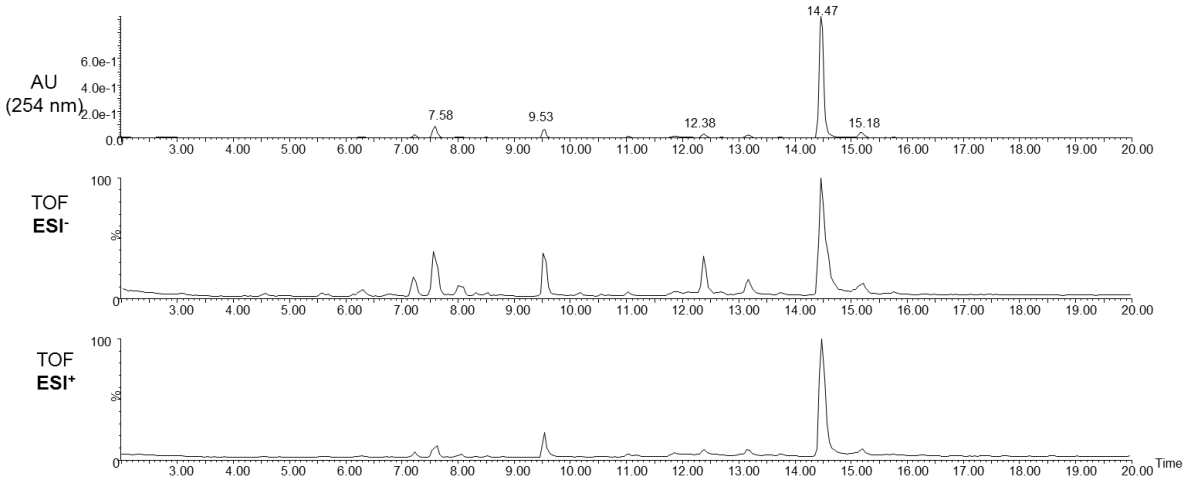
[**1a-2-3f**]
Retention time: 14.30 min.

Chemical Formula:
$C_{39}H_{46}N_{10}O_{17}S_6$

Exact Mass:
1118.1367



<u>Mixture of **1a**+**2**+**3f** (0.5, 0.5 and 2.5 mM respectively, pH.6.5)</u>



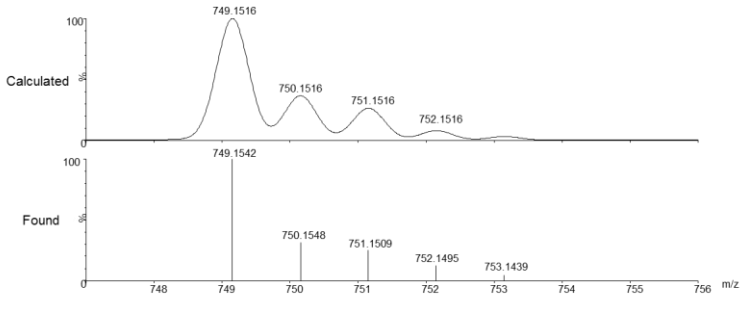
Identification of the products:

[**1a-3g$_2$**]
Retention time: 7.58 min.

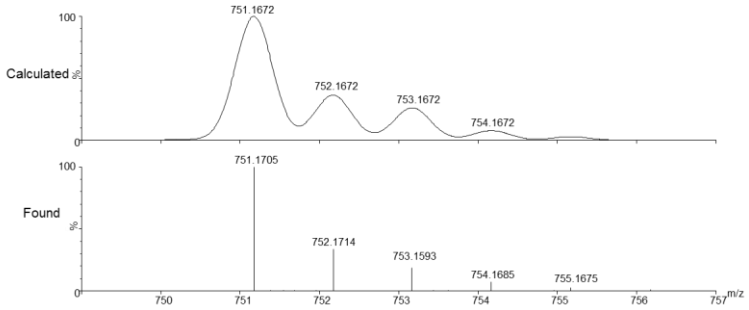
Chemical Formula:
$C_{26}H_{38}N_8O_{10}S_4$

Exact Mass:
750.1594

[M+H]$^-$



[M+H]$^+$

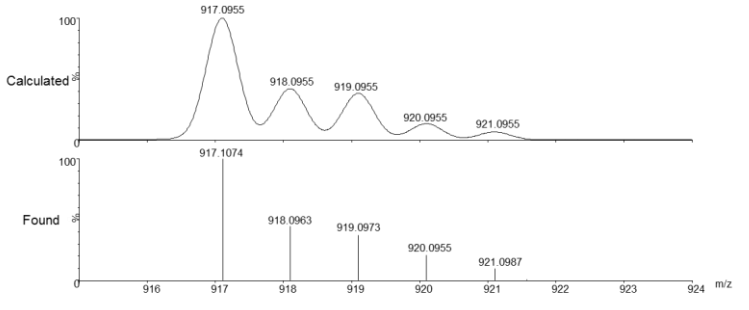


[**2-3g$_3$**]
Retention time: 9.53 min.

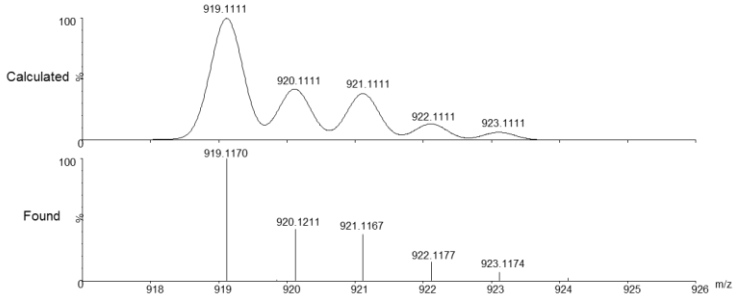
Chemical Formula:
$C_{30}H_{42}N_6O_{15}S_6$

Exact Mass:
918.1032

[M+H]$^-$



[M+H]$^+$



[**2_2**] (previously identified)
Retention time: 12.38 min.

Chemical Formula:
$C_{36}H_{36}N_6O_{18}S_6$

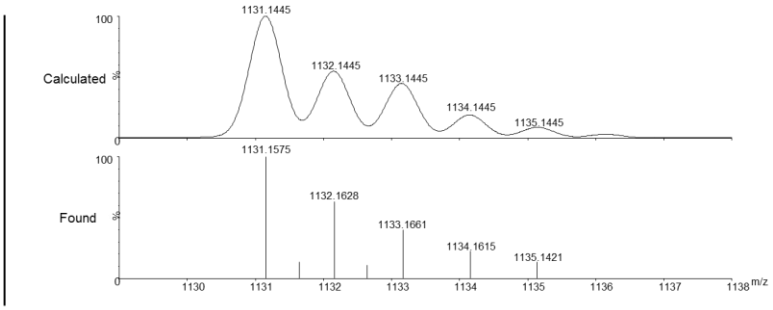
Exact Mass:
1032.0410

[1a-2-3g]
Retention time: 14.47 min.

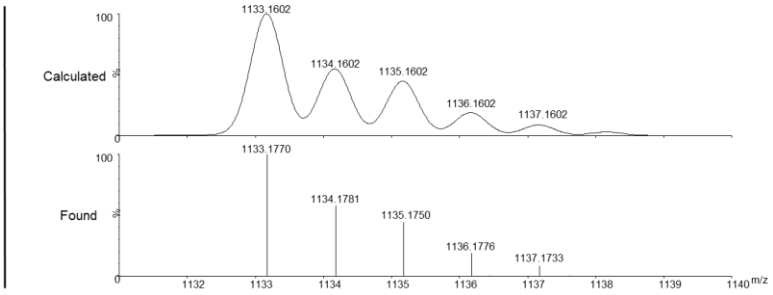
Chemical Formula:
$C_{40}H_{48}N_{10}O_{17}S_6$

Exact Mass:
1132.1523

[M+H]⁻



[M+H]⁺

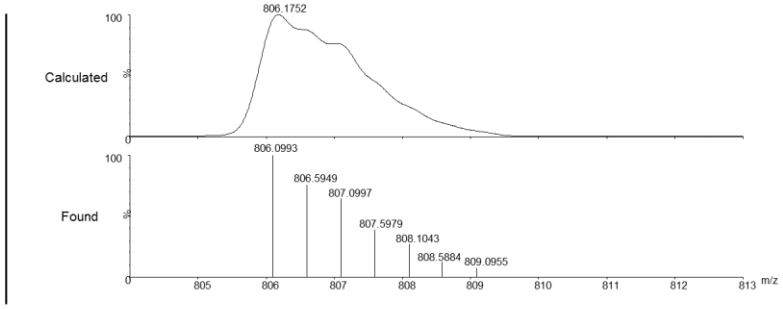


[**$1a_2$-2-3g**]
Retention time: 15.18 min.

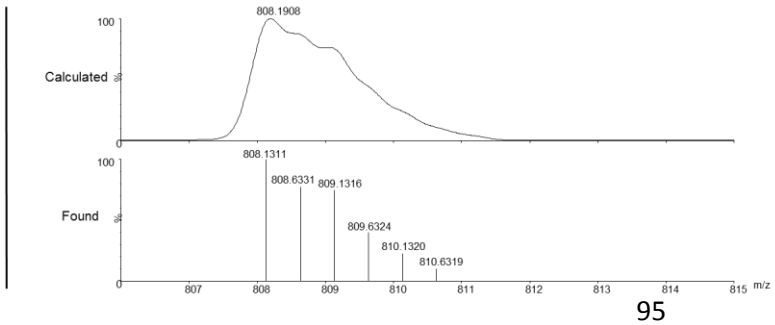
Chemical Formula:
$C_{58}H_{70}N_{16}O_{23}S_8$

Exact Mass:
1614.2565

[M+H]²⁻

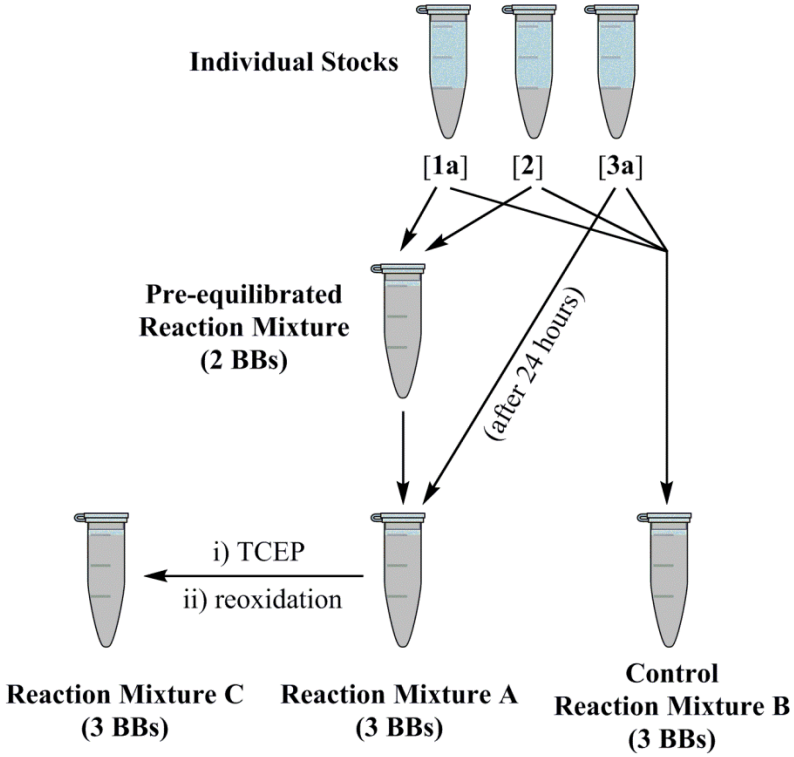


[M+H]²⁺



REVERSIBILITY TEST

Individual stocks of [**1a**, **2, 3a**] were prepared (4 mM in DMSO for **1a** and **2**, and 20 mM in milli-Q water for **3a**) (see scheme S1). From these, a **pre-equilibrated reaction mixture** was prepared by adding 40 µL of [**1a**] and [**2**] to 200 µL of a 66.7 mM BIS-Tris methane buffer (pH 6.5). The **individual stock** of [**3a**] was stored at -80 ºC. After 24 hours, the **reaction mixture A** was prepared by adding 30 µL of the **individual stock** of [**3a**] to 210 µL of the **pre-equilibrated reaction mixture**.



Scheme S1: Preparation of the reaction mixtures of the reversibility test.

Simultaneously, the **control reaction mixture B** was prepared by mixing 30 µL of each **individual stock** with 150 µL of a 66.7 mM BIS-Tris methane buffer (pH 6.5). After 24 hours, the **reaction mixture A**, the **control reaction mixture B** and the **pre-equilibrated reaction mixture** were analysed by HPLC (see figure S16).

Finally, the **reaction mixture C** was prepared by adding 0.35 equivalents of Tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl)[3] 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, 24 hours later, the **reaction mixture C** was analysed by HPLC (see figure S16).

Related references:

3. J. A. Burns, J. C. Butler, J. Moran, and G. M. Whitesides, *J. Org. Chem.*, **1991**, *56*, 2648-2650.

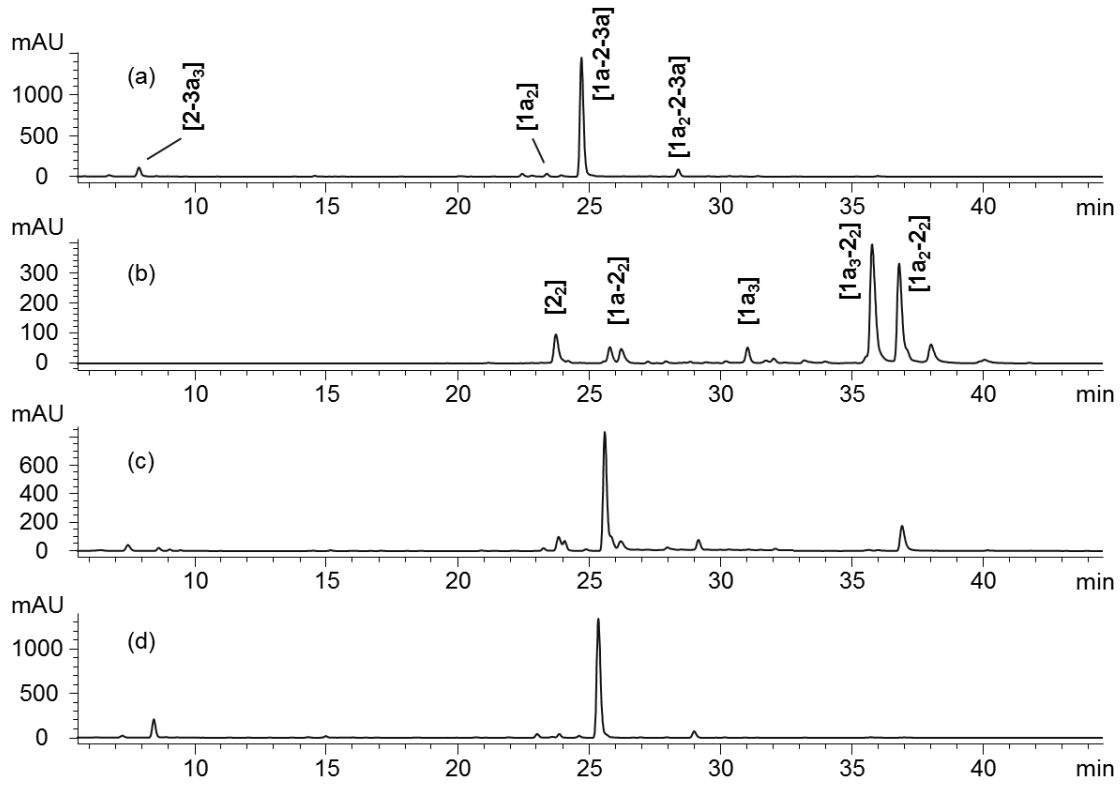
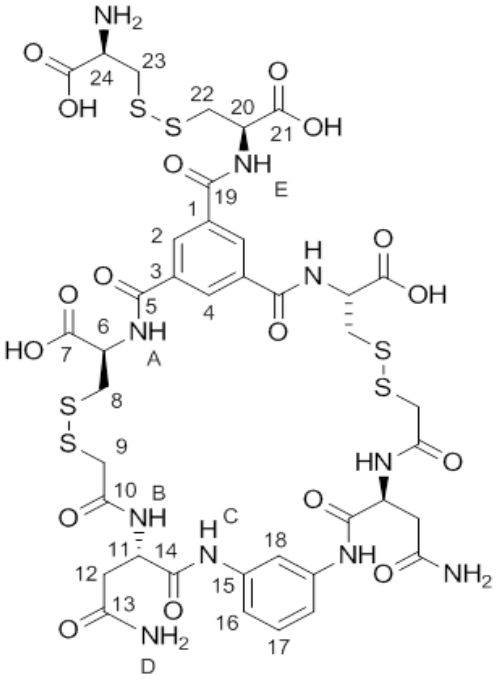


Figure S16: HPLC traces of the **control reaction mixture B** (a), the **pre-equilibrated reaction mixture** (b), the **reaction mixture A** (c) and the **reaction mixture C** (d).

SYNTHESIS AND NMR CHARACTERIZATION OF **1a23a** (**Ia**)



For the synthesis of the trimer **1a23a** (**Ia**), [**1a**] (4.8 mg, 0.010 mmol), [**2**] (5.2 mg, 0.010 mmol) and [**3a**] (2.4 mg, 0.20 mmol) were dissolved in 2 mL of MeOH and 1 mL of milli-Q water. Then, the pH of the solution was adjusted to 7.0 by the addition of NaOH (aq). After 6 days at room temperature, complete oxidation of the starting reagents was observed by HPLC. The MeOH was evaporated *in vacuo* and the residue was lyophilized. The resulting white solid was purified by reversed-face flash chromatography (gradient: from 5% to 15% CH_3CN in H_2O) and 3.8 mg of pure [**Ia**] (34% yield) were obtained as a white solid. HRMS (ESI-) calcd. for $C_{39}H_{46}N_{10}O_{17}S_6$ $[M-H]^-$ (m/z): 1117.1289, found: 1117.1323. ^{1}H NMR (500 MHz, H_2O:DMSO-d_6 (85:15)): δ = 8.69 (s, 2H, H_C), 8.62 (s, 1H, H_4), 8.54 (d, J = 8.4 Hz, 2H, H_B), 8.46 (d, J = 8.8 Hz, 2H, H_A), 8.38 (d, J = 8.2 Hz, 1H, H_E), 8.29 (d, J = 1.7 Hz, 2H, H_2), 7.58 (s, 2H, H_D), 7.20 (d, J = 1.9 Hz, 1H, H_{18}), 6.94 (t, J = 8.2 Hz, 1H, H_{17}), 6.89 (d, J = 8.2 Hz, 2H, H_{16}), 6.76 (s, 2H, $H_{D'}$), 4.05 – 4.00 (m, 1H, H_{24}), 4.01 (d, J = 14.8 Hz, 2H, H_9), 3.47 (dd, J = 15.2 Hz, J = 2.4 Hz, 1H, H_8), 3.45 (d, J = 2.2 Hz, 1H, $H_{9'}$), 3.35 (dd, J = 9.9 Hz, J = 3.9 Hz, 1H, H_{22}), 3.32 (dd, J = 10.7 Hz, J = 4.3 Hz, 1H, H_{23}), 3.19 (dd, J = 15.2, 11.0 Hz, 1H, H_8), 3.04 (dd, J = 15.2, 9.3 Hz, 1H, $H_{22'}$), 2.99 (dd, J = 15.8, 9.9 Hz, 1H, $H_{23'}$), 2.93 – 2.82 (m, 2H, H_{12}) ppm. ^{13}C NMR (125 MHz, D_2O:DMSO-d_6 (85:15)): δ = 174.6 (CO, C_{13}), 172.1 (CO, C_{10}), 169.6 (CO, C_{14}), 167.2 (CO, C_{19}), 166.8 (CO, C_5), 137.0 (C, C_{15}), 134.0 (C, C3), 133.4 (C, C1), 129.7 (CH, C_4), 129.65 (CH, C_{16}), 129.60 (CH, C_2), 116.3 (CH, C_{17}), 111.5 (CH, C_{18}), 55.8 (CH, C_6), 54.4 (CH, C_{20}), 51.0 (CH, C_{11}), 53.4 (CH, C_{24}), 45.0 (CH_2, C_8), 41.2 (CH_2, C_9), 39.5 (CH_2, C_{22}), 37.9 (CH_2, C_{23}), 35.9 (CH_2, C_{12}) ppm.

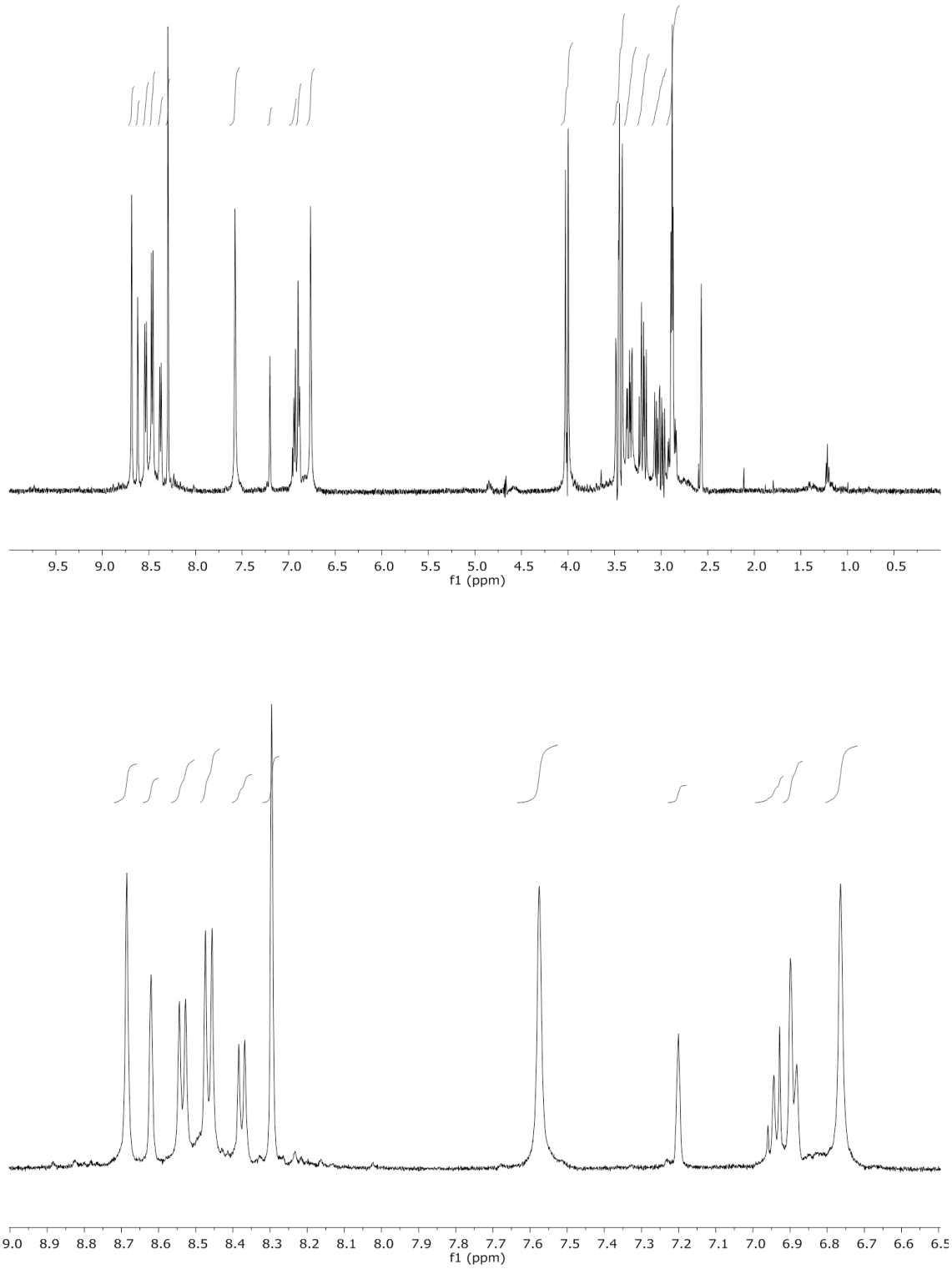


Figure S17: ^{1}H (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectrum of **Ia** (8.5 mM phosphate buffer, pH 6.5), and expansion of the amide region (9.0 – 6.5 ppm).

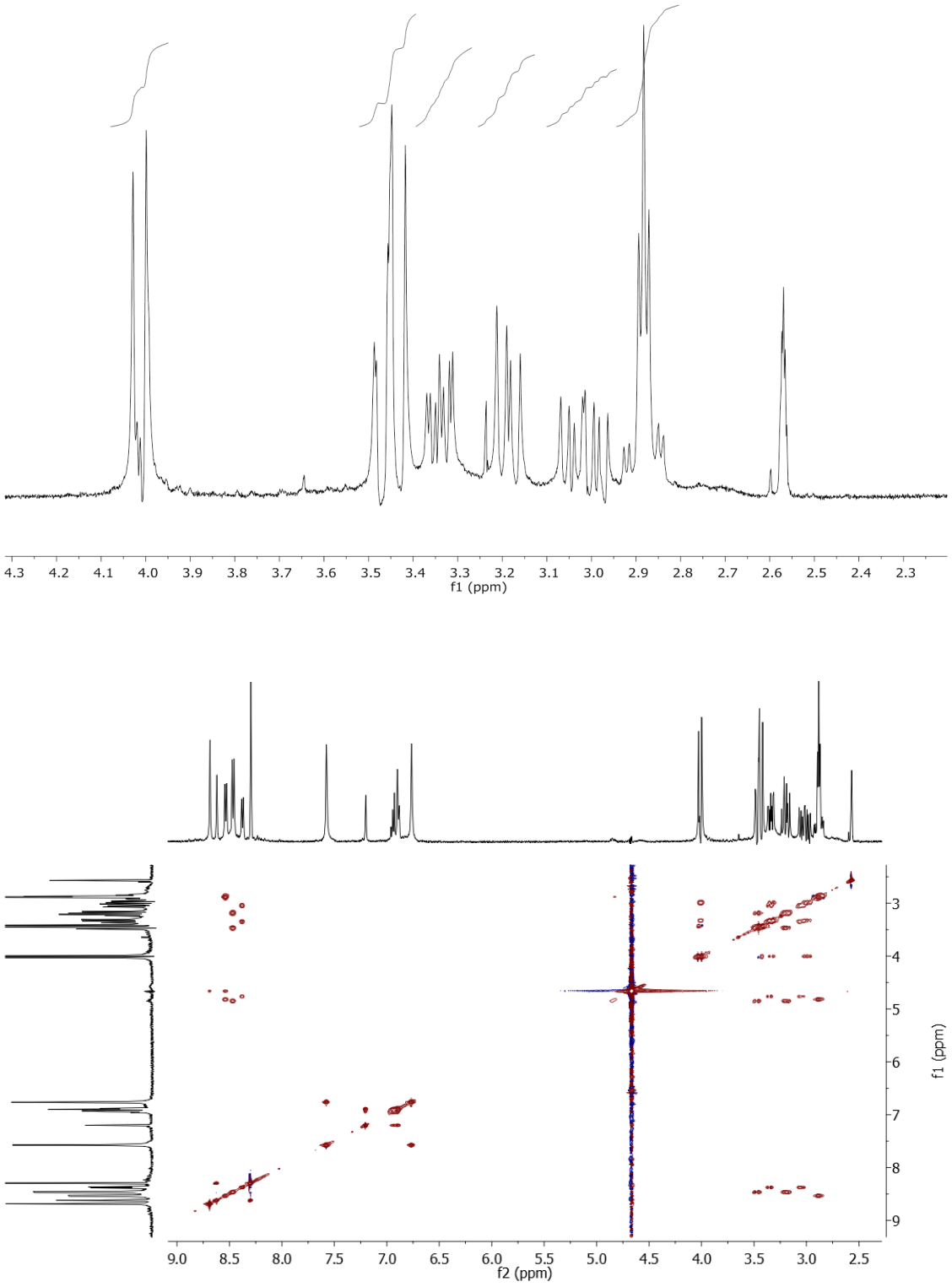


Figure S18: Expansion of the aliphatic region (4.3 – 2.3 ppm) of ^{1}H (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) and ^{1}H 2D TOCSY (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectra of **Ia** (8.5 mM phosphate buffer, pH 6.5).

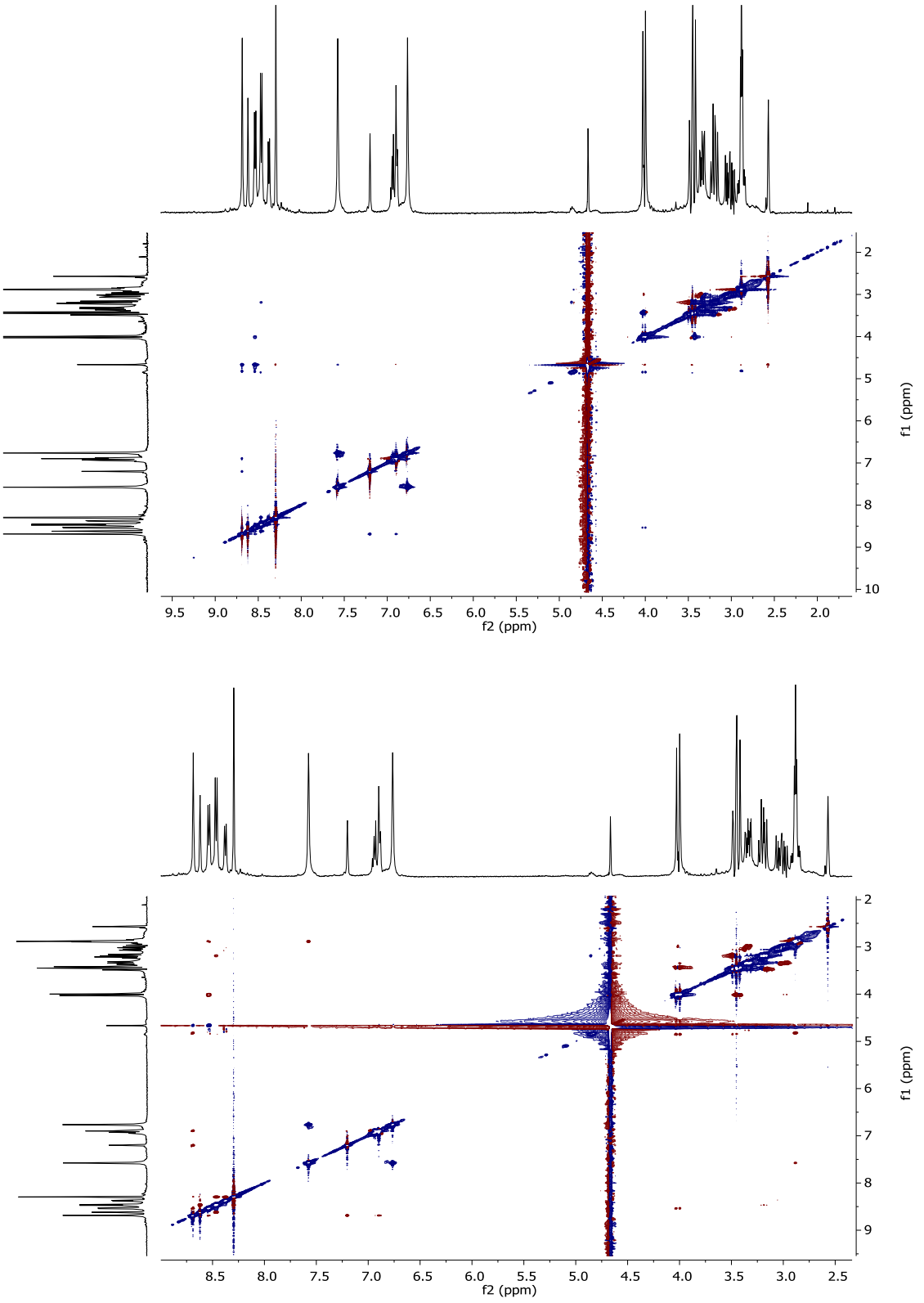


Figure S19: ^{1}H NOESY (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) and ^{1}H ROESY (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectra of **Ia** (8.5 mM phosphate buffer, pH 6.5).

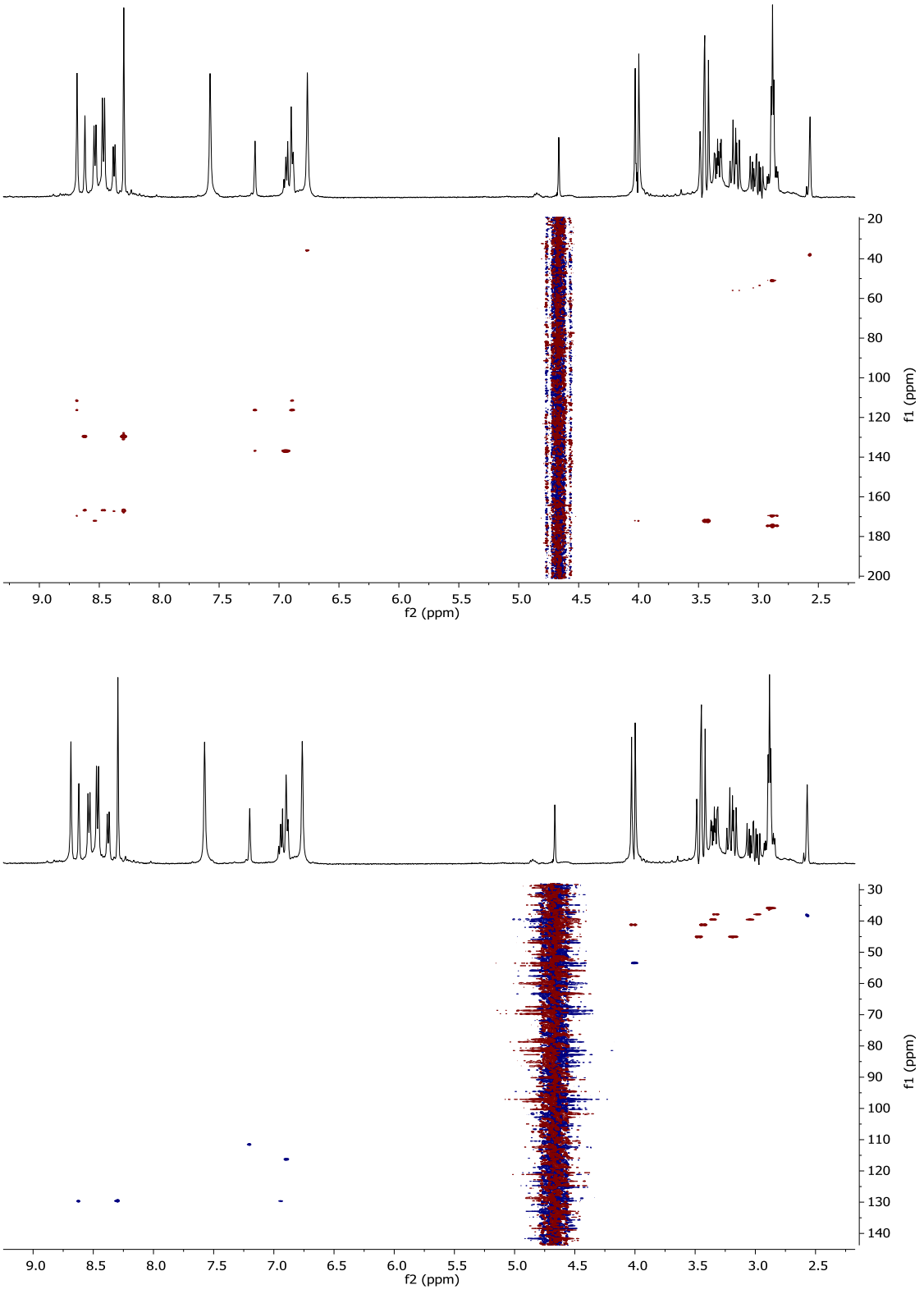


Figure S20: $^{1}H/^{13}C$ gHMBC (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) and $^{1}H/^{13}C$ gHSQC (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) of **Ia** (8.5 mM phosphate buffer, pH 6.5).

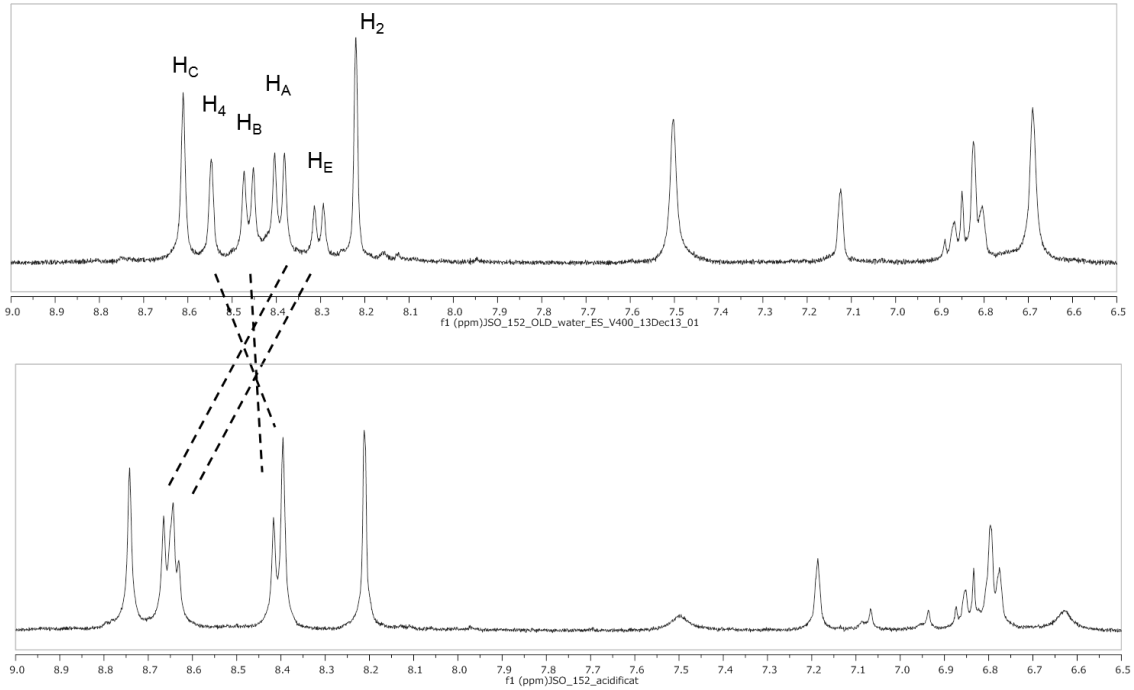


Figure S21: Expansion of the amide region (9.0 – 6.5 ppm) of ^{1}H (400 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectrum of **Ia**. Top: 8.5 mM phosphate buffer, pH 6.5. Bottom: after acidification with TFA, pH ~2. The differences in chemical shift for protons H_A, H_E (moving downfield) and H_4 (moving upfield) are consistent with an unfolding of the proposed conformation for **1a23a**.

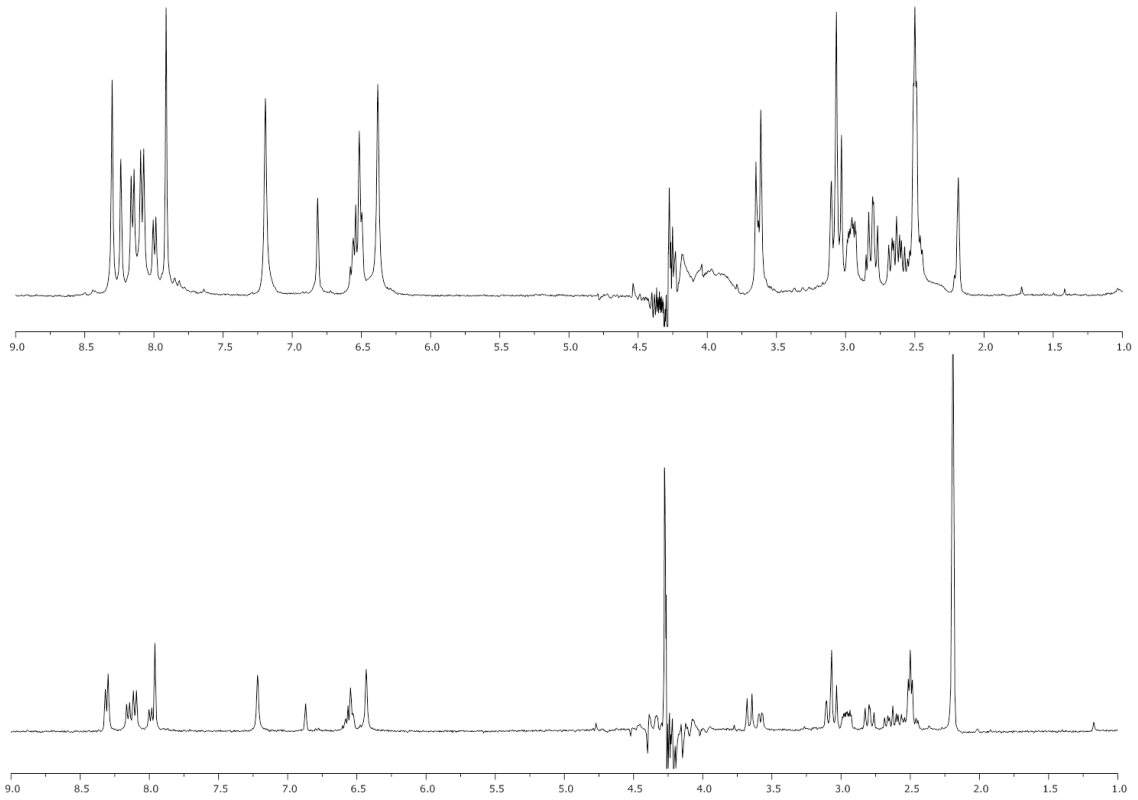


Figure S22: ^{1}H-NMR (400 MHz, 298K) with water suppression by excitation sculpting for **1a23a**. Top: H_2O:DMSO-d_6 (85:15, 8.5 mM phosphate buffer, pH 6.5). Bottom: H_2O:DMSO-d_6 (75:25, 8.5 mM phosphate buffer, pH 6.5). Subtle changes are appreciated, in a more organic solvent the folded structure of is more favored as can be seen by downfield shift for protons H_4 and H_2.

MOLECULAR MODELLING OF **1a23a** (**Ia**)

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. This force field has proved to be the most suitable for the conformational analysis of pseudopeptide and peptoid molecules.[4] The obtained local minima were ordered following the corresponding MMFFaq energies. The process was repeated several times starting 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. We considered all the carboxylic groups as carboxylate anions and the amine function of the pendant cysteine as the ammonium cation. The global minimum for the **1a23a** (**Ia**) compund is shown in Figure 2 of the manuscript. The superposition of the energetically accessible local minima found for this molecule is shown in figure S23.

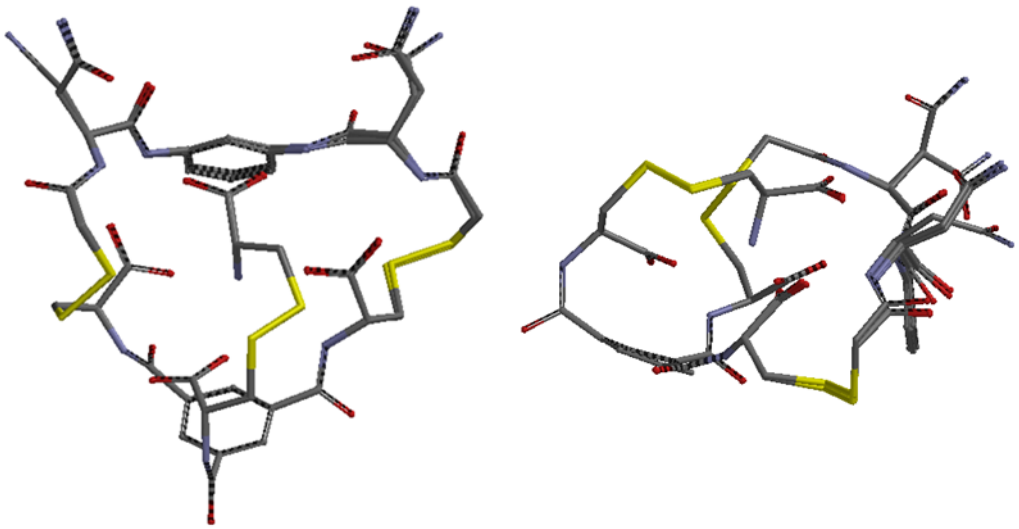


Figure S23: Top and side views of the superposition of the accessible local minima for **1a23a** (**Ia**). Hydrogen atoms have been omitted for clarity.

Related references:

4. (a) W. Brandt, T. Herberg, L. Wessjohann, *Biopolymers (Protein Science)*, **2011**, *96*, 651-667; (b) C. F. Rodriquez, G. Orlova, Y. Guo, X. Li, C.-K. Siu, A. C. Hopkinson, K. W. M.Siu, *J. Phys. Chem. B* **2006,** *110,* 7528-7537; (c) E. F. Strittmatter, E. R. Williams, *J. Phys. Chem. A* **2000,** *104,* 6069-6076; (d) M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, R. A. Friesner, *J. Am. Chem. Soc.* **1997,** *119,* 5908-5920.

2.4. PUBLICATION B:

Adaptive Correction from Virtually Complex Dynamic Libraries: The role of noncovalent interactions in structural selection and folding.

Chem. Eur. J. **2015**, *21*, 17002-17009

Publiation Date (online): 8 October 2015

Dynamic Combinatorial Chemistry

Adaptive Correction from Virtually Complex Dynamic Libraries: The Role of Noncovalent Interactions in Structural Selection and Folding

Maria Lafuente, Joan Atcher, Jordi Solà,* and Ignacio Alfonso*[a]

Abstract: The hierarchical self-assembling of complex molecular systems is dictated by the chemical and structural information stored in their components. This information can be expressed through an adaptive process that determines the structurally fittest assembly under given environmental conditions. We have set up complex disulfide-based dynamic covalent libraries of chemically and topologically diverse pseudopeptidic compounds. We show how the reaction evolves from very complex mixtures at short reaction times to the almost exclusive formation of a major compound, through the establishment of intramolecular noncovalent interactions. Our experiments demonstrate that the systems evolve through error-check and error-correction processes. The nature of these interactions, the importance of the folding and the effects of the environment are also discussed.

Introduction

Living organisms present a vast variety of complex chemical structures such as membranes, proteins and nucleic acids to name a few.[1] These entities are usually self-assembled from simpler parts through the cooperative action of noncovalent interactions,[2] such as ion pairing, hydrogen bonding or solvophobic effects, bringing local order to intrinsically complex systems.[3] The molecular recognition between counterparts is therefore responsible for the self-organization processes. These high levels of sophistication have long fascinated chemists that took Nature as a source of inspiration. In particular, supramolecular chemists have tried to develop artificial molecular structures that somehow mimic the structure and/or function of the naturally occurring systems.[4] However, due to the complexity of living organisms, it is almost impossible to design such assemblies from scratch. Although powerful new computational tools have rendered some impressive results,[5] alternative strategies have emerged to cut corners in a chase where Nature has several millions of years lead thanks to the evolutionary selection processes.

A better understanding of the factors that guide the assembly of simple molecules into bigger structures can give some clues about similar events that occur in biological systems.[6] In the last years, among other strategies, constitutional dynamic chemistry (CDC)[7] and dynamic combinatorial chemistry (DCC)[8] have become popular methods for the synthesis of self-assembled structures, usually under thermodynamic control. Accordingly, the distribution of products can be altered by the presence of external stimuli, such as ligand, so DCC has found a broad application in the synthesis of new molecular receptors.[9] Moreover, because the networks are intrinsically complex and dynamic, new properties may arise from the mixture. Consequently, *systems chemistry* has rapidly developed as a new area of study.[10] Systems chemistry studies the complexity and the emergence of new properties in chemical systems. The careful study of dynamic combinatorial libraries (DCLs) provides an excellent model to help chemists understand what factors govern the formation of a given structure.

In aqueous media, disulfide exchange provides an excellent tool for the generation of dynamic libraries.[11] This methodology has rendered some impressive results, with the formation of some remarkable receptors for biologically relevant structures such as carbohydrates,[12] anions,[13] and DNA G-quadruplexes.[14] In addition, disulfide chemistry has shown its potential in the formation of complex linked structures such as catenanes[15] and knots.[16]

An additional source of structural diversity can arise by combining building blocks (BBs) with different numbers of thiol functionalities. Most of the systems studied to date present BBs that are complementary in terms of valence, with dithiols being the most common.[17] These molecules form more or less complex macrocycles that can always, in principle, close a link. The introduction of multifunctional components leads to more complex libraries in terms of topology and could expand the tridimensional space. However, few examples have been reported that combine molecules with different valence. The formation of organic cages by the combination of tripodal and bipodal components has been described[18] as well as the effect of different monothiols upon a central tripodal moiety.[19] In ad-

[a] M. Lafuente, J. Atcher, Dr. J. Solà, Dr. I. Alfonso
Department of Biological Chemistry and Molecular Modelling
Institute of Advanced Chemistry of Catalonia (IQAC-CSIC)
Jordi Girona 18–26, 08034, Barcelona (Spain)
E-mail: jordi.sola@iqac.csic.es
ignacio.alfonso@iqac.csic.es

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201501415.

dition, environmental effects such as pH or ionic strength have seldom been explored by scientists working on DCC. This can probably be explained by the precise conditions of pH that are required to generate libraries; for example, slightly basic pH for DCL based on disulfide chemistry. Nevertheless some examples describing the variation of a DCL in response to ionic strength have been described. Sanders' group used increasing amounts of $NaNO_3$ to favor the formation of a [2]-catenane driven by hydrophobic effects.[15c] More recently, our group reported the effects of salt concentration on libraries of macrocyclic pseudopeptides.[20] Finally, the effects of ion templation and solvent polarity have been investigated in the parallel or antiparallel aggregation patterns of gramicidin A-like peptides.[21]

We were prompted to study the effects of the media on potentially complex libraries. Moreover, we wanted to analyze the intricate molecular rearrangements that occur during the oxidation and equilibration processes. In a previous study, we investigated the formation of libraries that arise by combining building blocks of different valence.[22] Thus, a series of DCLs formed by the combination of a monopodal, bipodal and tripodal BBs were studied. Now we extend our investigations to put the focus on the variables that drive the system to its final composition. Our findings highlight the delicate conditions that produce effective recognition between counterparts.

Results and Discussion

The molecular components of the different libraries employed in this study are easily accessible by means of standard peptide synthesis protocols. Thus, tripodal building blocks **1a** and **1b** can be obtained from the corresponding triacids in just two steps by using the methodology described previously.[18a] On the other hand, dithiols **2** can be obtained in four steps from phenylenediamine (see the Supporting Information for more details). The monothiols **3** used for this study are all commercially available. These BBs contain diverse structural and chemical information; they also present a different number of thiol functionalities to generate compounds with topological variety (Figure 1).

The adaptive composition of topologically diverse libraries: The effect of pH on the oxidation rate in water/dimethyl sulfoxide (DMSO) mixtures was studied previously.[23] By using 10–25% DMSO in water we are able to ensure complete solubility of BBs, products and intermediates. Moreover, the use of DMSO as minor co-solvent accelerates both oxidation and disulfide exchange, allowing us to perform the reactions with short reaction times (24–48 h) even in slightly acidic media. When oxidising the BBs alone (0.5–2.0 mM) at pH 6.5 (aqueous 50 mM bis-tris buffer containing 10–25% DMSO) they mainly produced the corresponding homodimers (some trimers can be observed as minor compounds for bipodal compounds **2**). However, when BBs of different valence are mixed, very distinct product distributions can be seen, depending on the relative stability of the components. Thus, a mixture of tripodal **1a** and cysteine **3a** (1:5) mainly produced the 'linear' [**1a**–**3a**$_3$] hetero-oligomer, by combination of the two BBs. When bipodal **2a** and **3a** were mixed, the major final compounds were the homodimers (although cysteine was not visible in the UV-HPLC traces). Thus, in one case, the system efficiently combined the BBs whereas in the second, segregation of the linear and cyclic components was observed.

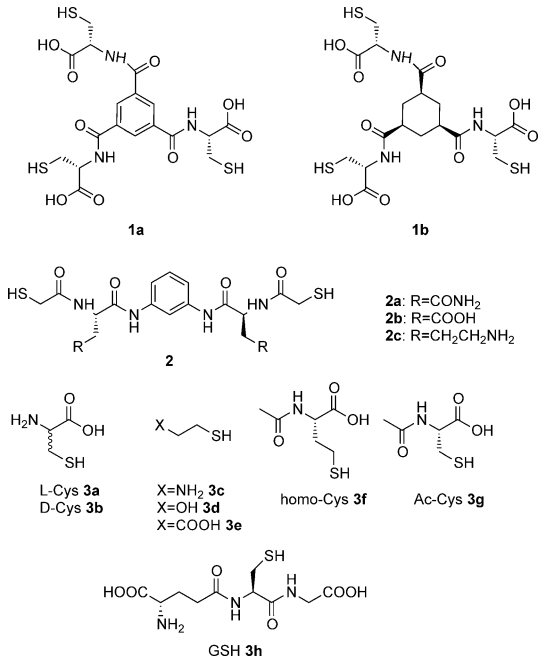


Figure 1. Constitutional building blocks used for the DCL studies.

We monitored the evolution of the two libraries over time. In the first case we observed a rapid formation of large oligomers [**1a**$_2$–**3a**$_4$], major species at short reaction times, [**1a**–**3a**$_3$] and [**1a**$_2$–**3a**$_2$], and the homodimer [**1a**$_2$] (see Figure S9 and S10). After 24 h, almost all the tripodal BBs form [**1a**–**3a**$_3$] with some homodimer [**1a**$_2$] present (Figure 2a). For the binary mixture of asparagine derivative **2a** and **3a**, at short reaction times, linear oligomers [**2a3a**$_2$], [**2a**$_2$**3a**$_2$], [**2a**$_3$**3a**$_2$], homodimer [**2a**$_2$] and oxidized monomer [**2a**] can be detected (Figure S12 and S13). This mixture changed rapidly, growing in proportion of [**2a**$_2$], which is the major compound at the end of the reaction, and the appearance of trimer [**2a**$_3$], not present at short reaction times. [**2a3a**$_2$] is also detected at the final composition, whereas [**2a**$_2$**3a**$_2$] almost disappeared. Larger oligomers rearrange and cannot be detected at the end of the reaction (Figure 2b right). Thus, in both cases, the final composition is a consequence of the dynamic rearrangement of initially very complex mixtures.

The diversity and complexity of the system increased when an equimolar mixture of tripodal **1a** and bipodal **2a** compounds (0.5 mM each) was equilibrated under the conditions described above. At short reaction times, homodimers [**1a**$_2$] and [**2a**$_2$] rapidly appeared together with unsaturated mixed compounds [**1a**–**2a**–SH], homotrimer [**2a**$_3$] and other unidentified oligomers. The mixture rearranged until reaching its final composition: together with homodimers [**1a**$_2$] and [**2a**$_2$], assemblies containing two units of **1a** and one to three units of **2a** in what we presume must be cage-like compounds could

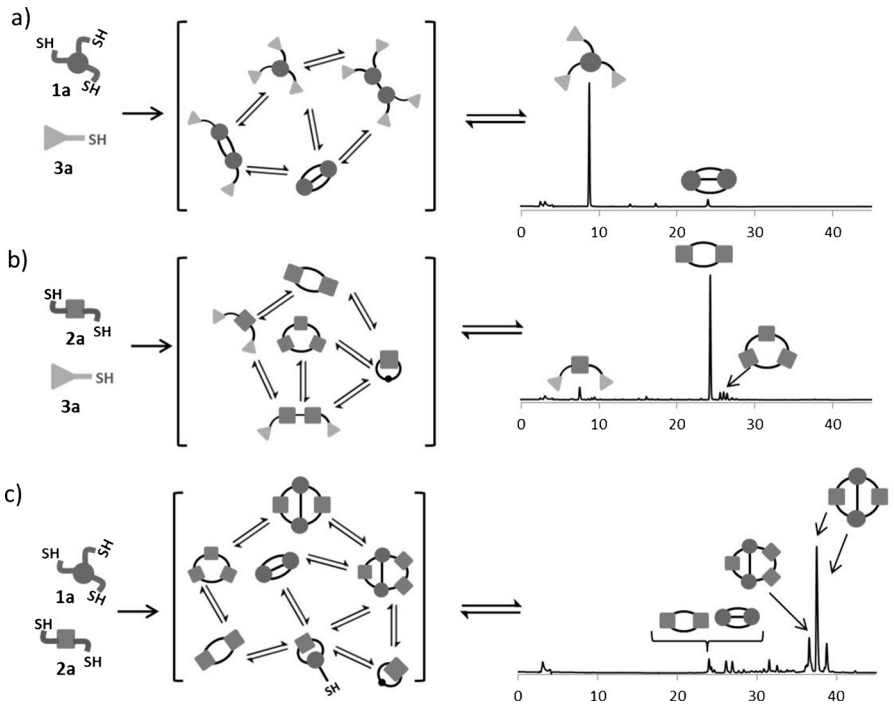


Figure 2. Schematic representation of the binary libraries obtained by the combination of BBs with different valence and HPLC traces at the end of the reaction (24 h): a) **1a**+**3a**, b) **2a**+**3a** and c) **1a**+**2a** (**1a**, **2a** 0.5 mM, **3a** 2.5 mM, bis-tris aqueous buffer at pH 6.5 with 25% DMSO).

be detected (Figure 2c). Interestingly, we observed mass-degenerated species corresponding to different constitutional isomers. Thus, the mixing of different constitutional information is transferred to generate a different library composition.

More remarkably, when the same experiment was performed by adding 2.5 mM cysteine to the mixture (enough to saturate all the library thiols) the final composition was completely different (Figure 3). At the first stages of the reaction, some kinetic compounds can be observed such as oxidized monomer [**2a**], linear oligomers [**1a**–**3a**$_3$] and [**2a**$_2$–**3a**$_2$] (surprisingly as the main product), homodimers [**1a**$_2$], [**2a**$_2$] and unsaturated mixed compounds [**1a**–**2a**–SH]. The mixture slowly rearranges: [**1a**$_2$–**2a**$_2$] transiently formed, the same trend was observed for homodimers, which increased from $t=0$ until 2 h but then started decreasing until they reached their final proportions. Strikingly, a compound containing the three BB compounds [**1a**–**2a**–**3a**] grew in proportion from short reaction times until it became by far the predominant species at the end of the reaction (Figure 3 and Figure S6). The larger oligomer [**1a**–**2a**$_2$–**3a**] also forms during the course of the reaction but in minor quantities.

As control experiments, we also tested ternary mixtures in which either **1a** or **2a** was double the concentration of the other component in a sample containing an excess of cysteine.[24] The heterotrimer was again selected from the mixture, consuming almost the entire limiting reagent, showing again that the selection toward this compound overcame the stoichiometric limitations. The excess reagent was rearranged in its most stable final oxidized compound as seen in the binary mixtures with cysteine: linear oligomer [**1a**–**3a**$_3$] for tripodal **1a** and homodimer [**2a**$_2$] for asparagine derivative **2a** (see the Supporting Information).

Thus, we observed that the chemical and topological diversity of the BBs was expressed in a large initial complexity, which surprisingly evolved toward an extraordinary simplicity. Furthermore, this transformation occurred through an error-check and error-correction iterative process, which is one of the main characteristics of the dynamic covalent chemistry procedures and has been described before.[25] However, to our knowledge, few examples have shown this property using disulfide chemistry,[16] and our experiments nicely present this corrective and adaptive nature of DCC.

Structural characterization of the expressed heterotrimer: We synthesized and characterized the compound [**1a**–**2a**–**3a**] in detail. 1D and 2D ^{1}H NMR (10 or 25% [D$_6$]DMSO, pH 6.5, phosphate buffer) spectroscopic analysis led us to conclude that [**1a**–**2a**–**3a**] is a symmetric struc-

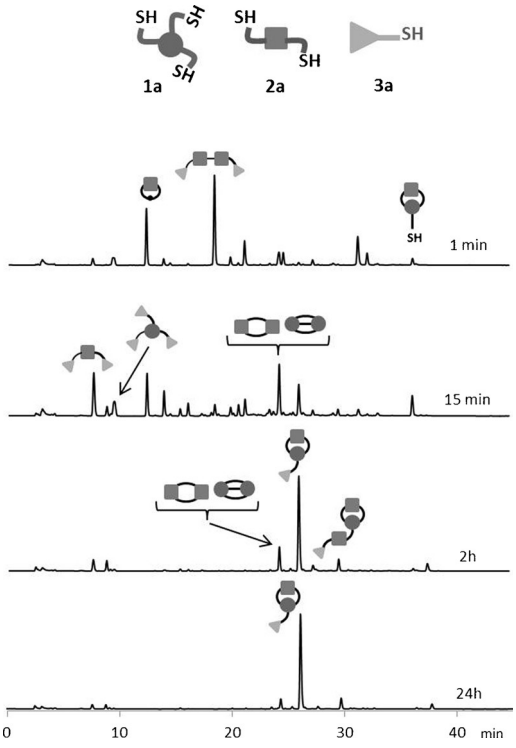


Figure 3. HPLC traces (UV detection at 254 nm) for the evolution of a sample containing **1a**, **2a** and **3a** (**1a**, **2a** 0.5 mM, **3a** 2.5 mM, bis-tris pH 6.5, 25% aqueous DMSO) over time.

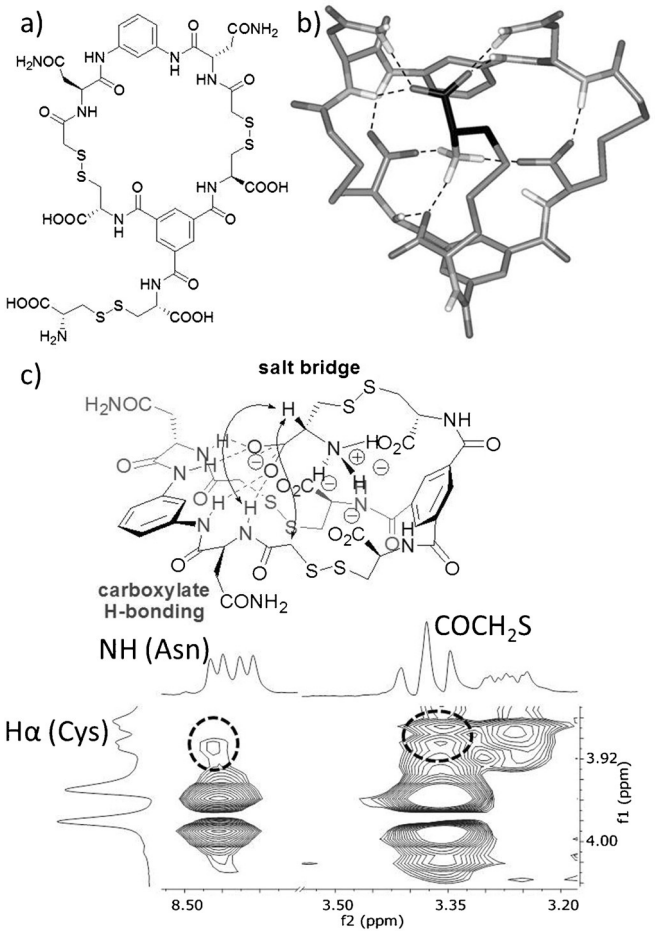


Figure 4. a) Chemical structure of [**1a**–**2a**–**3a**]. b) Folded conformation of [**1a**–**2a**–**3a**] obtained by molecular mechanics calculations. Nonpolar H atoms are omitted and possible H-bonds are highlighted by dashed lines. The C atoms of the pendant cysteine are shown in black. c) Schematic representation of the folded conformation, highlighting the proposed noncovalent interactions, the double headed arrows show the ROESY signals depicted in the selected regions of the ROESY spectrum.

ture in which **1a** and **2a** close a macrocycle and the remaining thiol of **1a** is linked to a molecule of cysteine (Figure 4a).

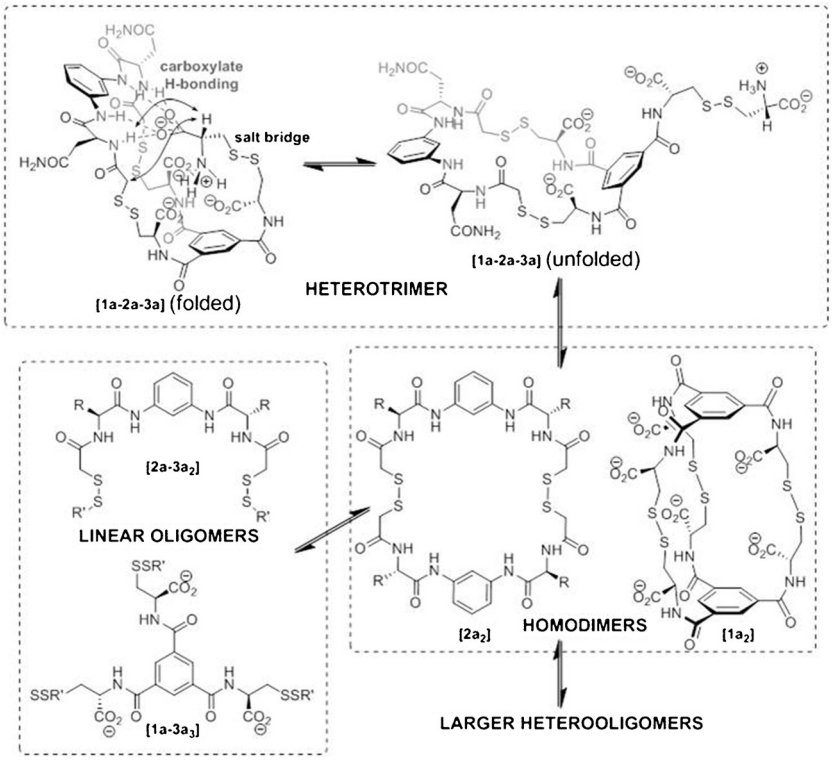
Molecular models suggested a series of interactions within the structure (Figure 4b). Thus, the three carboxylates from the tripodal moiety would be stabilized by the ammonium group of the cysteine (salt bridge interaction), forcing this part to fold into the macrocycle. Further hydrogen bonds of these carboxylates to amide protons would also contribute to the stability of the major compound. Similar hydrogen-bonding interactions would be established by the carboxylate of the pendant cysteine residue. The formation of this folded structure was further evidenced by ^{1}H NMR spectroscopic analysis (1 mM, pH 6.5, 50 mM phosphate buffer, H_2O, 25% [D_6]DMSO). The chemical shifts observed for the amide protons are consistent with the proposed H-bonding pattern. More importantly, the proton attached to the Cα of the pendant cysteine rendered cross-peaks in the ROESY spectrum with both the amide NH from the asparagine moiety and one of the protons of the methylene of the thioglycolic ($COCH_2S$) moiety (Figure 4c). These ROEs indicate the proximity between the pendant cysteine and the bipodal moiety of the macrocyclic structure and, thus, the presence of a folded structure in solution. Moreover, upon acidification of the sample, the position of some signals shifted, in agreement with an unfolding process (see the Supporting Information).

We believe that the [**1a**–**2a**–**3a**] species is highly amplified from the dynamic mixture because of the attractive interactions that are established between the different moieties of the corresponding building blocks, which can only take place in the folded conformation. In other words, a folding-directed selection of a component in a DCL occurs, shifting the equilibrium away from a statistical or stoichiometric distribution.[26] Folding processes together with folding-induced recognition have been shown to shift the equilibria in imine-based DCL.[27] In aqueous media, however, folding usually occurs through hydrophobic effects, not by polar interactions such as in our system.[15c,16,28]

Environmental effects on the library composition: An important consequence of the findings described in the previous sections should be that our topologically diverse dynamic library is able to express the noncovalent interactions established between the BBs, by the selection of an almost unique species from a virtually very complex mixture. We propose a dynamic equilibrium in solution as a consequence of the connection of noncovalent and covalent dynamic bonds (Scheme 1). Given that the major heterotrimer was amplified because of the polar salt bridges and H-bonds established by the pendant cysteine in the folded conformation, the breaking of these interactions in solution would unfold the molecules, destabilizing the structure relative to other potential members of the library.

We studied the effect of environmental parameters such as pH or salt concentration on the final composition of the ternary libraries containing **1a**, **2a** (0.5 mM) and **3a** (2.5 mM). Due to the electrostatic nature of the interactions promoting the formation of [**1a**–**2a**–**3a**], subtle changes in the environment should have a large impact on the library distribution. Accordingly, we repeated the same experiment in aqueous solution at different pH values (10 or 25% DMSO). To cover a broad range with a minimal change in the structure of the buffer, bis-tris propane was used between 6.5 and 9.5, whereas a formate buffer was employed at pH 4.5.[29] The members formed in the reaction were classified into four families, as shown in Scheme 1: 'linear' oligomers with cysteine [**1a**–**3a**$_3$] or [**2a**–**3a**$_2$], homodimers [**1a**$_2$] and [**2a**$_2$], the major heterotrimer [**1a**–**2a**–**3a**] and a larger hetero-oligomer of formula [**1a**–**2a**$_2$–**3a**]. The equilibrium concentrations of the major compounds are displayed in Figure 5. The optimal pH for the recognition is around 6.5. At this value, the carboxylic acids of **1a** are deprotonated and the cysteine moiety is a zwitterion, so the proposed salt-bridge and carboxylate–amide interactions can take place efficiently. Higher or lower pH values led to less efficient formation of the [**1a**–**2a**–**3a**] heterotrimer.

To further prove the importance of polar interactions for the selection of the major compound, we next turned our attention to the effect of ionic strength on our libraries. Thus, we re-



Scheme 1. General equilibrium pattern proposed for the DCLs obtained by mixing BBs with different topology.

almost exclusively. The higher ionic strength shielded the polar interactions, making the selection less effective (Figure 6a). Again, in 10% DMSO the trend was more pronounced and the amount of homodimers even exceeded that of the heterotrimer (Figure 6b) at 2 M NaCl.

Interestingly, the high selectivity toward heterotrimer formation was retained down to 2% DMSO. Lower concentrations of DMSO produced some precipitation, which precluded a reliable analysis of the composition of the libraries. Overall, the results from the analysis of the effects of environmental variables strongly support our model for the selection mechanism and further underscore the adaptation ability of dynamic libraries.

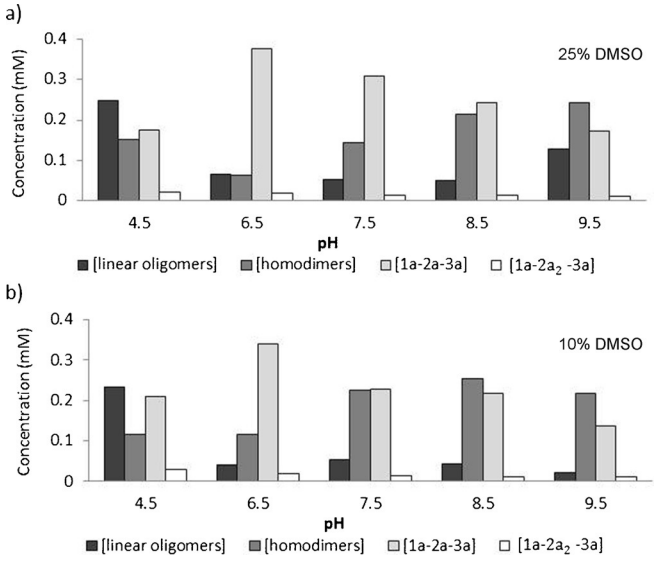


Figure 5. Compound distribution of the DCLs at different pH values: **1a**, **2a** (0.5 mM), **3a** (2.5 mM), aqueous bis-tris propane buffer (50 mM, pH 6.5–9.5) or formate buffer (50 mM, pH 4.5), containing a) 25% or b) 10% DMSO.

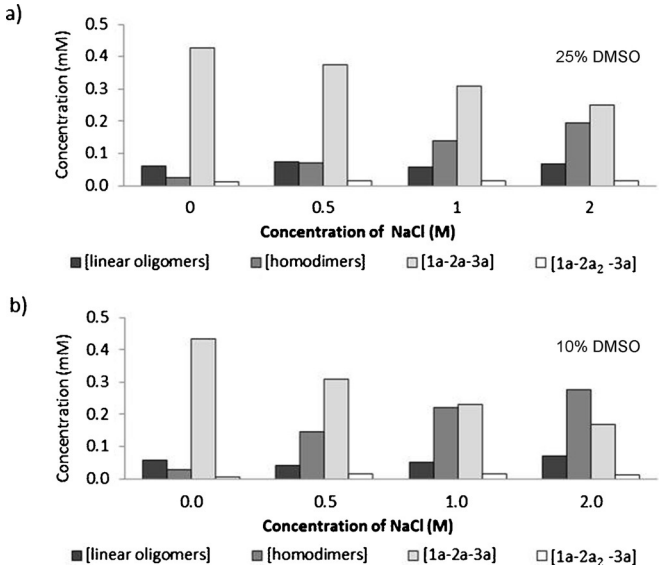


Figure 6. Influence of ionic strength on the composition of the library formed by **1a**, **2a** (0.5 mM) and **3a** (2.5 mM) in aqueous bis-tris buffer (50 mM, pH 6.5), containing a) 25% or b) 10% DMSO.

peated our experiments with increasing amounts of sodium chloride. It has been described that salt concentration can influence the establishment of salt bridges in small peptides.[30] As expected, the formation of the heterotrimer was affected by the amount of salt: at higher ionic strength the selection was lower. This is a further indication of the polar nature of the interactions driving the formation of a single compound

Effect of the structure of the BBs on the library composition: Subtle structural changes can be decisive in the outcome of a DCL.[31] It is expected that suppression of possible interactions or the establishment of new interactions will clearly modify the composition of the library (as in Scheme 1). When changing L-cysteine with D-cysteine, no major changes were observed in the composition of the library. Replacing cysteine

with homocysteine had a little effect on the preferred assembly of the corresponding heterotrimer but the appearance of additional products was also observed. However, a certain degree of structural recognition could be observed when these three amino acids were used in a competition experiment. Thus, L-cysteine was slightly preferred over D-cysteine, and both were preferred over homocysteine (see the Supporting Information).

Systematic variation of the chemical functions on the monothiol component provided further evidence of the nature of the interactions favoring the major compound. Thus, the inclusion of aminoethanethiol **3c** still led to a relatively simple library. The replacement of the amino group by noncharged alcohol functionality (**3d**) or an anionic residue (**3e**) had a detrimental effect on recognition. Accordingly, when acetylated cysteine was used, the system evolved toward a more complex library. From these data, we concluded that both the carboxylate group and the ammonium ion are important for the recognition, with the latter being critical. Moreover, when glutathione **3h**, which also possess a zwitterionic moiety, was used, some larger oligomers could be seen in the LC-traces together with the heterotrimer, thus evidencing that structure complementary is also necessary.

Substitution of the central tripodal building block by a related tricarboxylic acid with a different geometrical disposition was also performed.[32] The use of cyclohexanetricarboxylic derivative **1b** had a major impact on the mixture containing just this moiety and a bipodal structure. Thus, a very complex HPLC-trace was obtained. We were able to identify some of the components of the mixture by LC-MS analysis, with homodimer [$\mathbf{2a}_2$] predominating, and several compounds of formula [$\mathbf{1b}_2$–$\mathbf{2a}$] being detected (Figure 7). Some other components remained unidentified, probably due to the formation of larger oligomers. It is worth mentioning that only products containing **2a** will appear in the UV-HPLC traces, because now the central tripodal structure does not absorb at 254 nm. However, the LC-MS analysis showed a very small amount of homodimer [$\mathbf{1b}_2$], meaning that most of the tripodal moiety **1b** was consumed as oligomers with **2a**. When the reaction was repeated, adding cysteine to the mixture, the library presented a similar distribution of products to those containing **1a** but with a lower selectivity toward the corresponding heterotrimer. Nevertheless the HPLC traces simplify impressively (Figure 7). A compound of formula [**1b**–**2a**–**3a**] formed predominantly, but other products, mainly [**1b**–$\mathbf{2a}_2$–**3a**] and the homodimer of [$\mathbf{2a}_2$] were also observed. A certain amount of [**1b**–$\mathbf{3a}_3$] could also be detected by LC-MS. The effect of the monothiol structure with the aliphatic tripodal building block **1b** was also studied. Not surprisingly, the different DCL prepared gave similar results to those for **1a**, confirming the nature of the interactions between the structure counterparts (see the Supporting Information).

We tested the influence of the presence of charged residues in the bipodal moiety (Figure 8). When mixing **1a** and **2b**, both negatively charged at the reaction pH, the main species were the homodimers of each component, which reduced the concentration of charges of the same sign in every member. When the experiment was carried out in the presence of cysteine, although some heterotrimer [**1a**–**2b**–**3a**] was formed, other species still dominate the HPLC traces. In this library, compound **1a** tends to form the linear tetrameric structure with three cysteines [**1a**–$\mathbf{3a}_3$] although some homodimer could still be observed. On the other hand aspartate derivative

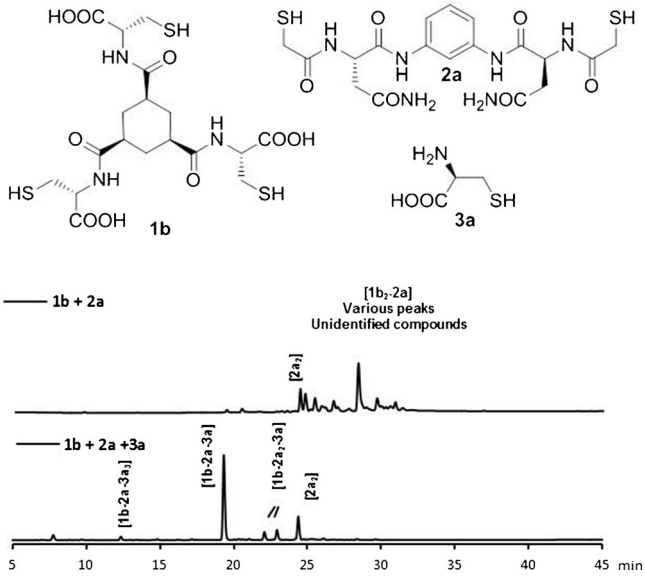


Figure 7. HPLC traces (UV detection at 254 nm) for the DCLs of compounds **1b** and **2a** (0.5 mM) at pH 6.5 in the absence (top) and presence (bottom) of 2.5 mM **3a**.

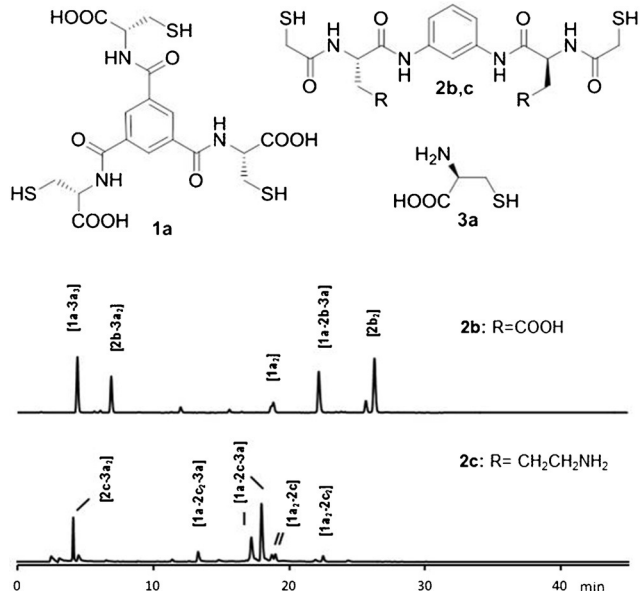


Figure 8. Effect of charged residues on the bipodal BB: HPLC traces (UV detection at 254 nm) of the equilibrium composition for DCLs formed by mixing tripodal (**1a**, 0.5 mM) with monopodal (**3a**, 2.5 mM), and either anionic (**2b**, 0.5 mM, upper trace) or cationic (**2c**, 0.5 mM, lower trace) bipodal BBs.

2b formed the homodimer [**2b**$_2$] or the linear compound with two cysteines [**2b**–**3a**$_2$] with slight preference for the latter.

In the case of the positively charged ornithine derivative **2c**, the library presented a variety of compounds in which again the heterotrimer [**1a**–**2c**–**3a**] was the major product but this was formed in a less effective manner than with **2a**, and two constitutional isomeric structures were observed. Other species including [**1a**$_2$–**2c**], [**1a**$_2$–**2c**$_2$], [**2c**$_2$], [**2c**–**3a**$_2$] and [**1b**–**2c**$_2$–**3a**] were also present (Figure 8).

These experiments clearly highlight the importance of having the correct interactions for the selection of a single geometry. Subtle changes in structure result in profound changes in the nature of the products formed.

Conclusion

We described the dynamic covalent libraries formed by the combination of compounds with different number of functionalities using disulfide chemistry in aqueous media. The systematic study of their time-evolution and equilibrium composition depending on intrinsic (chemical structure and valence of the building blocks) and extrinsic (pH, ionic strength and organic co-solvent) factors rendered important information about noncovalent intramolecular interactions. Surprisingly, a single species was expressed almost exclusively from a library displaying an initially very large complexity. This remarkable assembly is dictated by the cooperative action of weak polar interactions leading to a folded structure, as confirmed by NMR analysis and a battery of control experiments. Thus, the final outcome of the dynamic chemical system is selection through positive recognition and folding, as a result of the delicate interactions between its members, which exchange through a dynamic correction toward the best fitted molecule.

More importantly, our results highlight the power of dynamic covalent chemistry to discover unexpectedly stable assemblies and to unravel the reasons for their formation. In this particular case, a recognition macrocyclic motif for zwitterionic species has been serendipitously identified. This fact further confirms DCC as an extremely useful tool for systems chemistry research. Furthermore, structural insights regarding noncovalent interactions and folding properties of peptide-like molecules in highly competitive media can be extracted from carefully designed DCC assays. The further use of this knowledge should lead to the discovery of new materials for which structural selection will translate into functional selection.

Experimental Section

Full experimental details and full characterization of compound [**1a**–**2a**–**3a**] are available in the Supporting Information. Details of materials, synthesis of building blocks, NMR spectra, identification and MS of library members and HPLCs traces are also available in the Supporting Information.

Acknowledgements

This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO, CTQ2012-38543-C03-03), Generalitat de Catalunya (AGAUR, 2014 SGR 231). J. S. acknowledges MINECO (Ramon y Cajal contract) and EU (FP7-PEOPLE-2012-CIG-321659) for funding fellowship. M. L. thanks MINECO for pre-doctoral fellowship (FPI: BES-2013-063128). J. A. thanks CSIC for financial support (JAE-predoc fellowship). We thank Dr. Yolanda Pérez for helpful assistance with NMR experiments.

Keywords: conformation analysis · molecular evolution · noncovalent interactions · peptidomimetics · self-assembly

[1] H. F. Lodish, A. Berk, C. Kaiser, M. Krieger, A. Bretscher, H. Ploegh, A. Amon and M. P. Scott, *Molecular Cell Biology*, W. H. Freeman and Company, New York, **2012**.

[2] a) J.-M. Lehn, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 1304 – 1319; *Angew. Chem.* **1990**, *102*, 1347 – 1362; b) G. M. Whitesides, B. Grzybowski, *Science* **2002**, *295*, 2418 – 2421; c) J.-M. Lehn, *Angew. Chem. Int. Ed.* **2013**, *52*, 2836 – 2850; *Angew. Chem.* **2013**, *125*, 2906 – 2921.

[3] J. Boissonade in *Origins of Life: Self-Organization and/or Biological Evolution*, EDP Sciences, Paris, **2008**.

[4] a) E. R. Kay, D. A. Leigh, F. Zerbetto, *Angew. Chem. Int. Ed.* **2007**, *46*, 72 – 191; *Angew. Chem.* **2007**, *119*, 72 – 196; b) B. Lewandowski, G. De Bo, J. W. Ward, M. Papmeyer, S. Kuschel, M. J. Aldegunde, P. M. E. Gramlich, D. Heckmann, S. M. Goldup, D. M. D'Souza, A. E. Fernandes, D. A. Leigh, *Science* **2013**, *339*, 189 – 193; c) C. P. Collier, E. W. Wong, M. Belohradský, F. M. Raymo, J. F. Stoddart, P. J. Kuekes, R. S. Williams, J. R. Heath, *Science* **1999**, *285*, 391 – 394; d) S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173 – 180; e) D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893 – 4012; f) I. H. Stefan Hecht, *Foldamers: Structure, Properties, and Applications*, Wiley-VCH, Weinheim, **2007**.

[5] a) F. V. Cochran, S. P. Wu, W. Wang, V. Nanda, J. G. Saven, M. J. Therien, W. F. DeGrado, *J. Am. Chem. Soc.* **2005**, *127*, 1346 – 1347; b) L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas, D. Hilvert, K. N. Houk, B. L. Stoddard, D. Baker, *Science* **2008**, *319*, 1387 – 1391; c) H. Kries, R. Blomberg, D. Hilvert, *Curr. Opin. Chem. Biol.* **2013**, *17*, 221 – 228.

[6] G. M. Whitesides, R. F. Ismagilov, *Science* **1999**, *284*, 89 – 92.

[7] J.-M. Lehn, *Chem. Soc. Rev.* **2007**, *36*, 151 – 160.

[8] a) S. Otto, R. Furlan, L. E. Ricardo, J. K. M. Sanders, *Curr. Opin. Chem. Biol.* **2002**, *6*, 321 – 327; b) S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders, J. F. Stoddart, *Angew. Chem. Int. Ed.* **2002**, *41*, 898 – 952; *Angew. Chem.* **2002**, *114*, 938 – 993; c) P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652 – 3711; d) J. N. H. Reek, S. Otto, *Dynamic Combinatorial Chemistry*, Wiley-VCH, Weinheim, **2010**; e) J. Li, P. Nowak, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 9222 – 9239; f) A. Herrmann, *Chem. Soc. Rev.* **2014**, *43*, 1899 – 1933.

[9] a) J. K. M. Sanders, *Pure Appl. Chem.* **2000**, *72*, 2265 – 2274; b) S. Otto, K. Severin, *Topics Curr. Chem.* **2007**, *277*, 267 – 288; c) O. Ramström, L. Amorim, R. Caraballo and O. Norberg, in *Dynamic Combinatorial Chemistry*, (Eds. J. Reek Otto), Wiley-VCH, Weinheim, **2010**, pp. 109 – 150.

[10] a) R. F. Ludlow, S. Otto, *Chem. Soc. Rev.* **2008**, *37*, 101 – 108; b) J. R. Nitschke, *Nature* **2009**, *462*, 736 – 738; c) R. A. R. Hunt, S. Otto, *Chem. Commun.* **2011**, *47*, 847 – 858; d) S. Otto, *Acc. Chem. Res.* **2012**, *45*, 2200 – 2210; e) K. Ruiz-Mirazo, C. Briones, A. de La Escosura, *Chem. Rev.* **2014**, *114*, 285 – 366.

[11] a) W. J. Lees, G. M. Whitesides, *J. Org. Chem.* **1993**, *58*, 642 – 647; b) H. Hioki, W. C. Still, *J. Org. Chem.* **1998**, *63*, 904 – 905.

[12] a) M. Rauschenberg, S. Bomke, U. Karst, B. J. Ravoo, *Angew. Chem. Int. Ed.* **2010**, *49*, 7340 – 7345; *Angew. Chem.* **2010**, *122*, 7498 – 7503; b) M. Rauschenberg, S. Bandaru, M. P. Waller, B. J. Ravoo, *Chem. Eur. J.* **2014**, *20*, 2770 – 2782.

[13] a) S. Otto, S. Kubik, *J. Am. Chem. Soc.* **2003**, *125*, 7804 – 7805; b) Z. Rodriguez-Docampo, E. Eugenieva-Ilieva, C. Reyheller, A. M. Belenguer, S. Kubik, S. Otto, *Chem. Commun.* **2011**, *47*, 9798 – 9800.

[14] A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2008**, *47*, 2677 – 2680; *Angew. Chem.* **2008**, *120*, 2717 – 2720.

[15] a) F. B. L. Cougnon, N. Ponnuswamy, N. A. Jenkins, G. D. Pantos, J. K. M. Sanders, *J. Am. Chem. Soc.* **2012**, *134*, 19129 – 19135; b) R. T. S. Lam, A. Belenguer, S. L. Roberts, C. Naumann, T. Jarrosson, S. Otto, J. K. M. Sanders, *Science* **2005**, *308*, 667 – 669; c) H. Y. Au-Yeung, G. D. Pantos, J. K. M. Sanders, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10466 – 10470.

[16] N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş, J. K. M. Sanders, *Science* **2012**, *338*, 783 – 785.

[17] a) S. Otto, R. L. E. Furlan, J. K. M. Sanders, *J. Am. Chem. Soc.* **2000**, *122*, 12063 – 12064; b) S. Otto, R. L. E. Furlan, J. K. M. Sanders, *Science* **2002**, *297*, 590 – 593.

[18] a) K. R. West, K. D. Bake, S. Otto, *Org. Lett.* **2005**, *7*, 2615 – 2618; b) A. R. Stefankiewicz, M. R. Sambrook, J. K. M. Sanders, *Chem. Sci.* **2012**, *3*, 2326 – 2329.

[19] A. R. Stefankiewicz, J. K. M. Sanders, *Chem. Commun.* **2013**, *49*, 5820 – 5822.

[20] a) J. Atcher, A. Moure, I. Alfonso, *Chem. Commun.* **2013**, *49*, 487 – 489; b) J. Atcher, A. Moure, J. Bujons, I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 6869 – 6878.

[21] K. B. Jadhav, R. J. Lichtenecker, A. Bullach, B. Mandal, H.-D. Arndt, *Chem. Eur. J.* **2015**, *21*, 5898 – 5908.

[22] J. Solà, M. Lafuente, J. Atcher, I. Alfonso, *Chem. Commun.* **2014**, *50*, 4564 – 4566.

[23] J. Atcher, I. Alfonso, *RSC Adv.* **2013**, *3*, 25605 – 25608.

[24] We also performed the reaction in concentrations from 0.2 to 2.0 mM of BB without significant changes in the product distribution, therefore excluding the formation of self-templated aggregates.

[25] a) B. Hasenknopf, J.-M. Lehn, N. Boumediene, E. Leize, A. Van Dorsselaer, *Angew. Chem. Int. Ed.* **1998**, *37*, 3265 – 3268; *Angew. Chem.* **1998**, *110*, 3458 – 3460; b) G. Rapenne, C. Dietrich-Buchecker, J.-P. Sauvage, *J. Am. Chem. Soc.* **1999**, *121*, 994 – 1001; c) J.-F. Ayme, J. E. Beves, D. A. Leigh, R. T. McBurney, K. Rissanen, D. Schultz, *Nat. Chem.* **2011**, *4*, 15 – 20.

[26] It is difficult to infer the outcome of the DCL in the strict absence of self-recognition, which is necessary for obtaining the statistical distribution. However, considering the outcome of the corresponding binary mixtures, a non-biased library should contain linear oligomers [**1a**–**3a**$_3$], [**2a**–**3a**$_2$]; homodimers [**1a**$_2$] and [**2a**$_2$], as well as larger oligomers [**1a**$_x$–**2a**$_y$–**3a**$_z$].

[27] a) T. Nishinaga, A. Tanatani, K. Oh, J. S. Moore, *J. Am. Chem. Soc.* **2002**, *124*, 5934 – 5935; b) K. Oh, K.-S. Jeong, J. S. Moore, *J. Org. Chem.* **2003**, *68*, 8397 – 8403.

[28] a) L. Roy, M. A. Case, *J. Am. Chem. Soc.* **2010**, *132*, 8894 – 8896; b) M. A. Case, G. McLendon, *Acc. Chem. Res.* **2004**, *37*, 754 – 762.

[29] Compound distribution is slightly affected by the coordination properties of the buffer used. Thus, bis-tris(propane) gives less selectivity than bis-tris so the results between different buffers are not fully comparable at the same pH.

[30] S. Sukenik, Y. Boyarski, D. Harries, *Chem. Commun.* **2014**, *50*, 8193 – 8196.

[31] S. W. Sisco, J. S. Moore, *Chem. Sci.* **2013**, *4*, 81 – 85.

[32] Aromatic amides tend towards co-planarity, see: a) B. Koenig, O. Moller, P. Bubenitschek, P. G. Jones, *J. Org. Chem.* **1995**, *60*, 4291 – 4293; some crystal structures are available in b) P. Jana, A. Paikar, S. Bera, S. K. Maity, D. Haldar, *Org. Lett.* **2014**, *16*, 38 – 41. In the cyclohexane moiety the amide bond planes are tilted from the main plane; for a crystal structure, see: c) B. R. Bhogala, P. Vishweshwar, A. Nangia, *Acta Crystallogr. Sect. C* **2007**, *63*, o591 – o593.

Received: April 11, 2015
Published online on October 8, 2015

2.4.1. Electronic Supplementary Information (ESI) for:

PUBLICATION B

Adaptive Correction from Virtually Complex Dynamic Libraries: The role of noncovalent interactions in structural selection and folding.

Maria Lafuente, Joan Atcher, Jordi Solà and Ignacio Alfonso**

Department of Biological Chemistry and Molecular Modelling.
Institute of Advanced Chemistry of Catalonia, IQAC-CSIC
Jordi Girona 18-26, 08034, Barcelona, Spain. Fax: (+34)932045904
E-mail: ignacio.alfonso@iqac.csic.es
jordi.sola@iqac.csic.es

Table of Contents:

GENERAL METHODS

General: Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka or Merck) and were used without further purification. Flash chromatographic purifications and preparative reversed-phase purifications were performed on a Biotage® Isolera Prime™ equipment. TLCs were performed using 6x3 cm SiO_2 pre-coated aluminium plates (ALUGRAM® SIL G/UV_{254}).

Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC) analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using:

i) For the characterization of [**1b**]: a reversed-phase X-Terra C_{18} (15 x 0.46 cm, 5 μm) column. (CH_3CN + 0.07% TFA and H_2O + 0.1% TFA) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelengths were set at 220 nm.

ii) For the analysis of the DCLs: a reversed-phase kromaphase C_{18} (25 x 0.46 cm, 5μm) column. (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelength was set at 254 nm.

Nuclear Magnetic Resonance (NMR) spectroscopic experiments were carried out on a Varian INOVA 500 spectrometer (500 MHz for ^1H and 125 MHz for ^{13}C) and a Varian Mercury 400 instrument (400 MHz for ^1H and 101 MHz for ^{13}C). The chemical shifts are reported in ppm relative to trimethylsilane (TMS), and coupling constants (*J*) are reported in Hertz (Hz).

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).

High Resolution Mass Spectrometry (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.

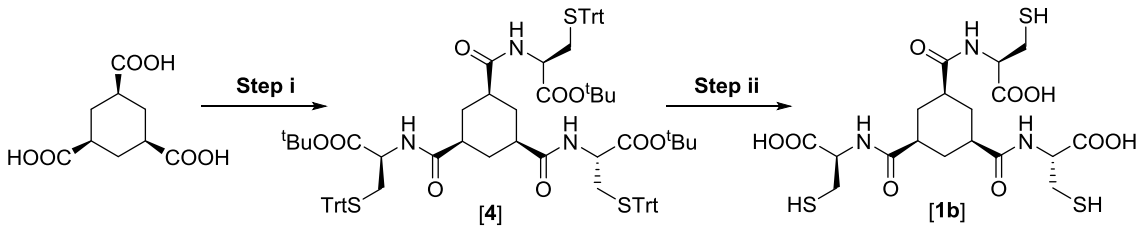
MOLECULAR MODELING

Molecular modeling studies 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. This force field has proved to be suitable for the conformational analysis of pseudopeptide and peptoid molecules.[1] Besides, we introduced the NOE constraints from the experimental cross-peaks observed in the ROESY experiment. The obtained local minima were ordered following the corresponding MMFFaq energies. We considered all the carboxylic groups as carboxylate anions and the amine function of the pendant cysteine as the ammonium cation (total charge: -3).

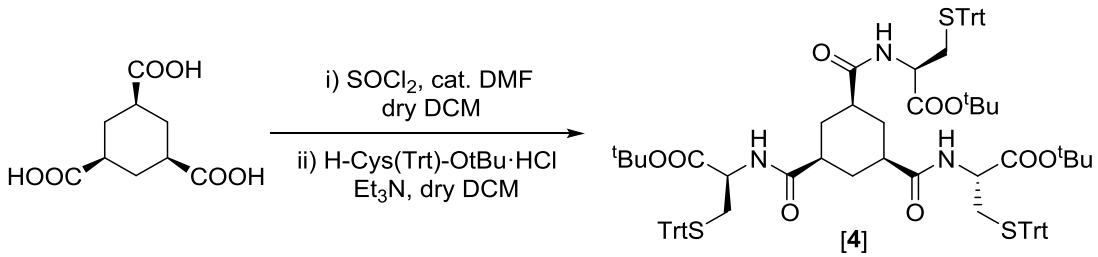
SYNTHESIS OF THE BUILDING BLOCKS

The trithiol [**1a**] was synthesised as previously descrived[2] and the dithiols [**2a-c**] were synthesized as we previously reported.[3] The monothiols [**3a-g**] were purchased from commercial suppliers (Sigma-Aldrich and Iris Biotech).

Synthetic scheme of [**1b**]

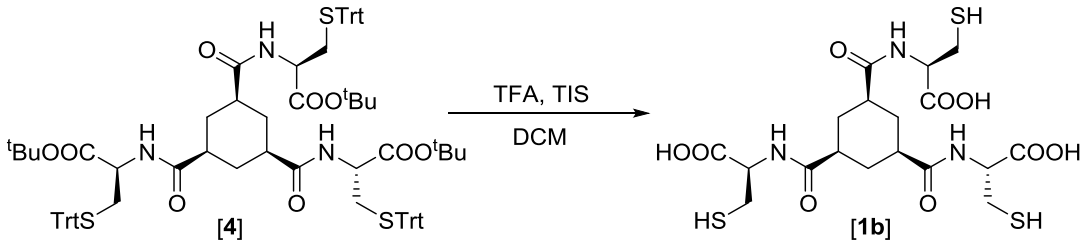


Step i: Experimental procedure for the synthesis of [**4**]



To a solution of cis-1,3,5-cyclohexanetricarboxylic acid (65 mg, 0.30 mmol) in dry DCM (10 mL), thionyl chloride (10 mL) and a drop of dry DMF were added. After 3 hours stirring at the reflux temperature, the excess of thionyl chloride was removed under reduced pressure followed by co-evaporation with toluene (5 mL). The residue was dissolved in dry DCM (5 mL) and added dropwise to a solution of H-L-Cys(Trt)-O^{t}Bu·HCl (500 mg, 1.10 mmol) and triethylamine (1.5 mL, 10.8 mmol) in dry DCM (20 mL) cooled to 0ºC in an ice-water bath. The mixture was stirred overnight at room temperature and then diluted with DCM (50 mL), washed with saturated aqueous $NaHCO_3$ (20 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography using hexane: AcOEt as eluent (from 20% to 40% AcOEt, Rf AcOEt/Hexane, 2:3 (v:v): 0.53) to give 335 mg of [**4**] (79% yield) as a yellow foam. HRMS (ESI+) calcd. for $C_{87}H_{93}N_3O_9S_3$ $[M+H]^+$ (m/z): 1420.6147, found: 1420.6246. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.39 (d, *J* = 7.2 Hz, 18H, CH_{Ar}), 7.27 (t, *J* = 7.5 Hz, 18H, CH_{Ar}), 7.19 (t, *J* = 7.2 Hz, 9H, CH_{Ar}), 6.00 (d, *J* = 7.7 Hz, 3H, CONH), 4.51 (dt, *J* = 7.7, 4.9 Hz, 3H, CONHC<u>H</u>), 2.59 (dd, *J* = 12.0, 5.3 Hz, 3H, C<u>H</u>$_2$STrt), 2.51 (dd, *J* = 12.0, 4.6 Hz, 3H, C<u>H</u>$_2$STrt), 2.19 (br t, *J* = 12.5 Hz, 3H, C<u>H</u>CONH), 2.06 (br d, *J* = 12.9 Hz, 3H, CHC<u>H</u>$_2$CH), 1.65 (q, *J* = 12.7 Hz, 3H, CH_2CHC<u>H</u>$_2$CH), 1.42 (s, 27H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 173.2 (3 x CO), 169.4 (3 x CO), 144.4 (9 x C_{Ar}), 129.6 (18 x CH_{Ar}), 128.1 (18 x CH_{Ar}), 127.0 (9 x CH_{Ar}), 82.7 (3 x <u>C</u>Me$_3$), 66.7 (3 x <u>C</u>Ph$_3$), 51.5 (3 x CONH<u>C</u>H), 44.1 (3 x <u>C</u>HCONH), 34.3 (3 x <u>C</u>H$_2$STrt), 31.4 (3 x CH<u>C</u>H$_2$CH), 28.1 (9 x CH_3).

Step ii: Experimental procedure for the synthesis of [**1b**]



To a solution of [**4**] (125 mg, 0.088 mmol) in DCM (2 mL), trifluoroacetic acid (TFA, 2 mL) and triisopropylsilane (TIS, 0.210 mg, 1.03 mmol) were added rapidly and under stirring. The reaction mixture was allowed to stir at room temperature for 3 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. The product was purified using reversed-phase flash chromatography (gradient: from 2% to 15% CH_3CN in H_2O) and 20 mg of [**1b**] (43% yield) were obtained as a white solid. HRMS (ESI+) calcd. for $C_{18}H_{27}N_3O_9S_3$ $[M+H]^+$ (m/z): 526.0988, found: 526.0992. ^{1}H NMR (400 MHz, MeOD-d_4): δ = 4.58 (dd, X of ABX, J_{BX} = 6.8, J_{AX} = 4.5 Hz, 3H, CONHC<u>H</u>), 2.98 (dd, A of ABX, J_{AB} = 13.9, J_{AX} = 4.6 Hz, 3H, C<u>H</u>$_2$SH), 2.88 (dd, B of ABX, J_{AB} = 13.9, J_{BX} = 6.8 Hz, 3H, C<u>H</u>$_2$SH), 2.51 (br t, J = 12.4 Hz, 3H, C<u>H</u>CONH), 2.04 (br d, J = 12.8 Hz, 3H, CHC<u>H</u>$_2$CH), 1.67 (q, J = 12.7 Hz, 3H). ^{13}C NMR (101 MHz, MeOD-d_4): δ = 177.3 (3 x CONH), 173.3 (3 x COOH), 55.9 (3 x CONH<u>C</u>H), 44.6 (3 x <u>C</u>HCONH), 32.8 (3 x CH<u>C</u>H$_2$CH), 26.7 (3 x CH$_2$SH).

NMR spectra, HRMS (ESI+) spectrum and HPLC trace of [**1b**]

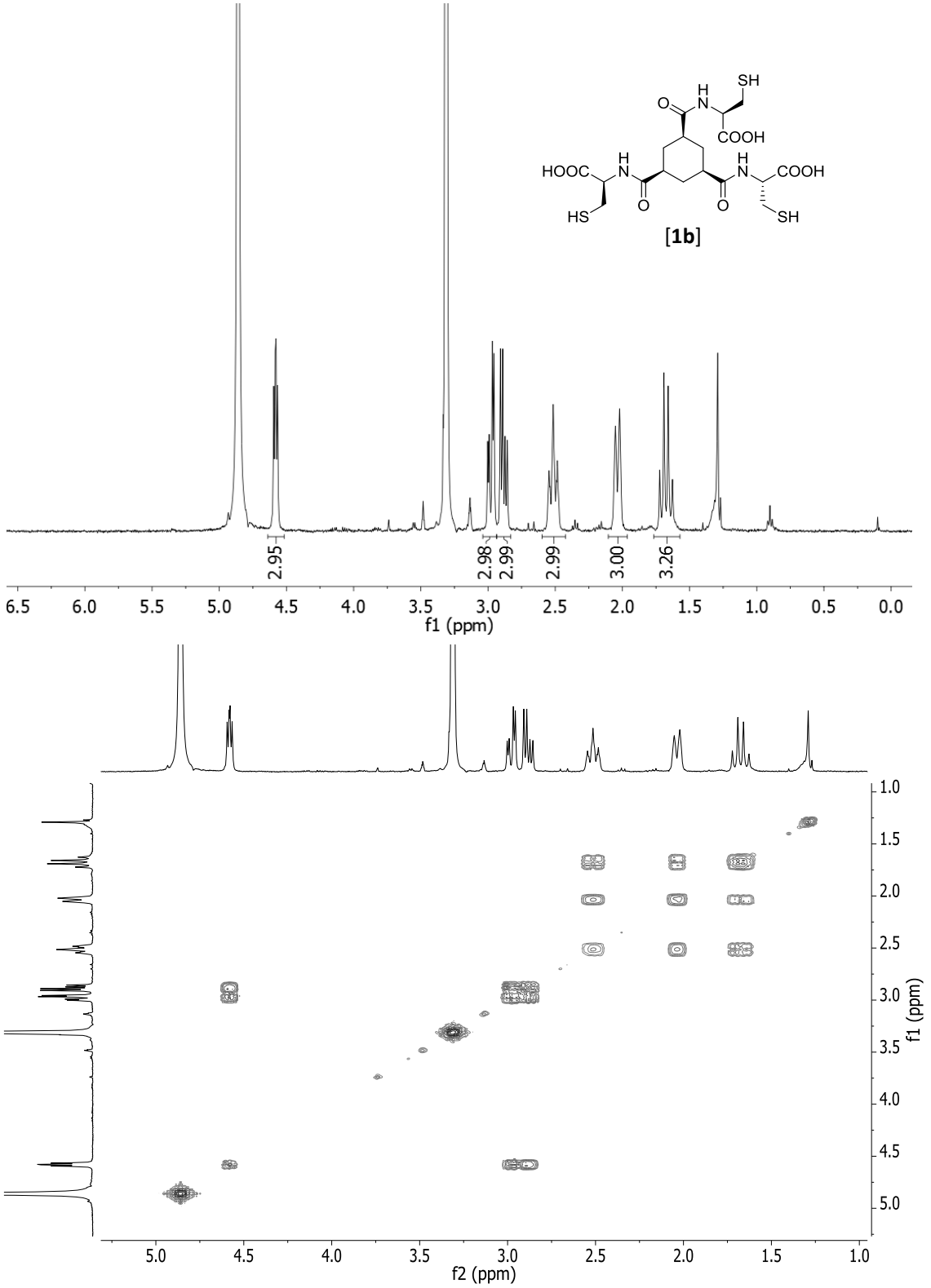


Figure S1: ^{1}H (400 MHz, 298 K in MeOD-d_4) and ^{1}H-^{1}H gCOSY (400 MHz, 298 K in MeOD-d_4) spectra of [**1b**].

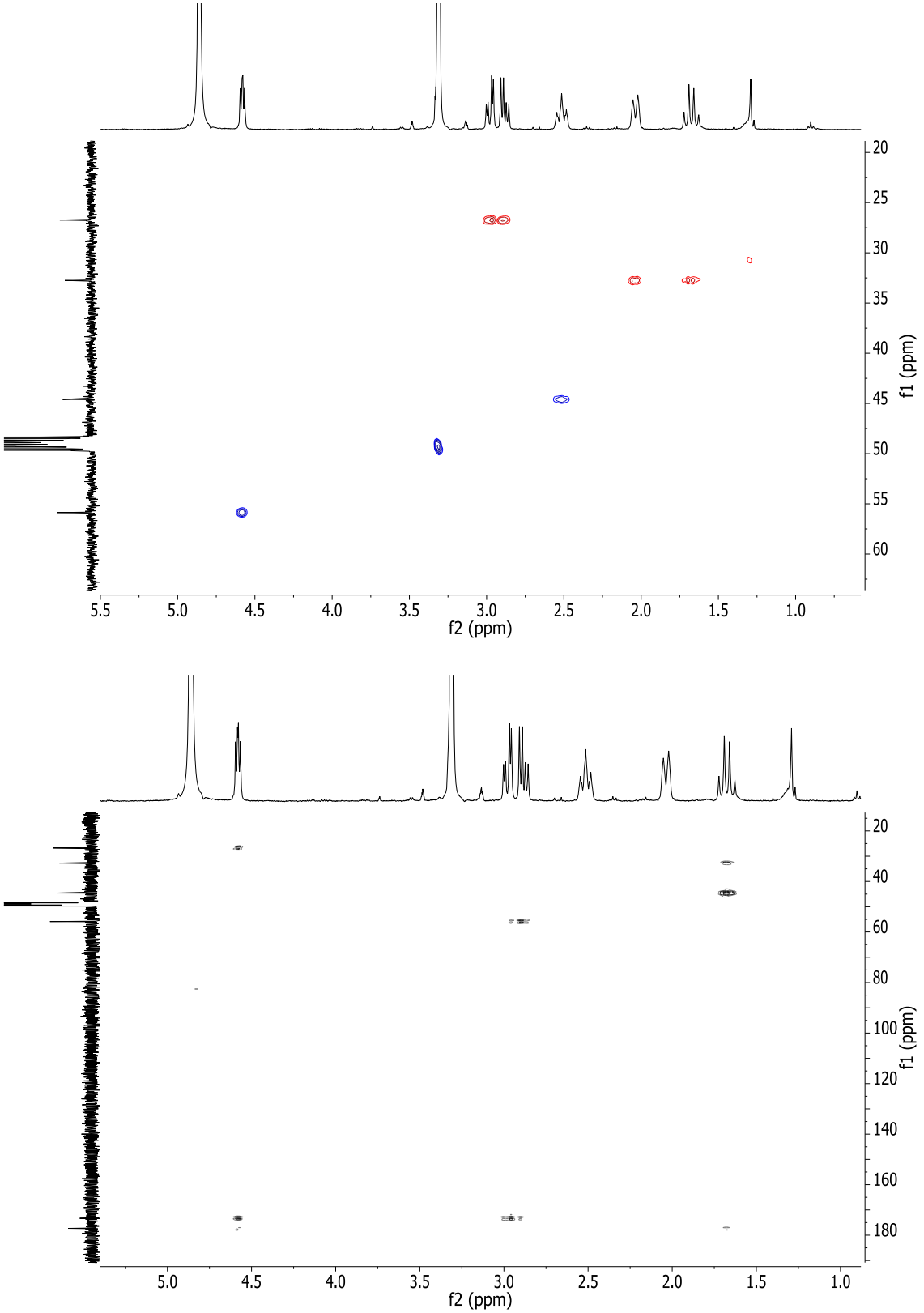


Figure S2: ^{1}H-^{13}C gHSQC (400 MHz, 298 K in MeOD-d_4) and ^{1}H-^{13}C gHMBC (400 MHz, 298 K in MeOD-d_4) spectra of [**1b**].

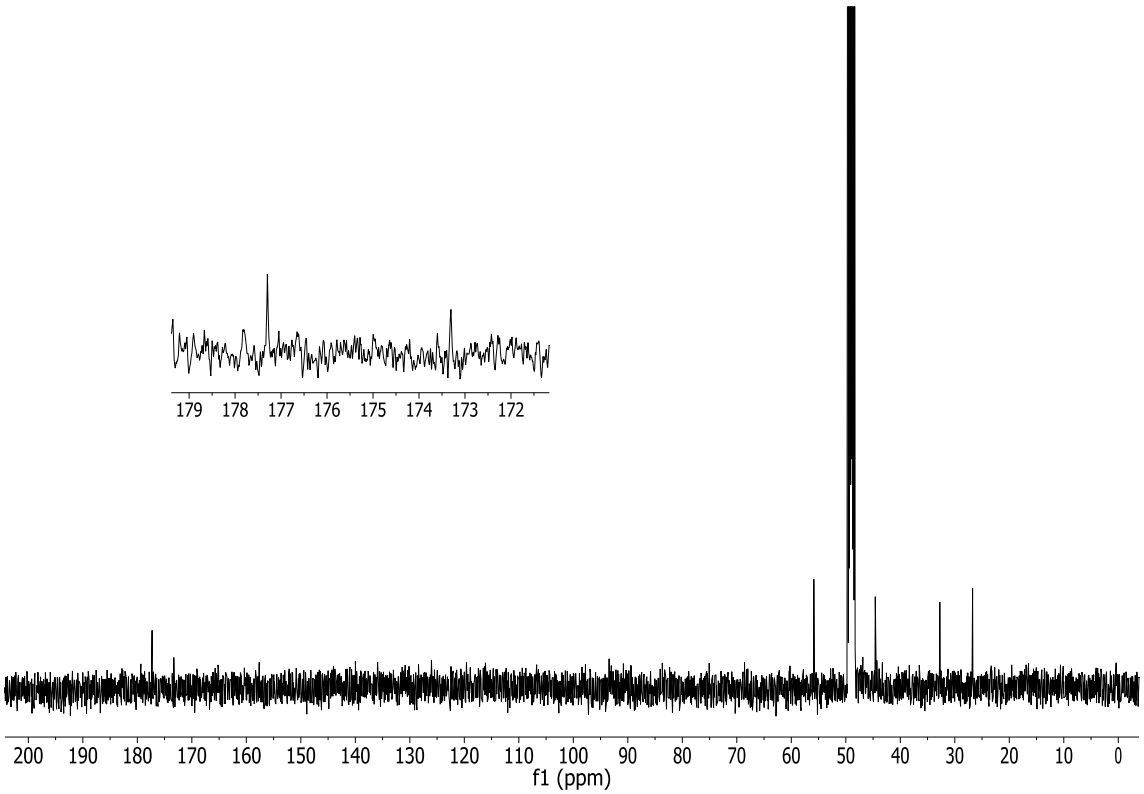


Figure S3: ^{13}C (101 MHz, 298 K in MeOD-d_4) spectrum of [**1b**].

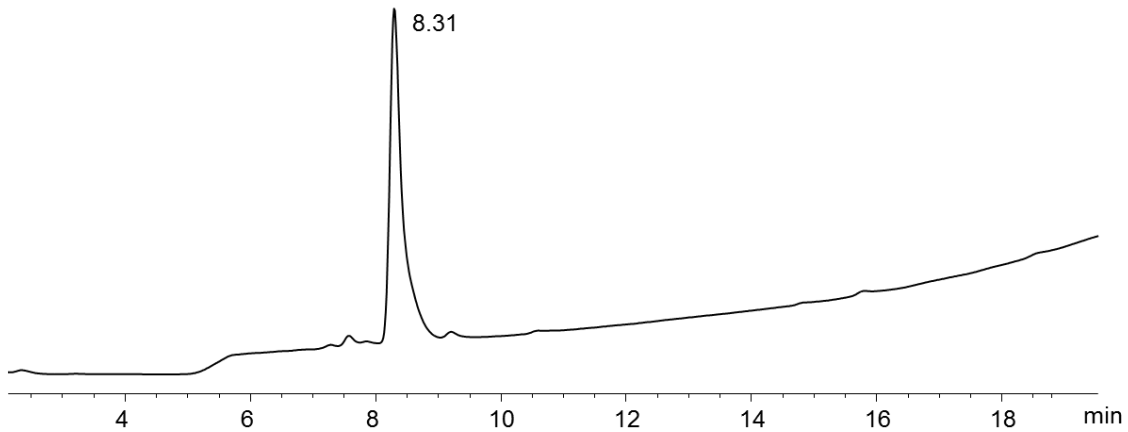


Figure S4: HPLC trace of [**1b**] (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

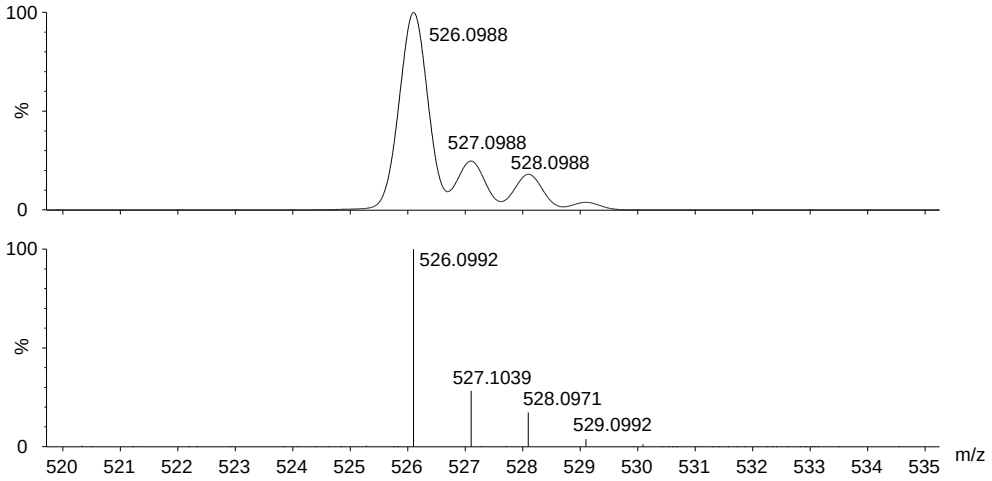


Figure S5: Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1b**].

DYNAMIC COMBINATORIAL LIBRARIES

General procedure for the preparation and HPLC analysis of the DCLs

A 66.7 mM BIS-Tris methane buffer was prepared by dissolving 1.39 g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq).

For the experiments performed at different pH values, a 100 mM BIS-Tris propane buffer was used for pH 6.5-9.5, and a 100 mM HCOOH/$HCOO^-$ buffer was used for pH 4.5. The BIS-Tris buffer was prepared by dissolving 2.82g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5, 7.5, 8.5 and 9.5 by the addition of HCl (aq). The HCOOH/$HCOO^-$ buffer was prepared by dissolving 0.45 mL of the formic acid in 100 mL of milli-Q water and adjusting the pH of the solution to 4.5 by the addition of NaOH (aq).

The reaction mixtures were prepared by dilution of individual stocks of the building blocks (BBs) **1a-b**, **2a-c** and **3a-g** in DMSO. For those experiments to be compared, the reaction mixtures were prepared by dilution of a stock mixture of the BBs, ensuring no differences in concentration between the reaction mixtures of the same batch. Unless otherwise specified, the DCLs were prepared at 0.5 mM of the di- and tripodal BBs (**1a-b** and **2a-b**) in a 50 mM BIS-Tris methane buffer (pH 6.5) with 25% DMSO. The concentrations of the monopodal BBs **3a-g** are specified for each of the experiments.

The mixtures were analysed by means of HPLC or LC-MS at different reaction times. Complete oxidation is achieved after 24h h for 25% DMSO or 48h for 10% DMSO. The HPLC samples were prepared by adding 40 μL of the corresponding reaction mixture to 65 μL of a solution of 89% H_2O, 10% CH_3CN and 1% TFA. Eluent used: 2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 40% CH_3CN over 48 min.

Mixture of **1a**+**2a**+**3a** at different reaction times with 25% DMSO

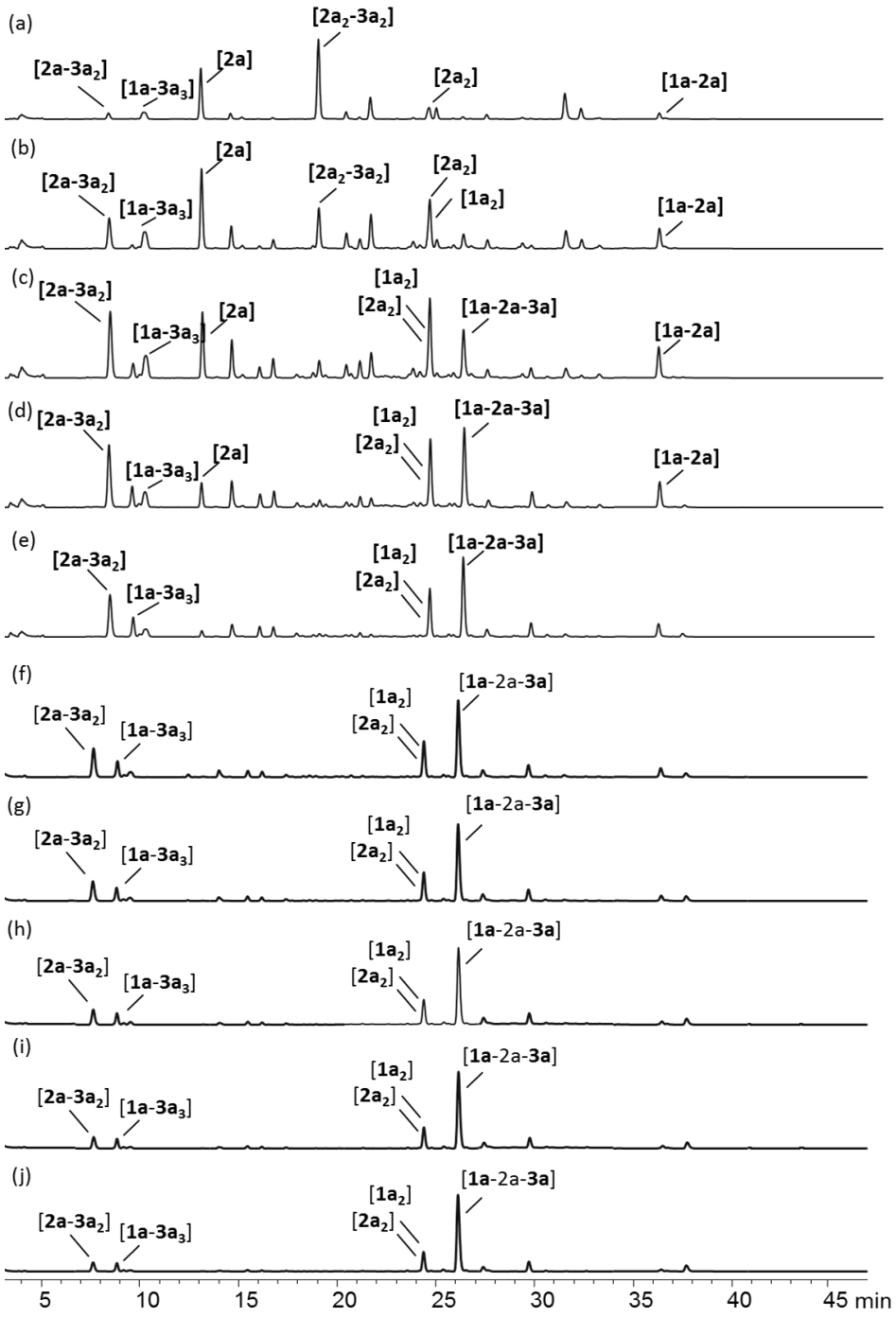


Figure S6: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) and **3a** (2.5 mM), at pH 6.5 with 25% of DMSO at different reaction times: 1 min (a), 5 min (b), 15 min (c), 30 min (d), 45 min (e), 1 hour (f), 1.25 hours (g), 1.50 hours (h), 1.75 hours (i), 2.0 h (j)

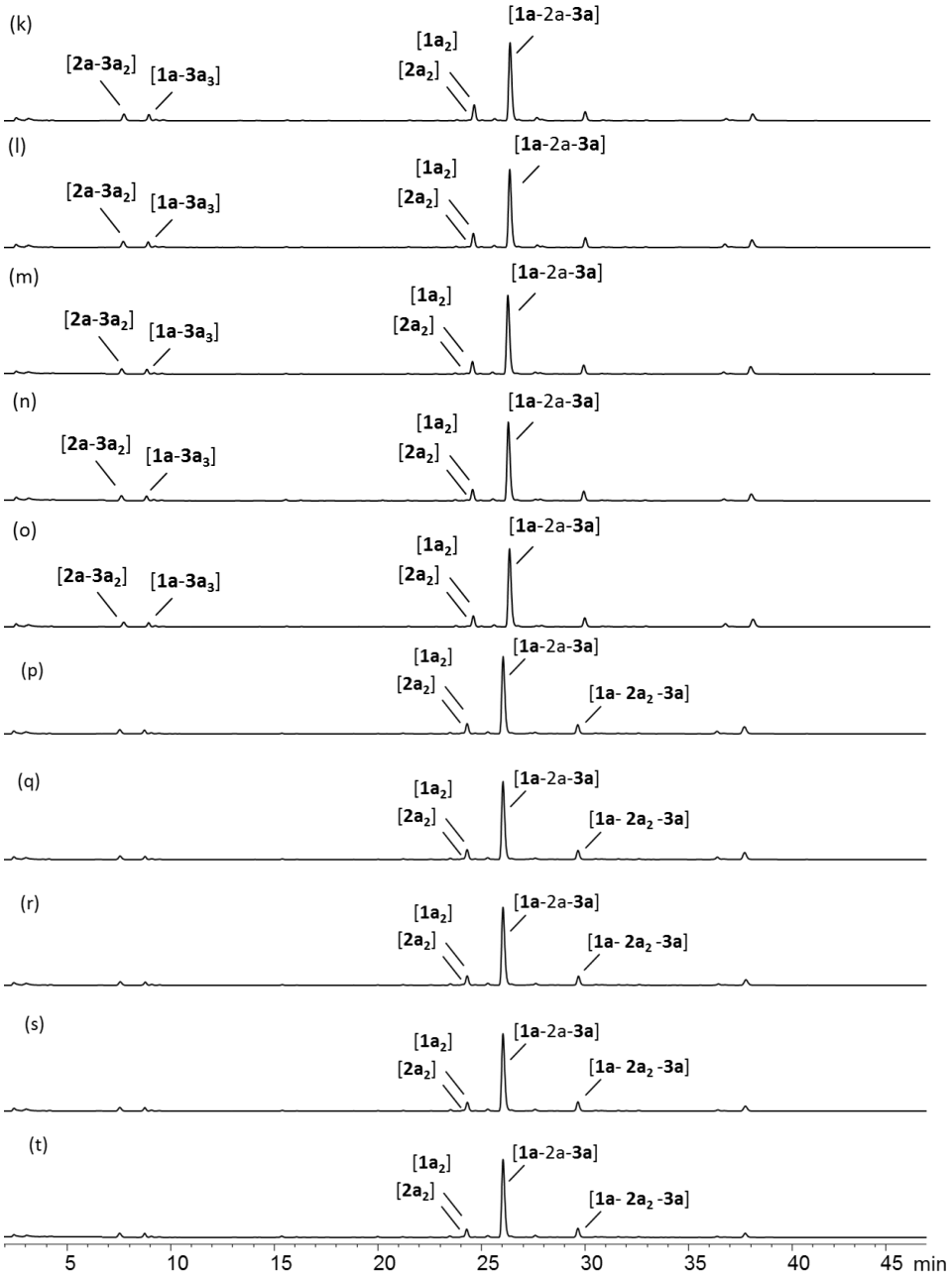


Figure S7: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) and **3a** (2.5 mM), at pH 6.5 with 25% of DMSO at different reaction times: 2.5 h (k), 3h (l), 3.30h (m), 4h (n), 4.30h (o), 5h (p), 5.30h (q), 6h (r), 7h (s), 24h (t)

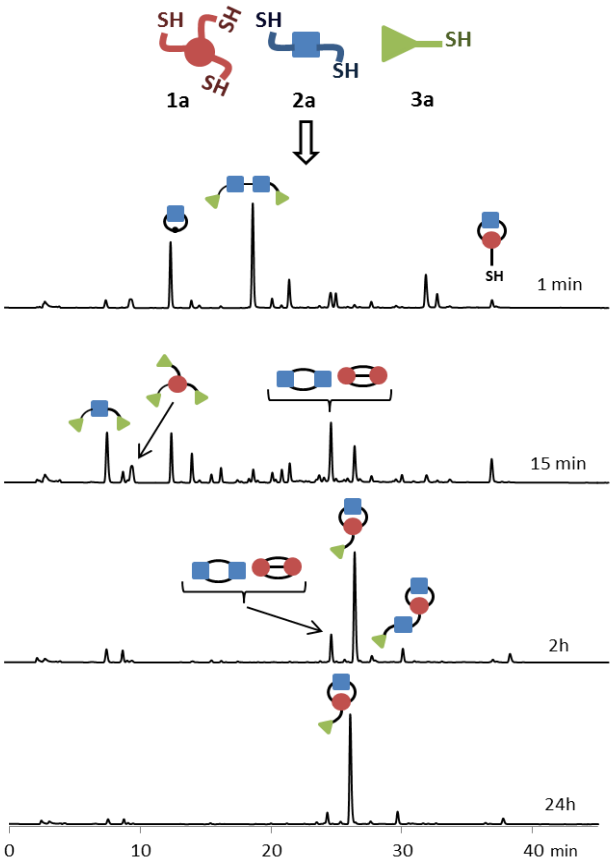


Figure S8: Summary cartoon of HPLC traces (UV detection at 254 nm) for the time evolution of a sample containing **1a**, **2a** and **3a** (**1a**, **2a** 0.5 mM each, **3a** 2.5 mM, bis-tris buffer, pH 6.5, 25% aqueous DMSO).

Mixture of **1a**+**3a** at different reaction times with 25%DMSO

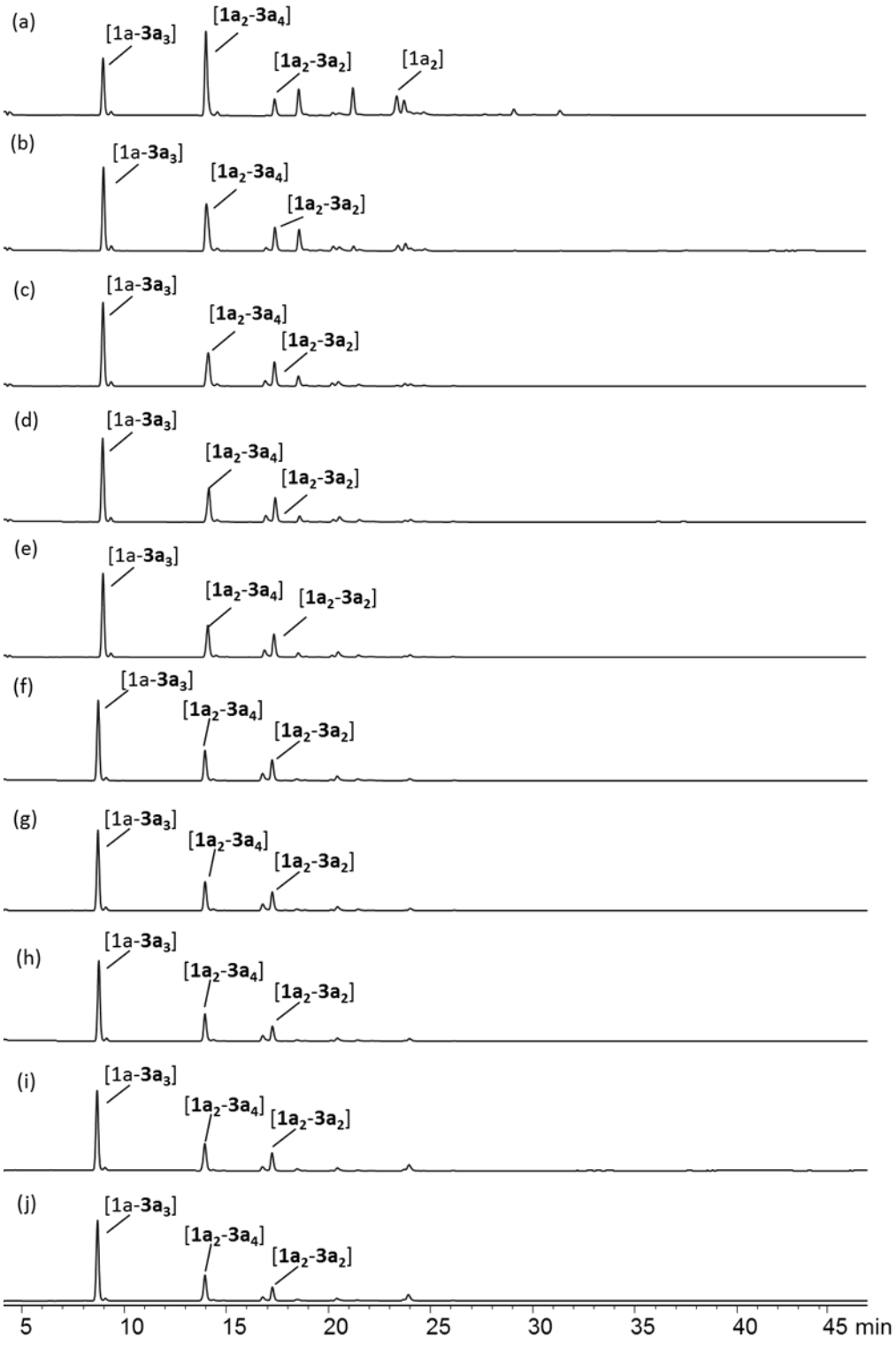


Figure S9: HPLC traces of the mixture of **1a** (0.5 mM) and **3a** (2.5 mM), at pH 6.5 with 25% of DMSO at different reaction times: 1 min (a), 5 min (b), 10 min (c), 15 min (d), 30 min (e), 45 min (f), 1 hour (g), 1.50 hours (h), 2.0 hours (i), 2.50 h (j)

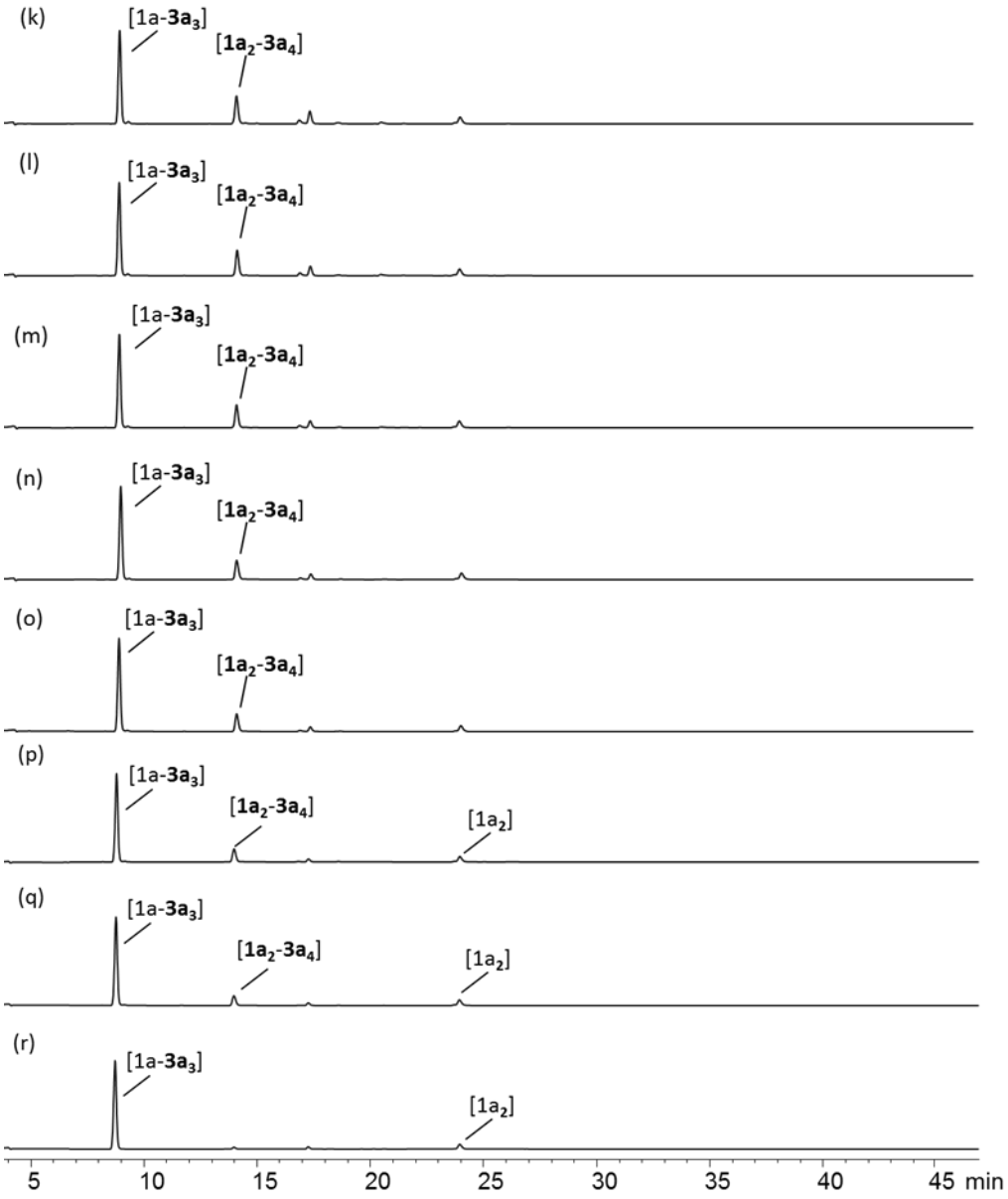


Figure S10: HPLC traces of the mixture of **1a** (0.5 mM) and **3a** (2.5 mM), at pH 6.5 with 25% of DMSO at different reaction times: 3 hours (k), 3.5 hours (l), 4 hours (m), 4.5 hours (n), 5 hours (o), 6 hours (p), 7 hours (q), 24 hours (r).

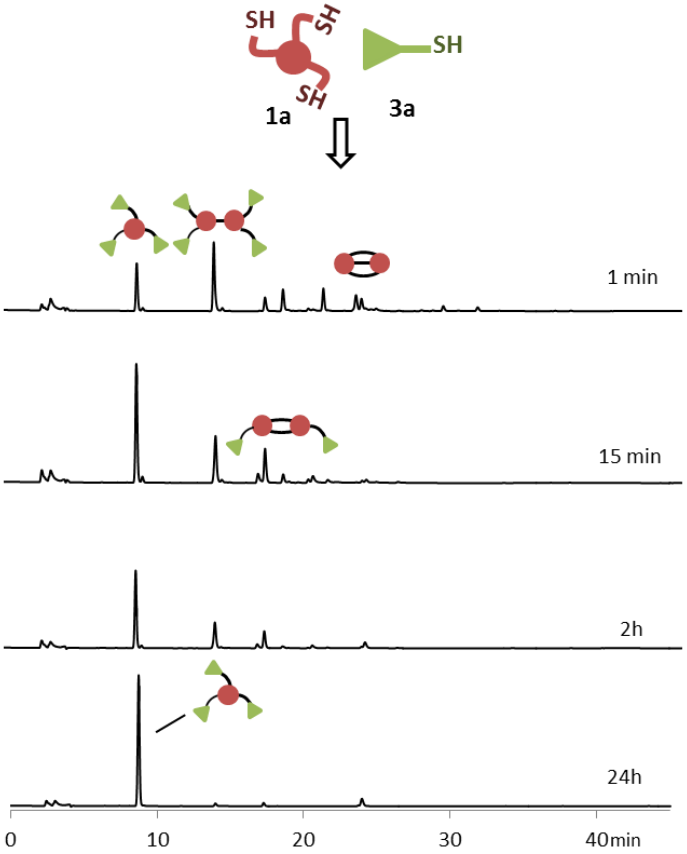


Figure S11: Summary cartoon of HPLC traces (UV detection at 254 nm) for the time evolution of a sample containing **1a** and **3a** (**1a** 0.5 mM, **3a** 2.5 mM, bis-tris buffer, pH 6.5, 25% aqueous DMSO).

Mixture of **2a**+**3a** at different reaction times with 25%DMSO

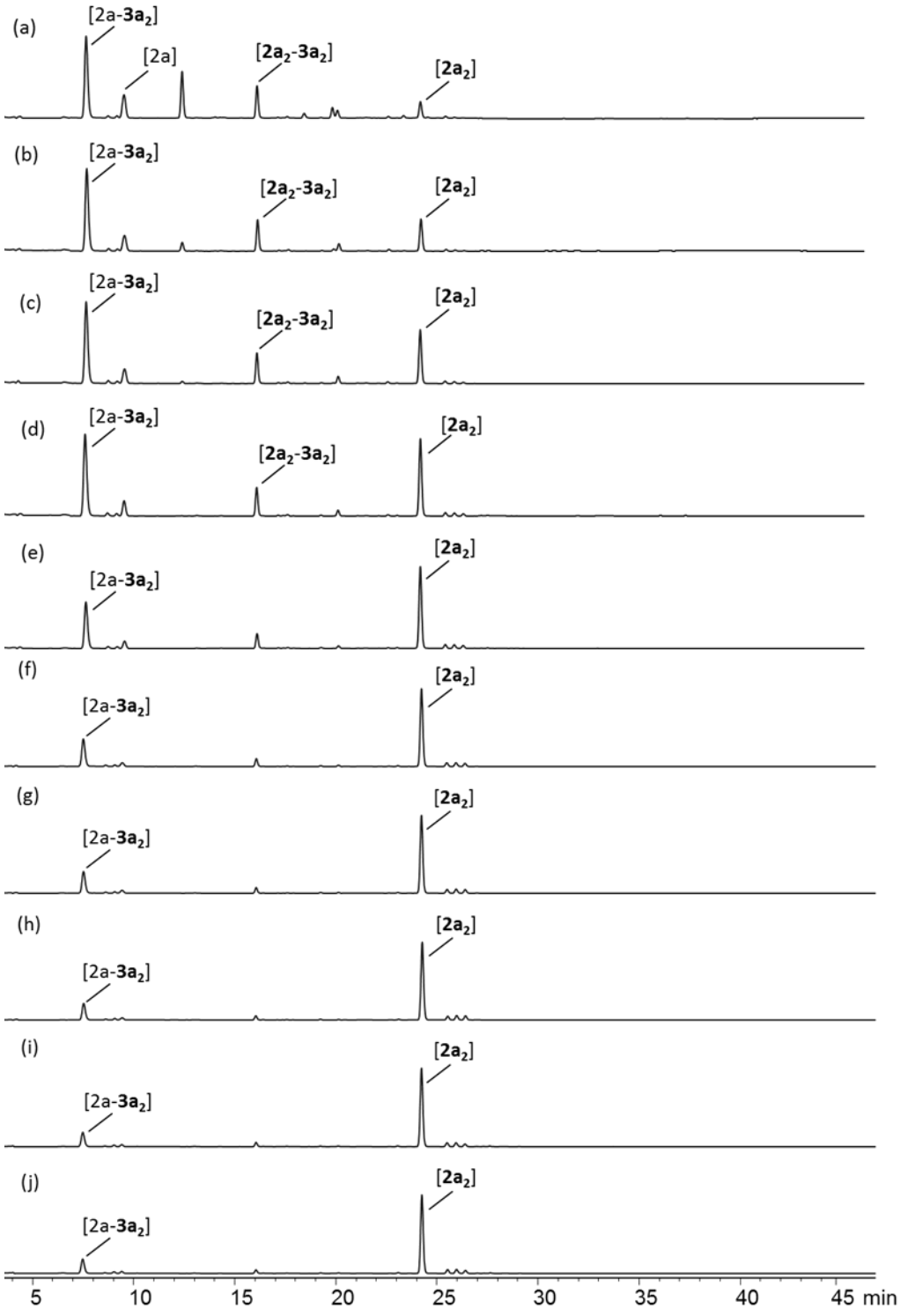


Figure S12: HPLC traces of the mixture of **2a** (0.5 mM) and **3a** (2.5 mM), at pH 6.5 with 25% of DMSO at different reaction times: 1 min (a), 5 min (b), 10 min (c), 15 min (d), 30 min (e), 45 min (f), 1 hour (g), 1.50 hours (h), 2 hours (i), 2.50 h (j)

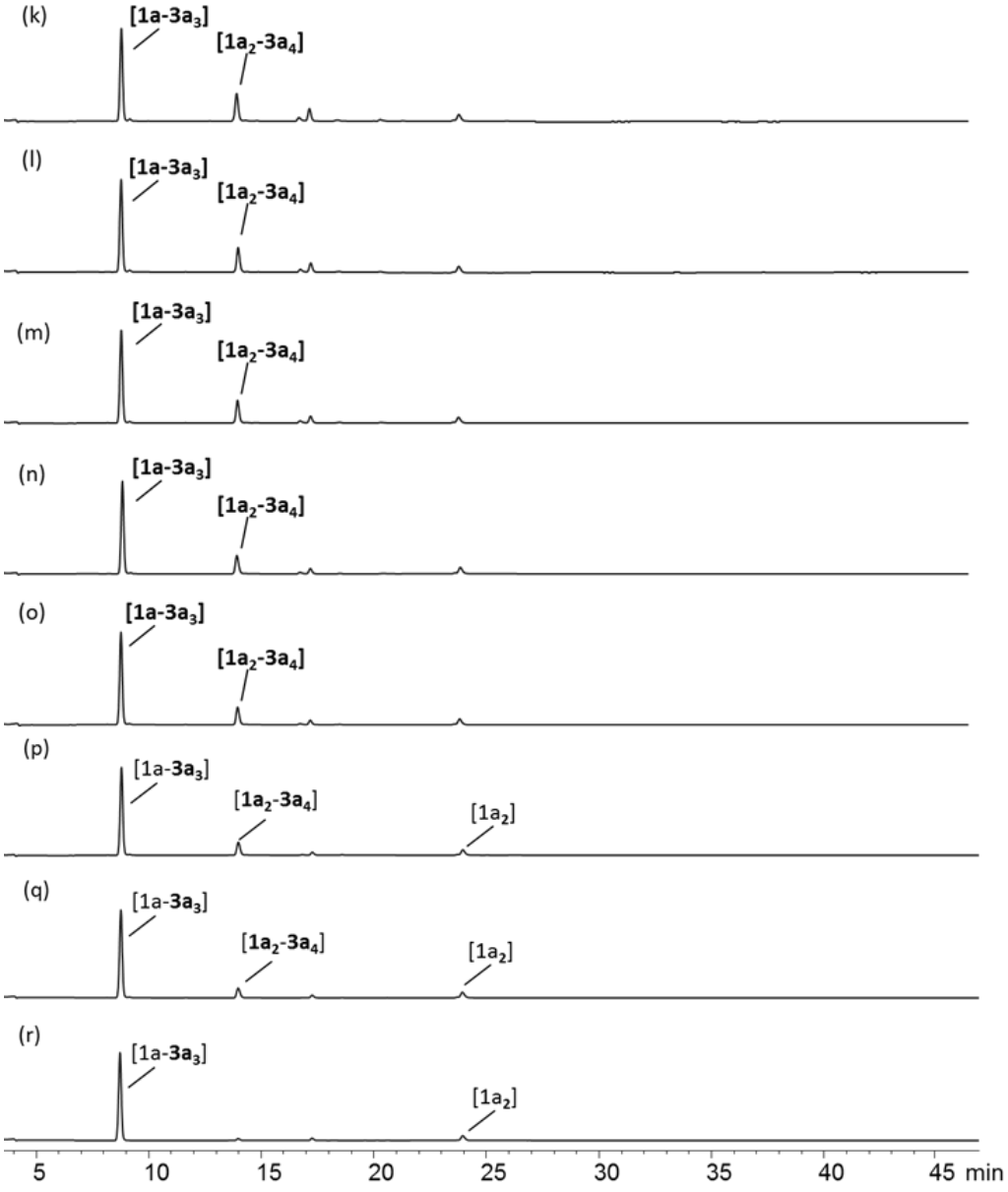


Figure S13: HPLC traces of the mixture of **2a** (0.5 mM) and **3a** (2.5 mM), at pH 6.5 with 25% of DMSO at different reaction times: 3 hours (k), 3.50 hours (l), 4 hours (m), 4.50 hours (n), 5 hours (o), 6 hours (p), 7 hours (q), 24 hours (r).

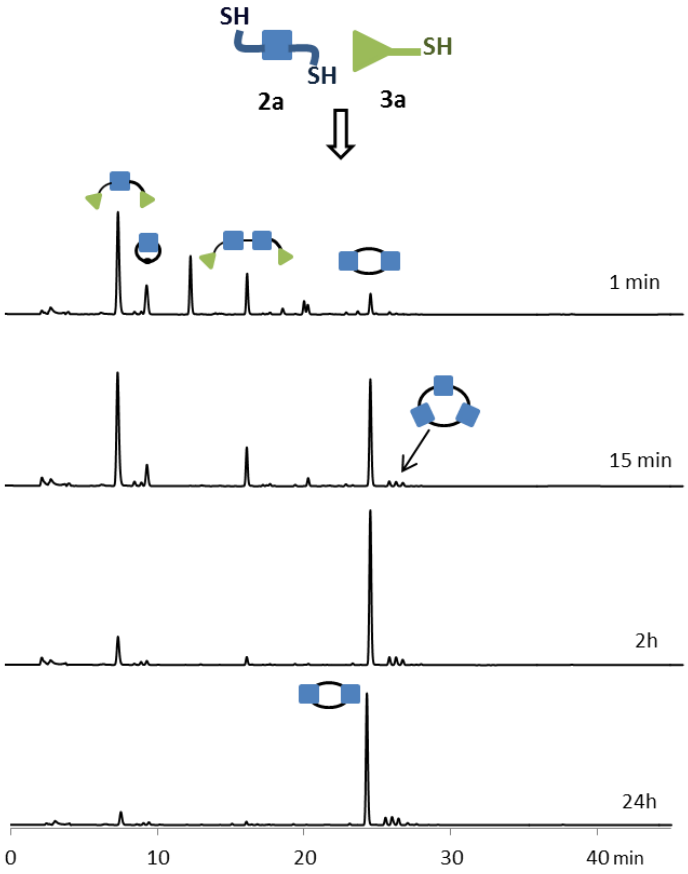


Figure S14: Summary cartoon of HPLC traces (UV detection at 254 nm) for the time evolution of a sample containing **2a** and **3a** (**2a** 0.5 mM, **3a** 2.5 mM, bis-tris buffer, pH 6.5, 25% aqueous DMSO).

Mixture of **1a**+**2a** at different reaction times with 25% DMSO

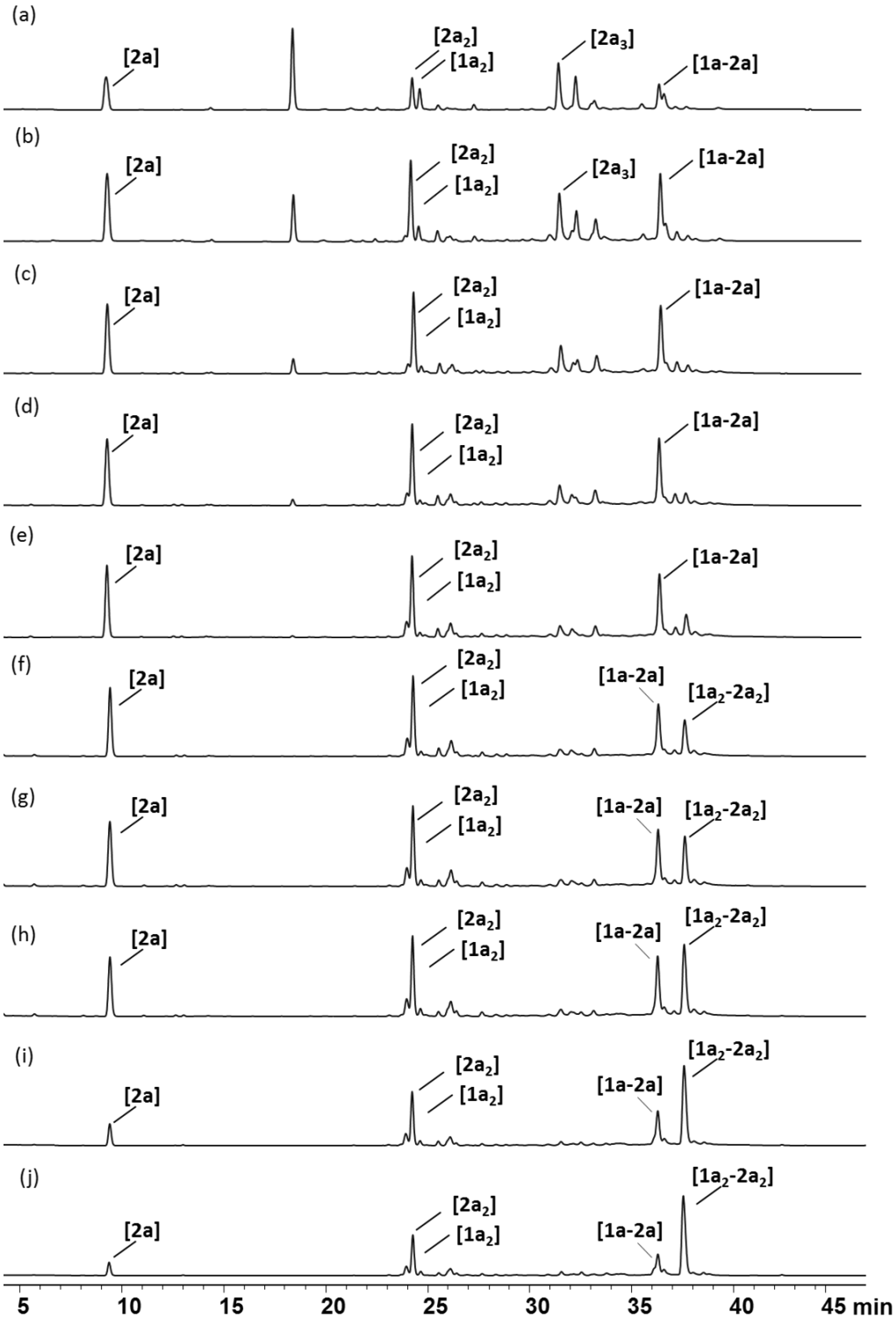


Figure S15: HPLC traces of the mixture of **1a+2a** (0.5 mM each), at pH 6.5 with 25% of DMSO at different reaction times: 1 min (a), 5 min (b), 10 min (c), 15 min (d), 30 min (e), 45 min (f), 1 hour (g), 1.50 hours (h), 2 hours (i), 2.50 h (j)

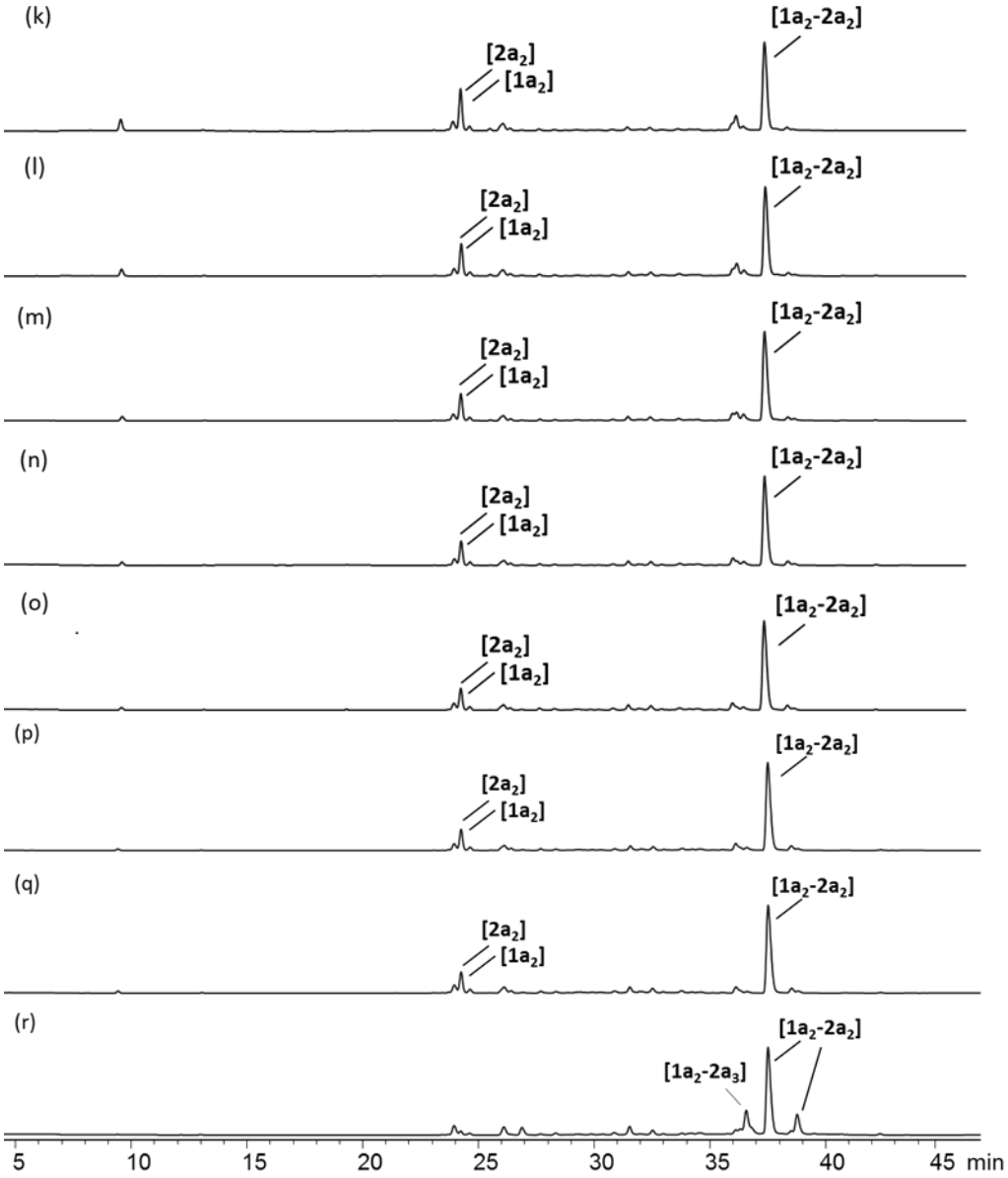


Figure S16: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each), at pH 6.5 with 25% of DMSO at different reaction times: 3 hours (k), 3.50 hours (l), 4 hours (m), 4.50 hours (n), 5 hours (o), 6 hours (p), 7 hours (q), 24 hours (r).

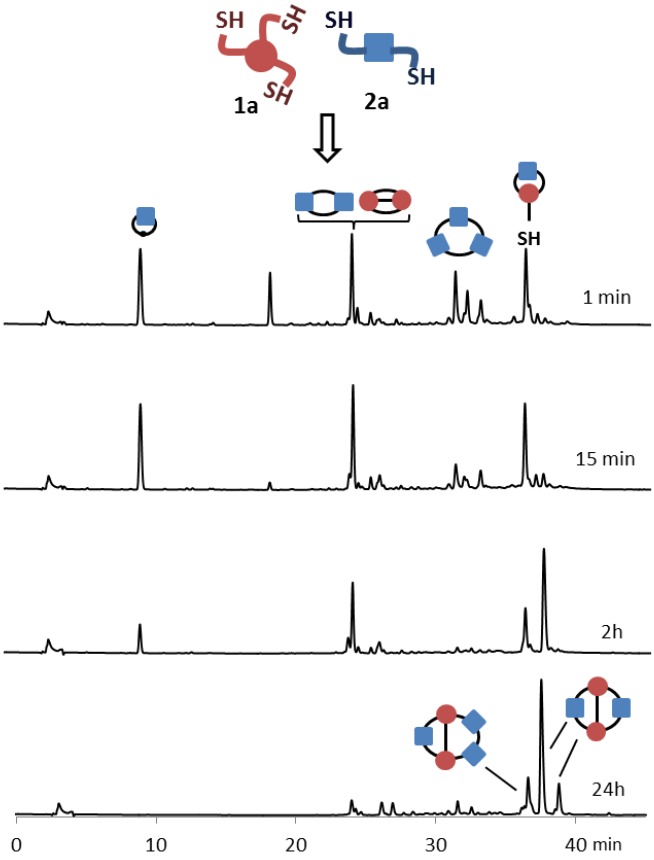


Figure S17: Summary cartoon of HPLC traces (UV detection at 254 nm) for the time evolution of a sample containing **1a** and **2a** (**1a, 2a** 0.5 mM each, bis-tris buffer, pH 6.5, 25% aqueous DMSO).

Mixture of **1a**+**2a**+**3a** at different pH values with 25%DMSO

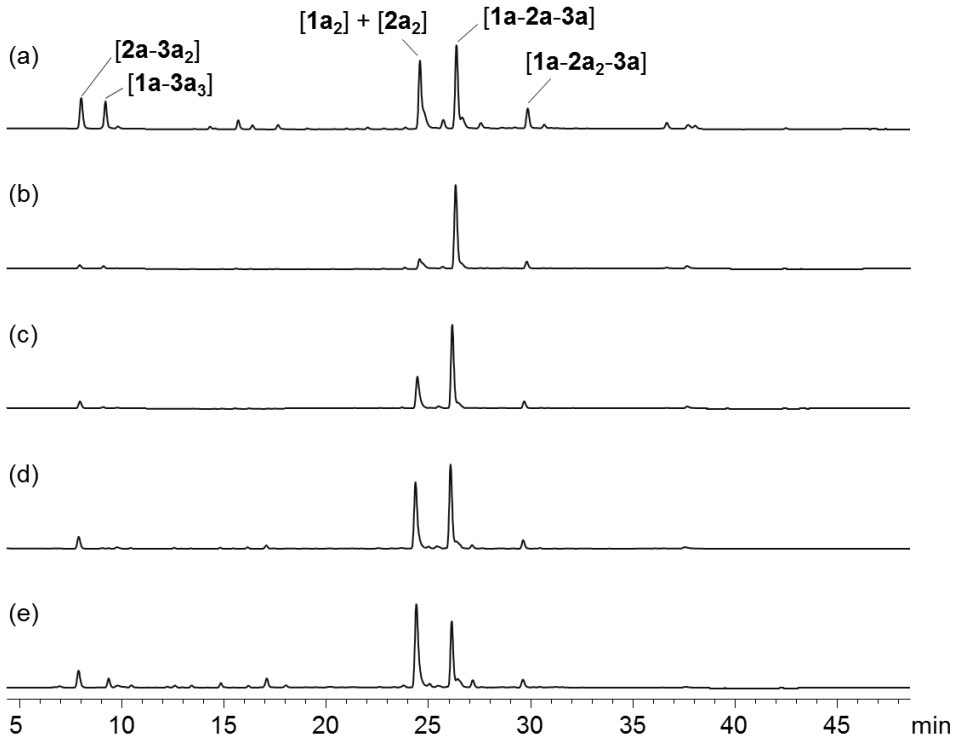


Figure S18: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) and **3a** (2.5 mM), with 25% of DMSO at different pH values: 4.5 (a), 6.5 (b), 7.5 (c), 8.5 (d) and 9.5 (e).

Mixture of **1a**+**2a**+**3a** at different pH values with 10%DMSO

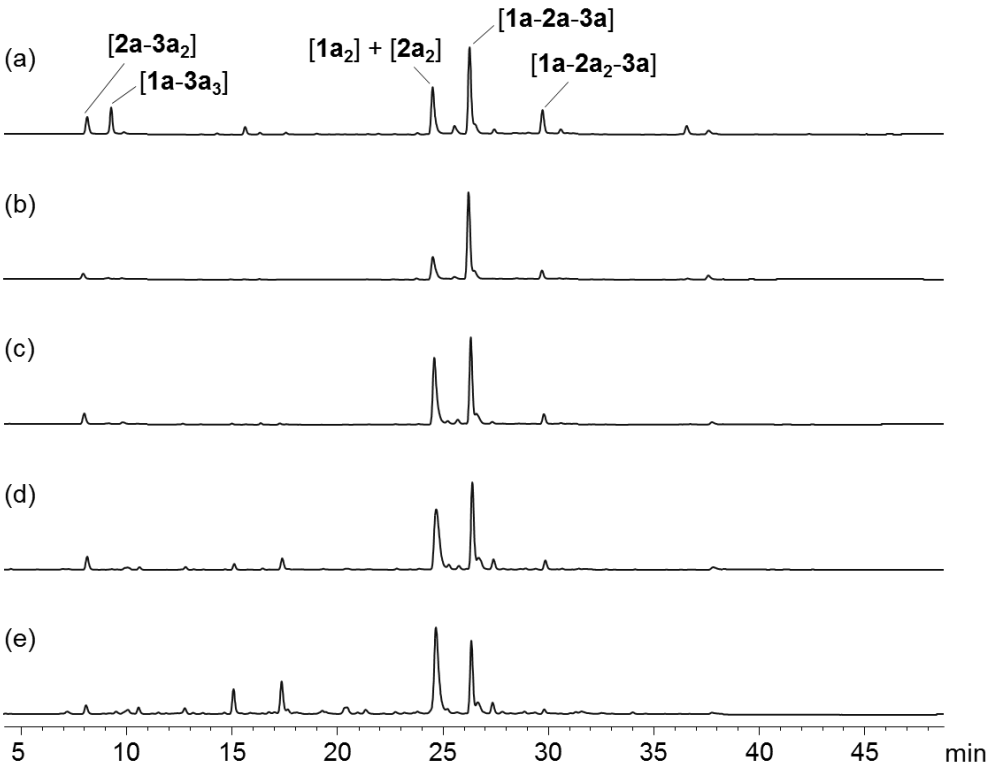


Figure S19: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) and **3a** (2.5 mM), with 10% of DMSO at different pH values: 4.5 (a), 6.5 (b), 7.5 (c), 8.5 (d) and 9.5 (e).

Mixture of **1a**+**2a**+**3a** at different NaCl concentrations with 25%DMSO

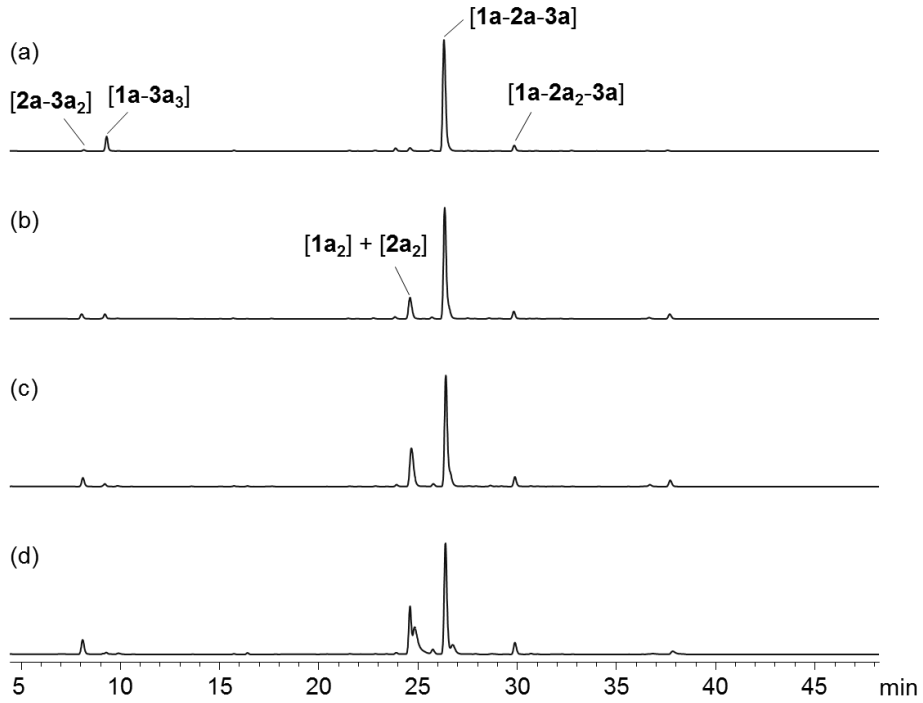


Figure S20: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) and **3a** (2.5 mM), at pH 6.5 with 25% of DMSO at different sodium chloride concentrations: 0.0 M (a), 0.5 M (b), 1.0 M (c) and 2.0 M (d).

Mixture of **1a**+**2a**+**3a** at different NaCl concentrations with 10%DMSO

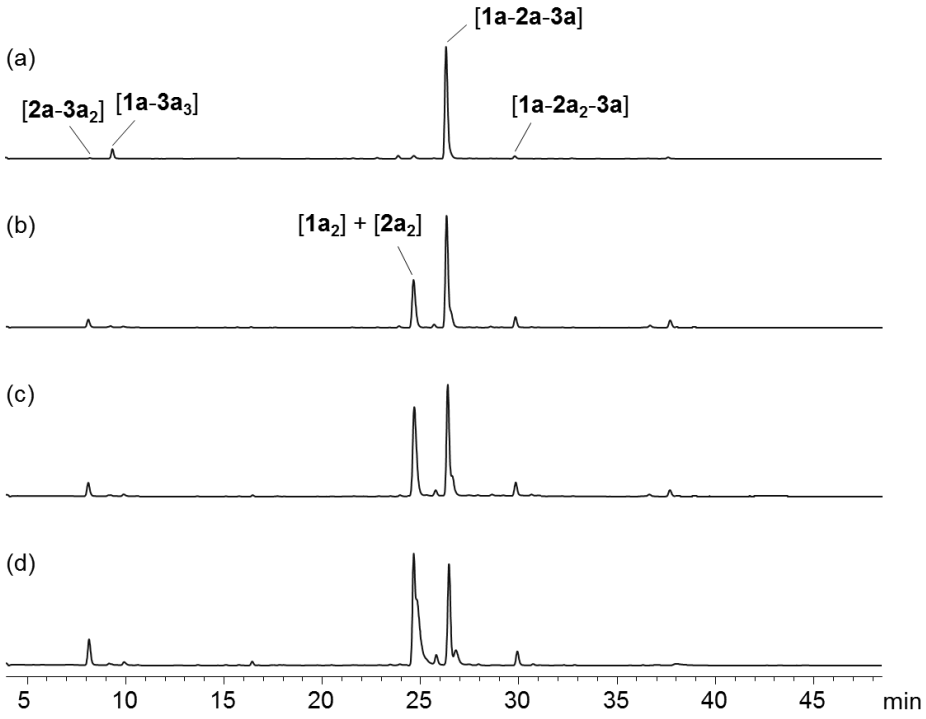


Figure S21: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) and **3a** (2.5 mM), at pH 6.5 with 10% of DMSO at different sodium chloride concentrations: 0.0 M (a), 0.5 M (b), 1.0 M (c) and 2.0 M (d).

Mixture of **1a**+**2a** with different monopodal BBs

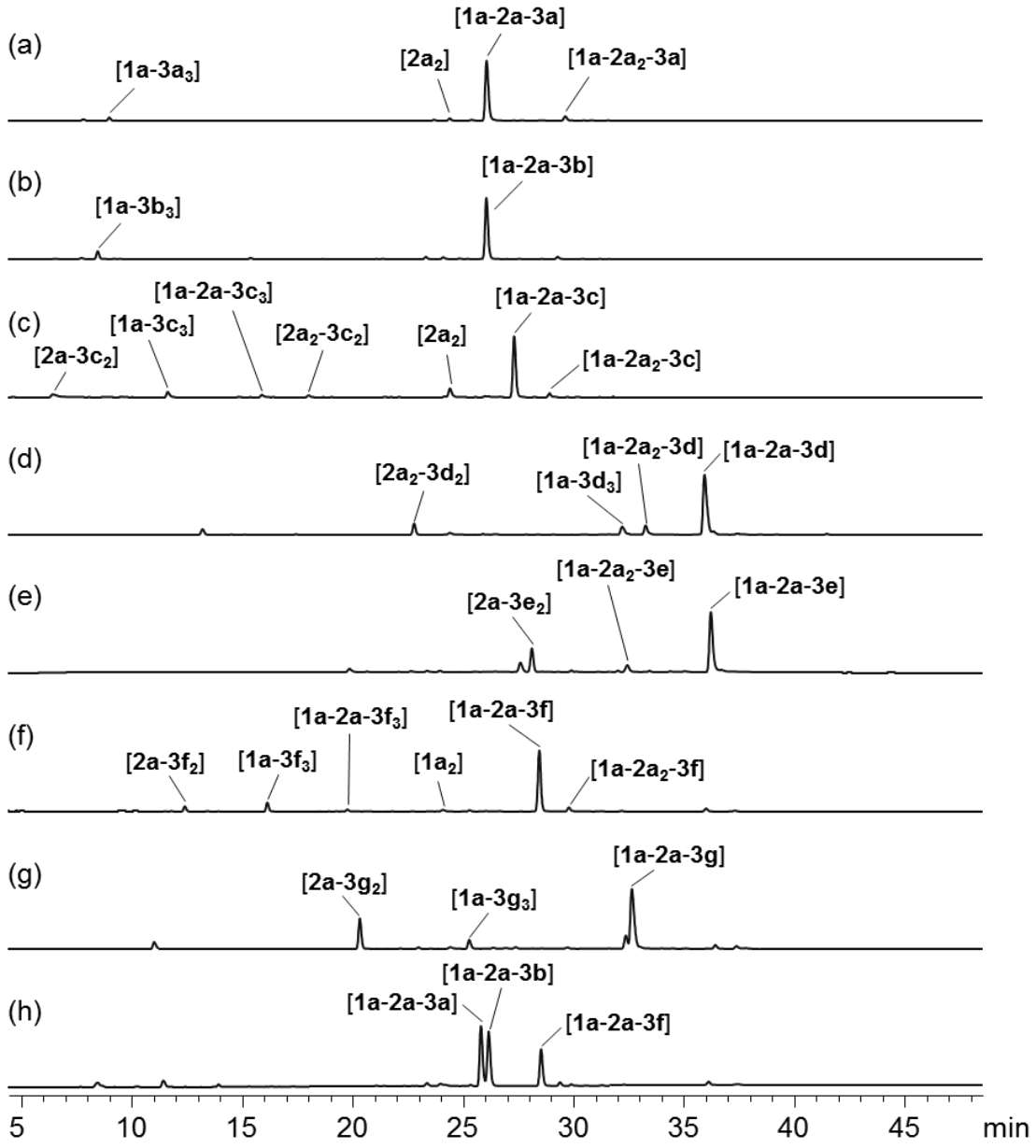


Figure S22: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) in the presence of 2.5 mM of **3a** (a), **3b** (b), **3c** (c), **3d** (d), **3e** (e), **3f** (f), **3g** (g). HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) with **3a**+**3b**+**3f** (1.0 mM each) (h).

Mixture of **1b**+**2a** with different monopodal BBs

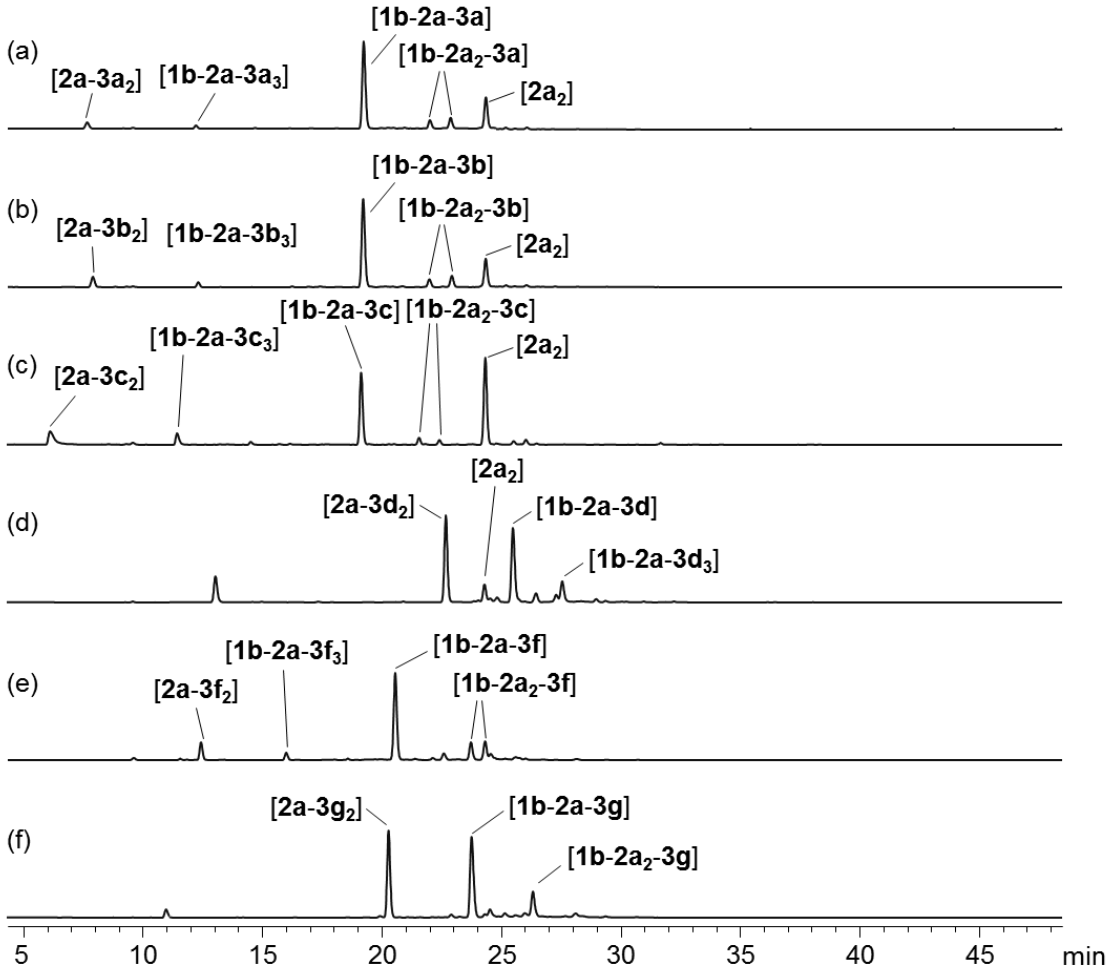


Figure S23: HPLC traces of the mixture of **1b**+**2a** (0.5 mM each) in the presence of 2.5 mM of **3a** (a), **3b** (b), **3c** (c), **3d** (d), **3f** (e) and **3g** (f).

Mixture of **1a**+**2a** with different monopodals (**3a** and **3h** for comparison):

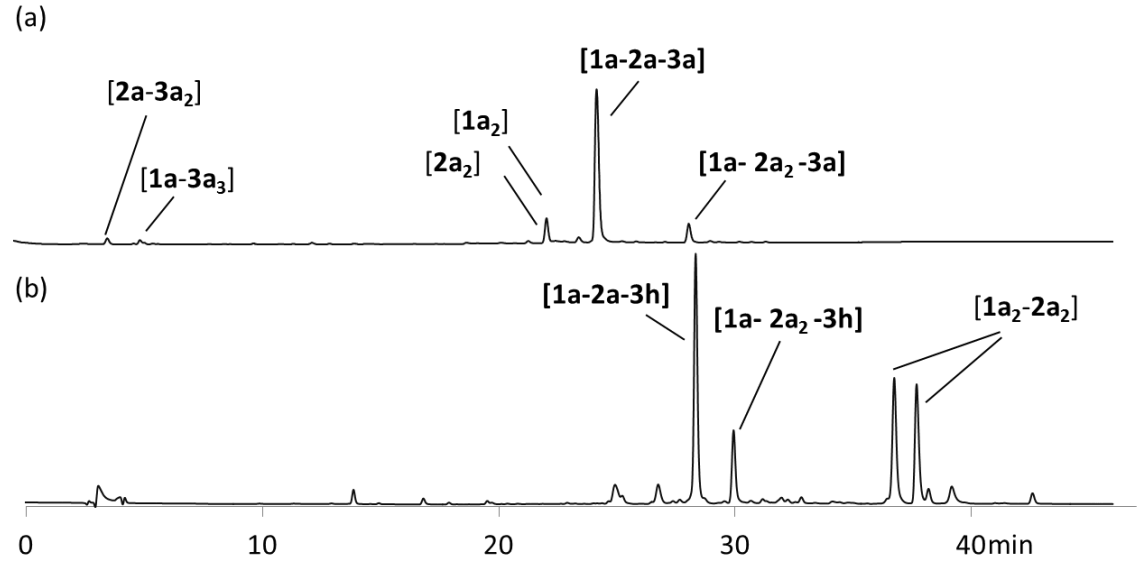


Figure S24: HPLC traces at 25% DMSO and pH 6.5 of the mixture of **1a**+**2a** (0.5 mM each) with different monopodals (2.5mM) **3a** (a), **3h** (b).

Mixture of **1a**+**3a** with different bipodal BBs

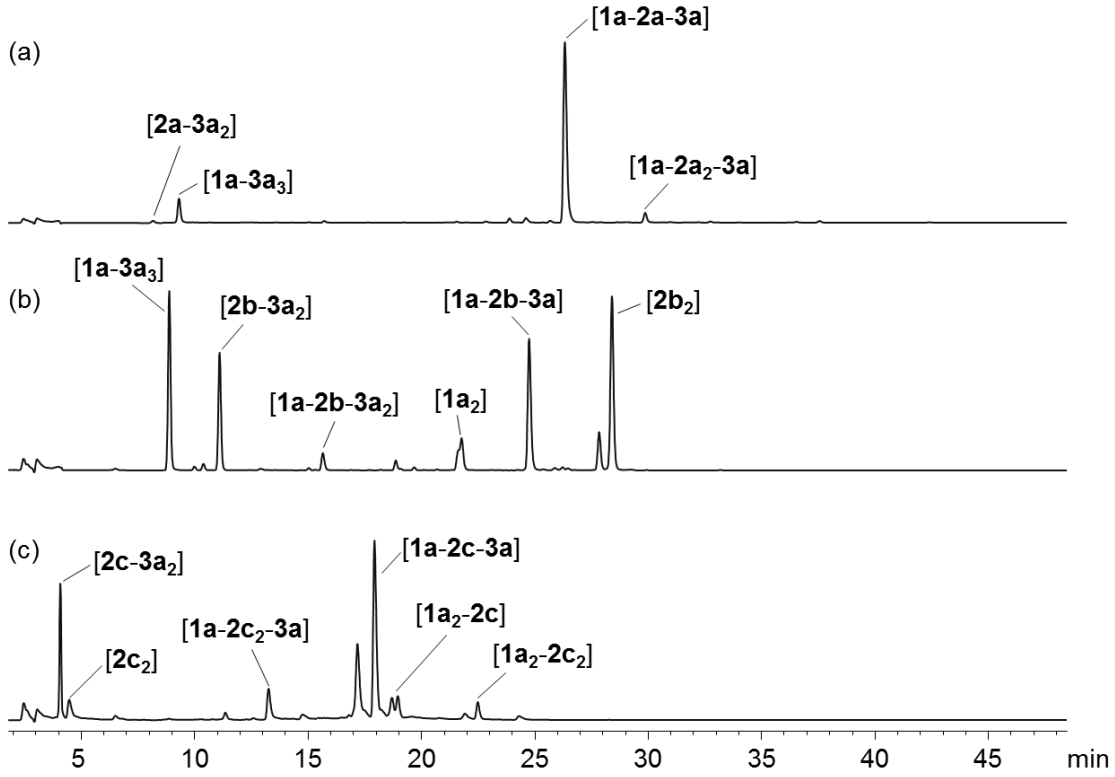


Figure S25: HPLC traces of the mixture of **1a**+**3a** (0.5 mM and 2.5 mM respectively) in the presence of 0.5 mM of **2a** (a), **2b** (b) and **2c** (c), at pH 6.5 with 25% DMSO.

Mixture of **2a**+**3a** with different tripodal BBs

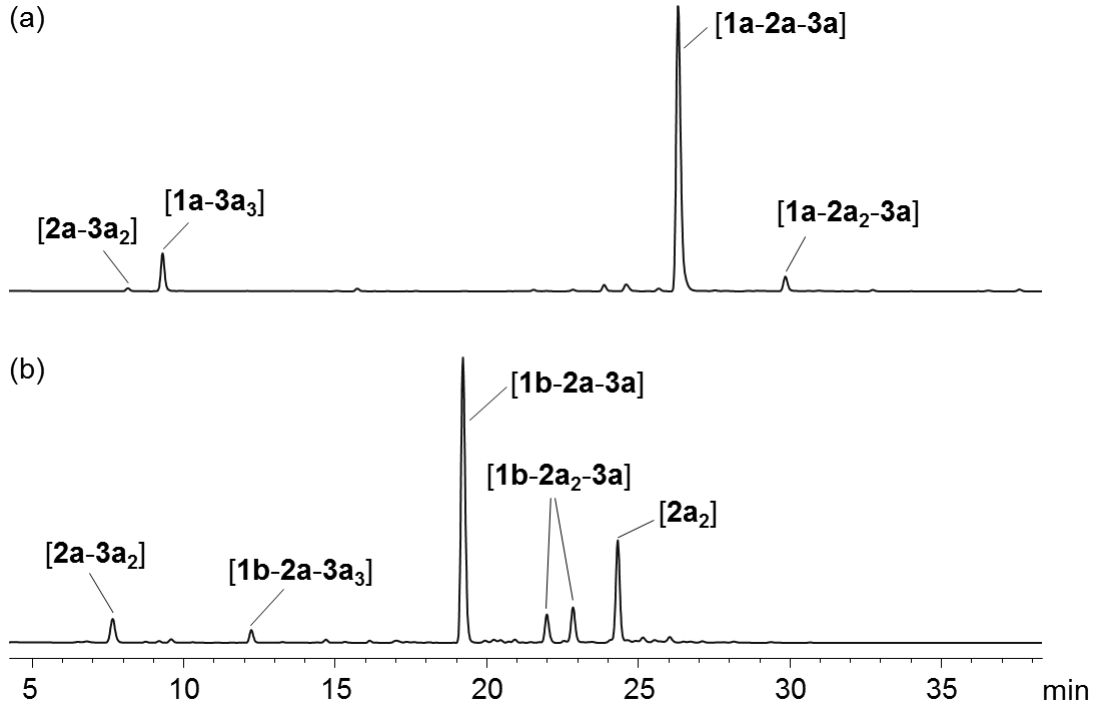


Figure S26: HPLC traces of the mixture of **2a**+**3a** (0.5 mM and 2.5 mM respectively) in the presence of 0.5 mM of **1a** (a) and **1b** (b), at pH 6.5 with 25% DMSO.

Mixture of **1a**+**2a**+**3a** with different percentage of DMSO

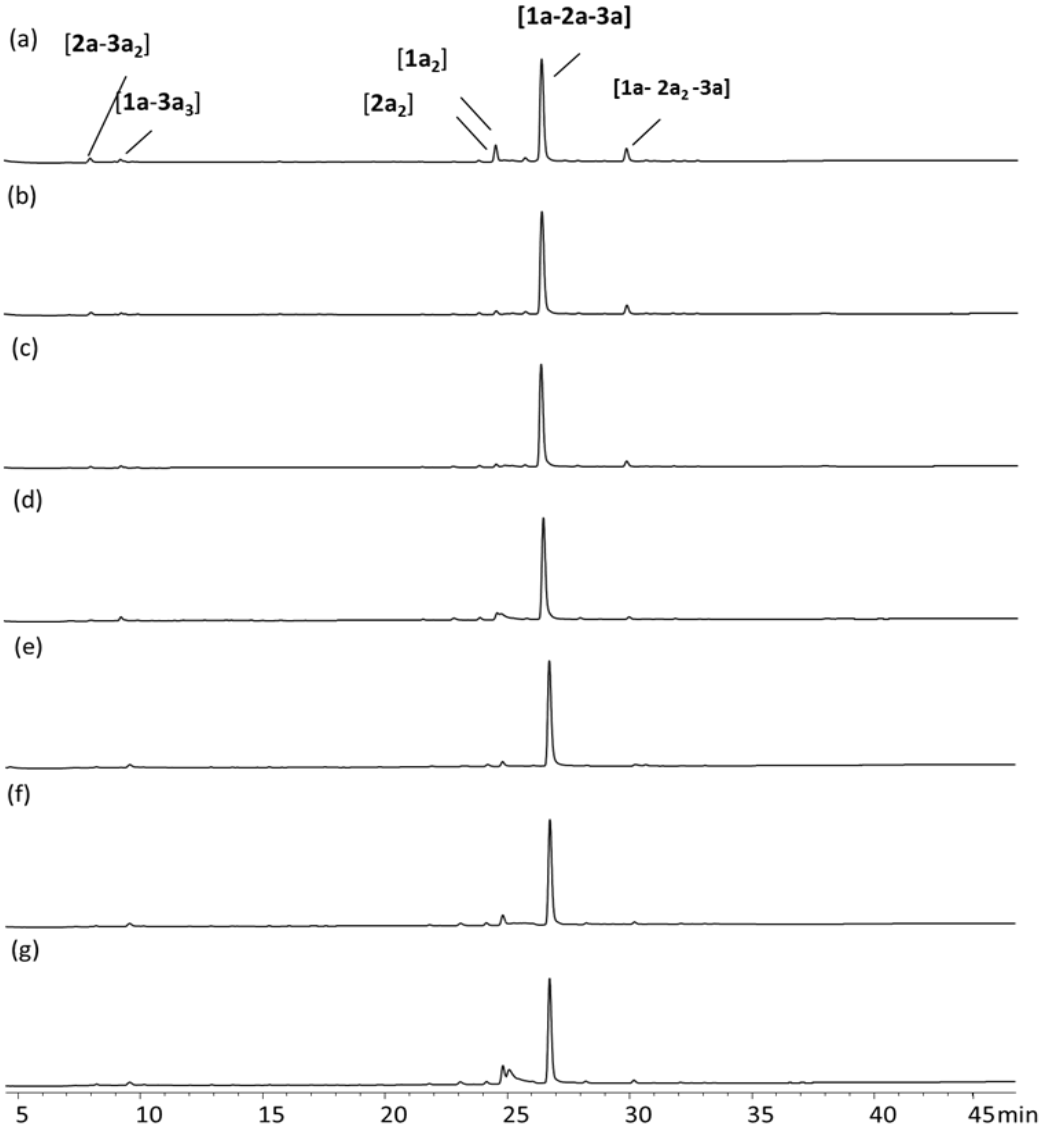


Figure S27: HPLC traces of the mixture of **1a**+**2a** (0.5 mM each) and **3a** (2.5 mM), at pH 6.5, after complate oxidation at different percentage of DMSO: 25 % (a), 10 % (b), 5 % (c), 2.5 % (d), 1.5 % (e), 1 % (f), 0.5 % (g)

Mixture of **1a**, **2a and 3a** at different concentrations:

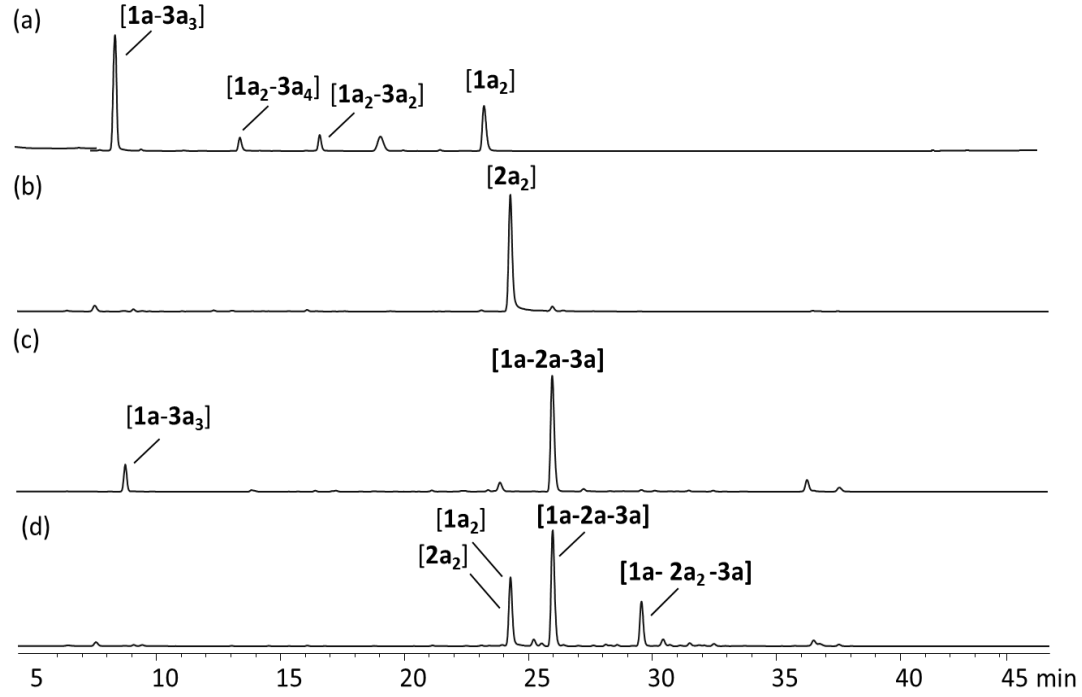


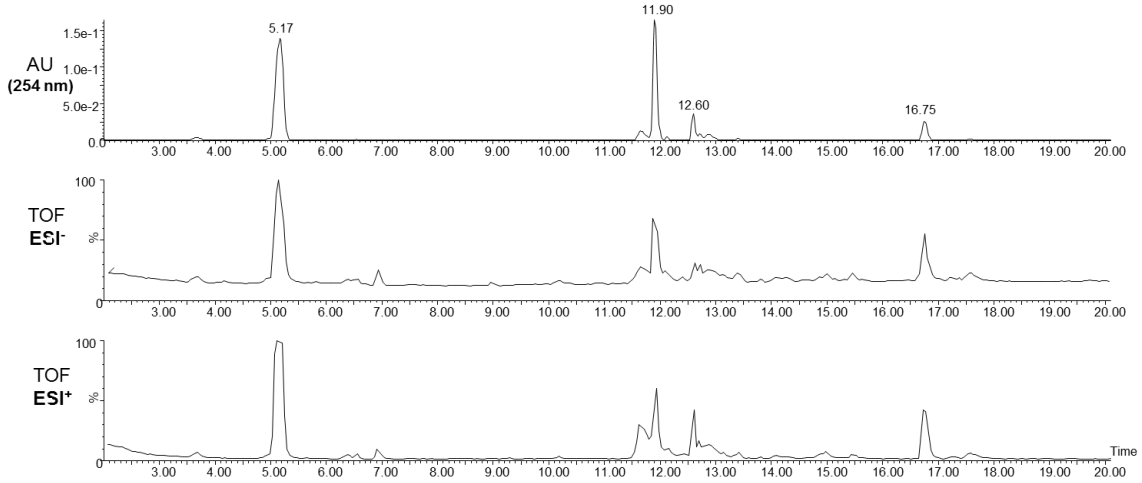
Figure S28: HPLC traces (Bis-tris aqueous buffer at pH 6.5 with 25% of DMSO) of the mixture of **1a** (0.5 mM) and **3a** (50 mM) (a), mixture of **2a** (0.5 mM) and **3a** (50 mM) (b), mixture of **1a+2a**(1.0 mM, 0.5 respectively) and **3a** (0.5 mM) (c), mixture of **1a+2a** (0.5 mM, 1.0 respectively) and **3a** (0.5 mM) (d).

MASS SPECTROMETRY

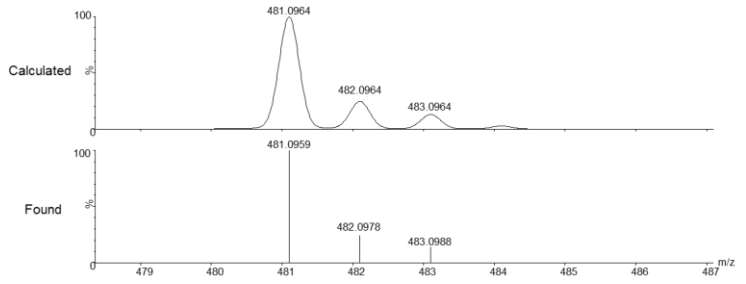
General procedure for the analysis of the DCLs by HRMS

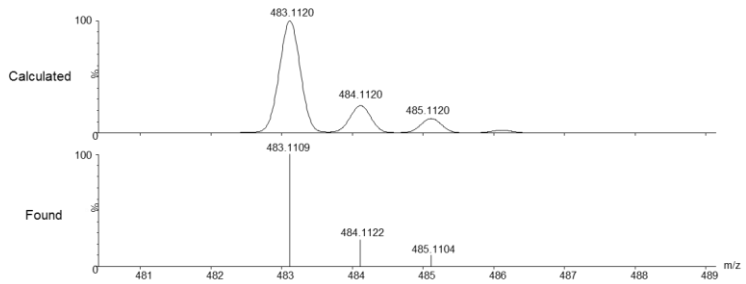
The HRMS (UPLC-ESI-TOF) samples were prepared by adding 20 μL of the corresponding reaction mixture to 40 μL of a solution of 89% H_2O, 10% MeCN and 1% TFA. Eluent used: 2.5 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 50% CH_3CN over 27.5 min. The individual MS for the identification of each compound show ESI negative and positive mode with the correspondent calculated and recorded isotope distributions.

Mixture of **1a**+**2a** (0.5 mM each, pH.6.5, *t* = 1min)

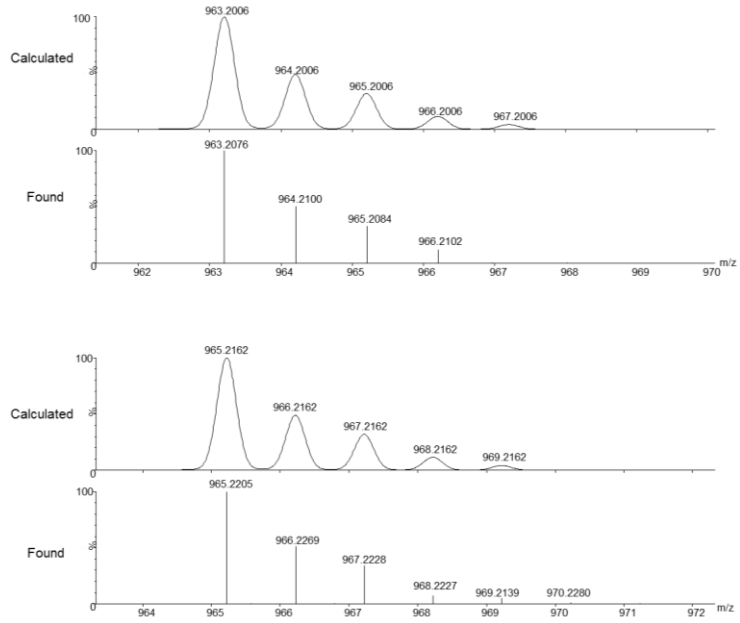


[**2a**]
Retention time: 5.17 min.
Chemical Formula: $C_{18}H_{22}N_6O_6S_2$
Exact Mass: 482.1042

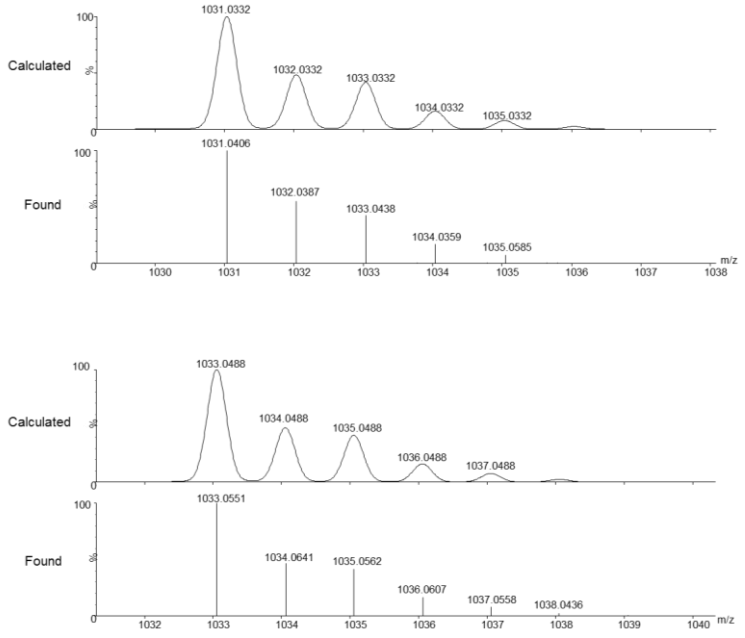




[**2a$_2$**]
Retention time: 11.90 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

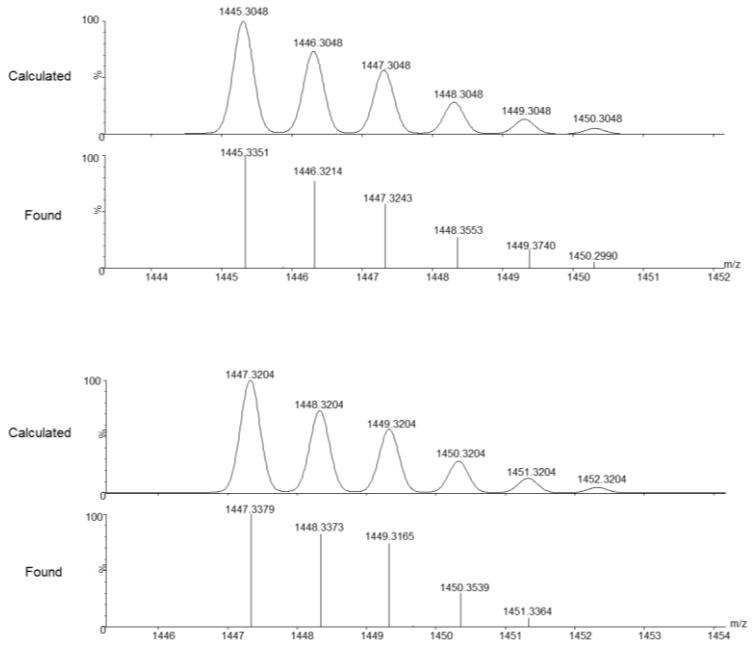


[**1a$_2$**]
Retention time: 11.90 min.
Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$
Exact Mass: 1032.0410

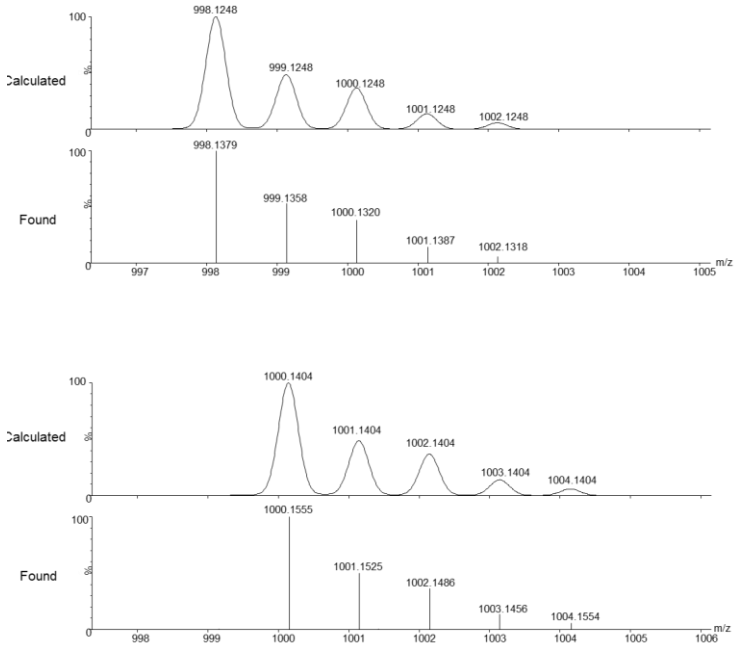


[**2a$_3$**]

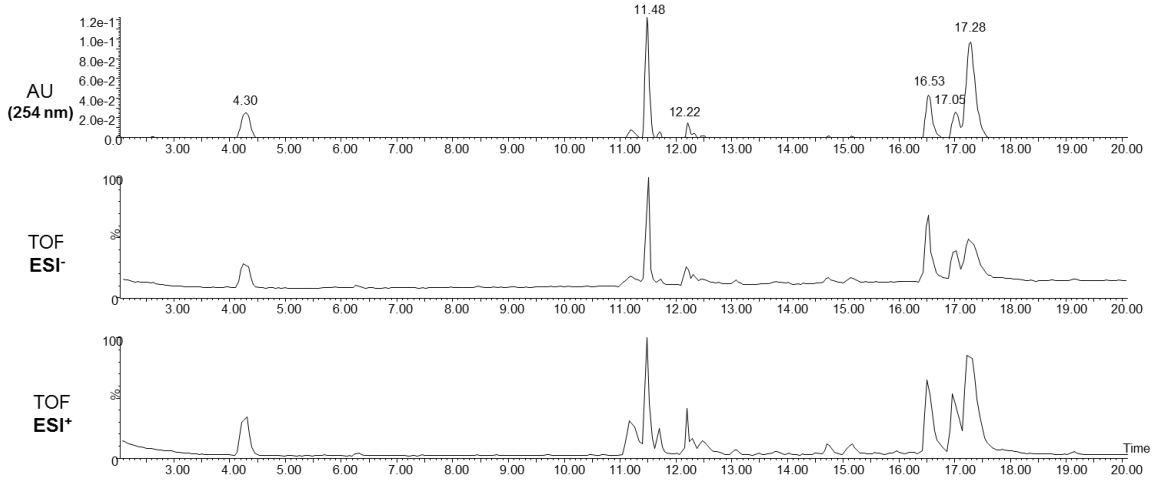
Retention time: 12.60 min.
Chemical Formula: $C_{54}H_{66}N_{18}O_{18}S_6$
Exact Mass:
1446.3127



[1a-2a-SH]
Retention time: 16.75 min.
Chemical Formula: $C_{36}H_{41}N_9O_{15}S_5$
Exact Mass:
999.1326



Mixture of **1a**+**2a** (0.5 mM each, pH.6.5, *t* = 2 h)



Identification of the products:

[**2a**] (previously identified)
Retention time: 4.30 min.

Chemical Formula: $C_{18}H_{22}N_6O_6S_2$
Exact Mass:
482.1042

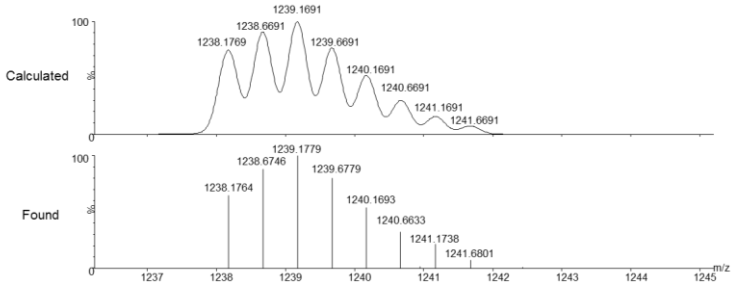
[**2a$_2$**] (previously identified)
Retention time: 11.48 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass:
964.2084

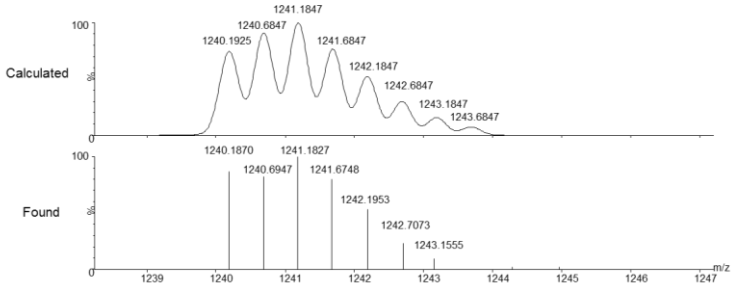
[**1a$_2$**] (previously identified)
Retention time: 11.48 min.
Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$
Exact Mass:
1032.0410

[**2a$_3$**] (previously identified)
Retention time: 12.22 min.
Chemical Formula: $C_{54}H_{66}N_{18}O_{18}S_6$
Exact Mass:
1446.3127

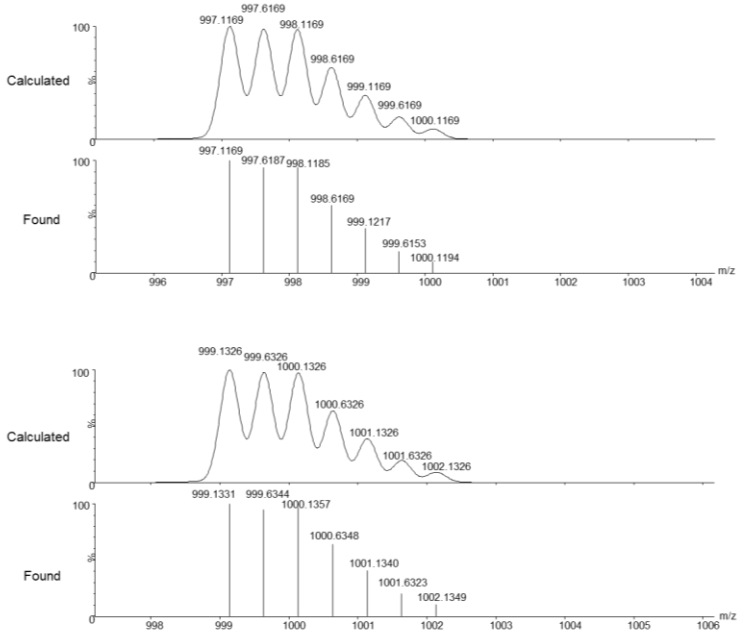
[**1a-2a-SH**] (previously identified)
Retention time: 16.53 min.
Chemical Formula: $C_{36}H_{41}N_9O_{15}S_5$
Exact Mass:
999.1326

[**1a$_2$-2a$_3$**]
Retention time: 17.05 min.
Chemical Formula: $C_{90}H_{102}N_{24}O_{36}S_{12}$
Exact Mass:
2478.3537

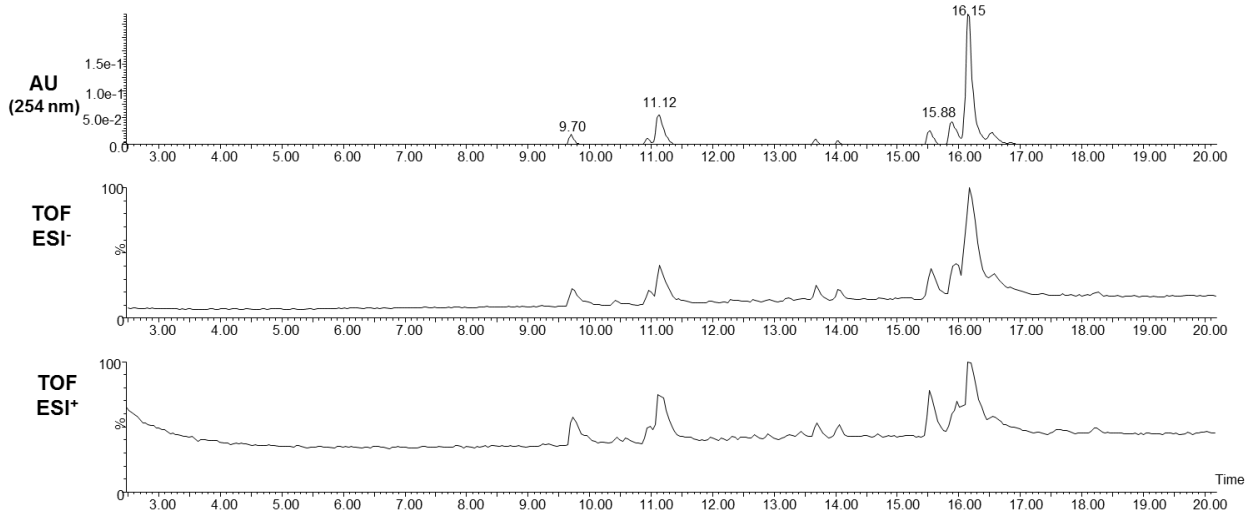




[**1a$_2$-2a$_2$**]
Retention time: 17.28 min.
Chemical Formula: $C_{72}H_{80}N_{18}O_{30}S_{10}$
Exact Mass: 1996.2495



Mixture of **1a**+**2a** (0.5 mM each, pH.6.5, *t* = 24 h)



Identification of the products:

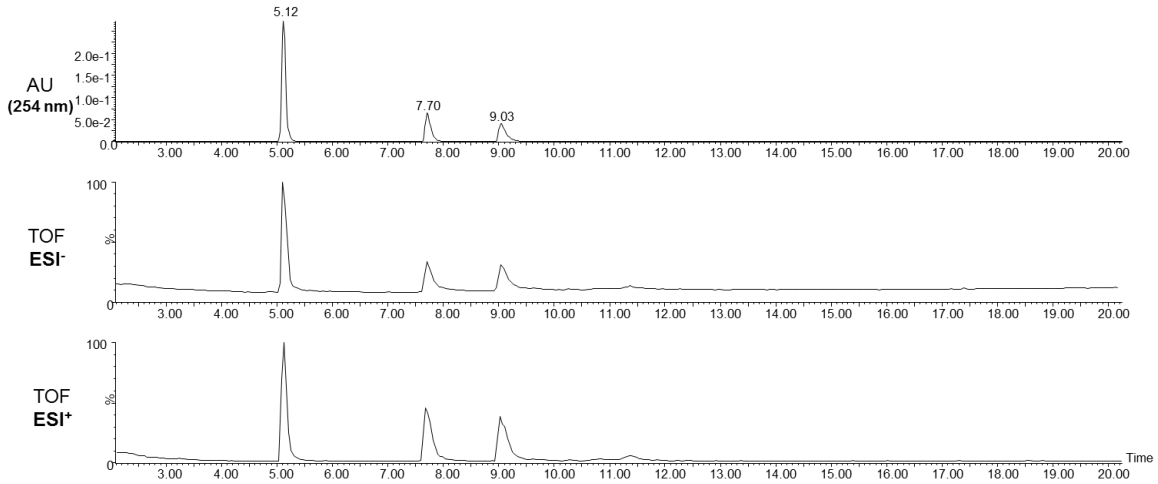
[**1a$_2$**] (previously identified)
Retention time: 9.70 min.
Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$
Exact Mass: 1032.0410

[**1a$_2$-2a-SH**] (previously identified)
Retention time: 11.12 min.
Chemical Formula: $C_{54}H_{58}N_{12}O_{24}S_8$
Exact Mass: 1514.1453

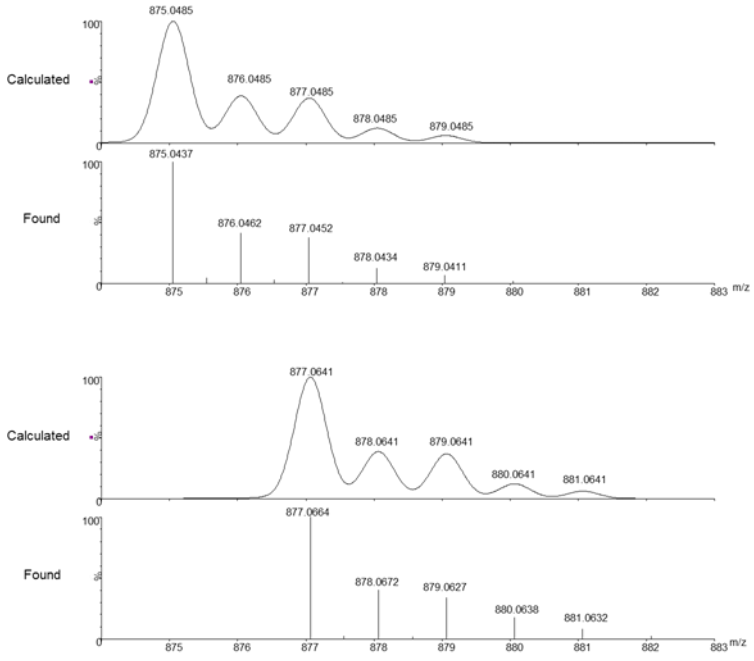
[1a$_2$-2a$_3$] (previously identified)
Retention time: 15.88 min.
Chemical Formula: $C_{90}H_{102}N_{24}O_{36}S_{12}$
Exact Mass: 2478.3537

[**1a$_2$-2a$_2$**] (previously identified)
Retention time: 16.17 min.
Chemical Formula: $C_{72}H_{80}N_{18}O_{30}S_{10}$
Exact Mass: 1996.2495

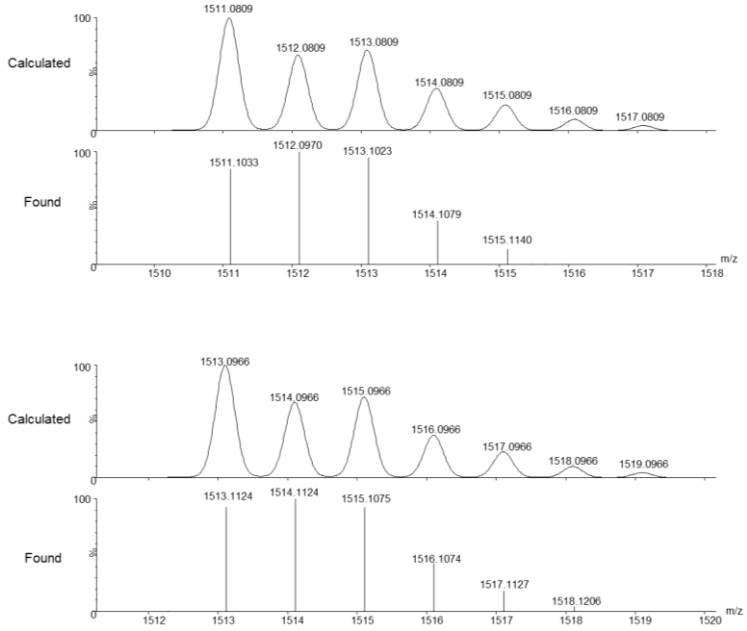
Mixture of **1a**+**3a** (0.5 and 2.5 mM respectively, pH.6.5, *t* = 1 min)



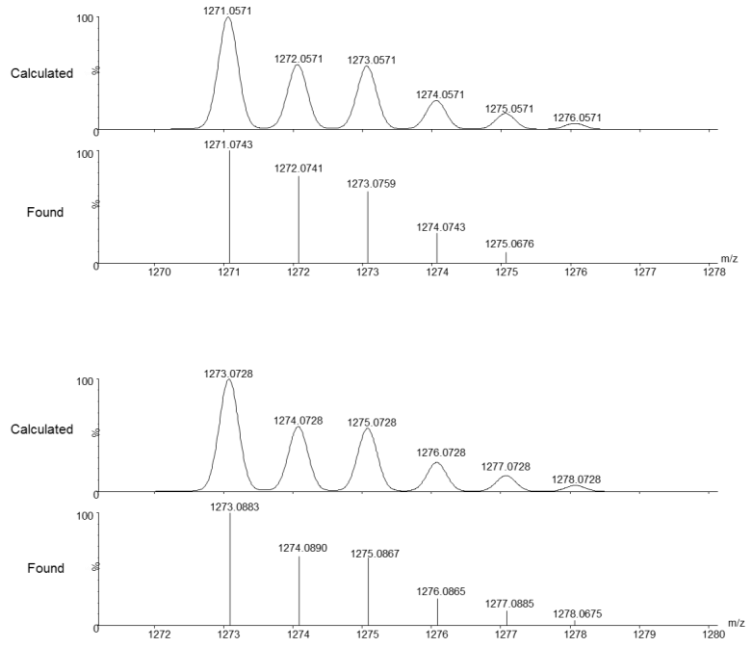
Identification of the products:
[**1a-3a$_3$**]
Retention time: 5.12 min.
Chemical Formula: $C_{27}H_{36}N_6O_{15}S_6$
Exact Mass: 876.0563



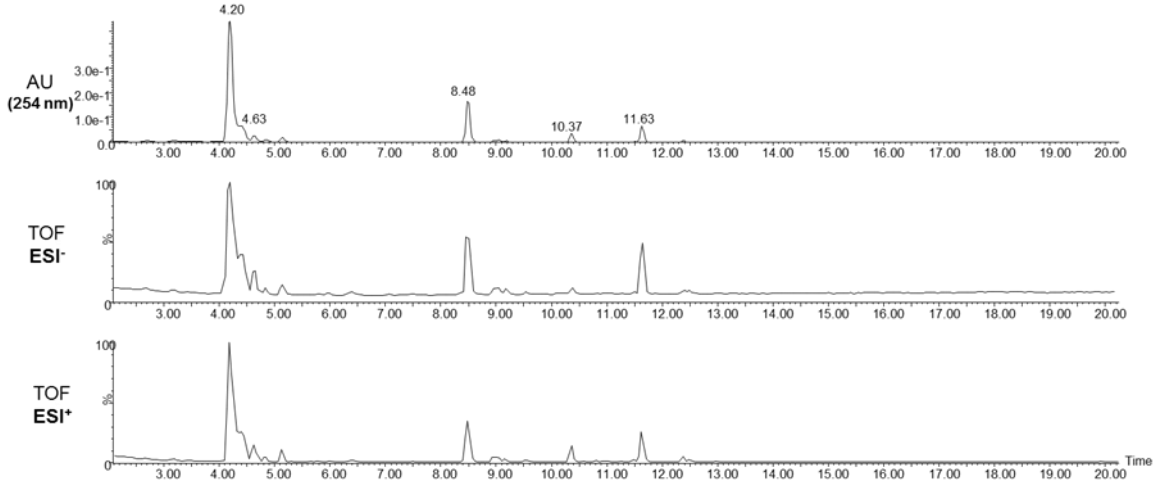
[**1a$_2$-3a$_4$**]
Retention time: 7.70 min.
Chemical Formula: $C_{48}H_{60}N_{10}O_{26}S_{10}$
Exact Mass: 1512.0887



[**1a$_2$-3a$_2$**]
Retention time: 9.03 min.
Chemical Formula: $C_{42}H_{48}N_8O_{22}S_8$
Exact Mass: 1272.0649

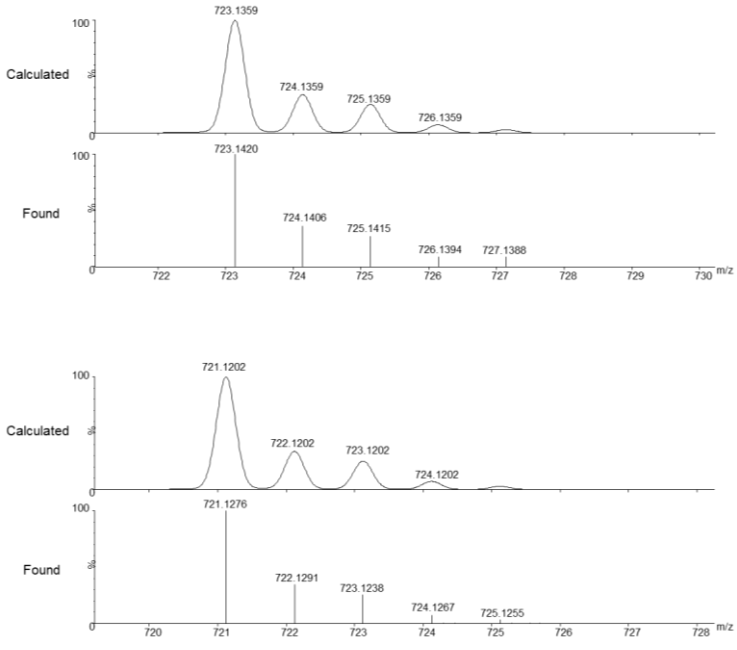


Mixture of **2a**+**3a** (0.5 and 2.5 mM respectively, pH.6.5, *t* = 1 min)



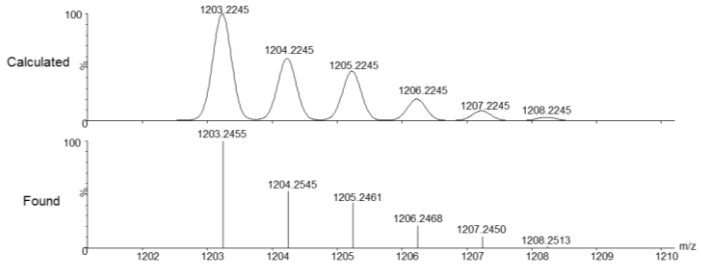
Identification of the products:

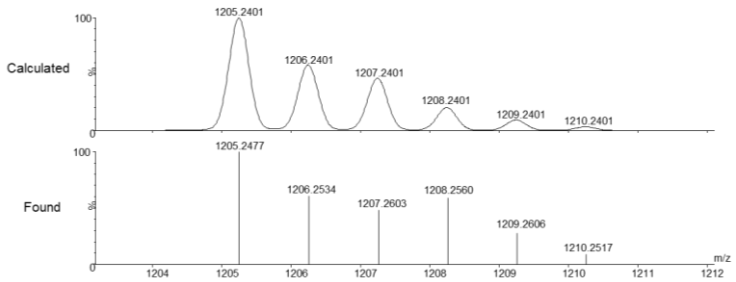
[2a-3a$_2$]
Retention time: 4.20 min.
Chemical Formula: $C_{24}H_{34}N_8O_{10}S_4$
Exact Mass: 722.1281



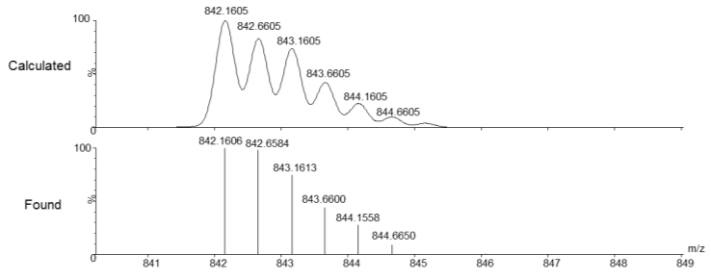
[**2a**] (previously identified)
Retention time: 4.63 min.
Chemical Formula: $C_{18}H_{22}N_6O_6S_2$
Exact Mass: 482.1042

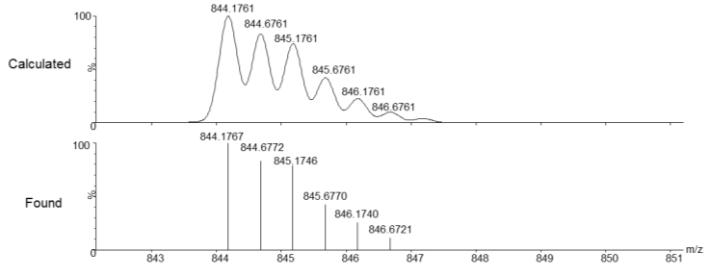
[**2a$_2$-3a$_2$**]
Retention time: 8.48 min.
Chemical Formula: $C_{42}H_{56}N_{14}O_{16}S_6$
Exact Mass: 1204.2323





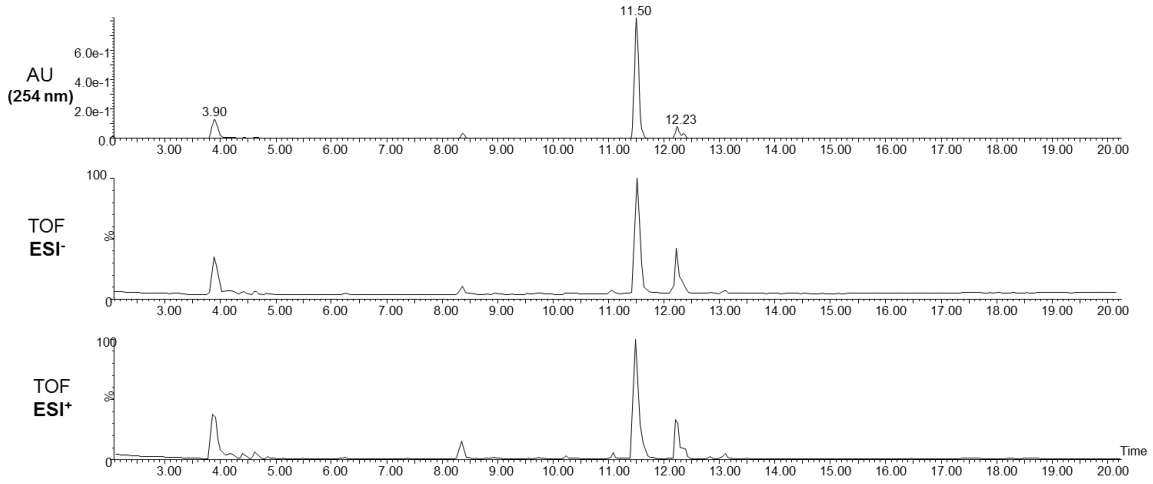
[2a$_3$-3a$_2$]
Retention time: 10.37 min.
Chemical Formula: $C_{60}H_{78}N_{20}O_{22}S_8$
Exact Mass: 1686.3365





[**2a$_2$**] (previously identified)
Retention time: 11.63 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

Mixture of **2a**+**3a** (0.5 and 2.5 mM respectively, pH.6.5, *t* = 2 h)



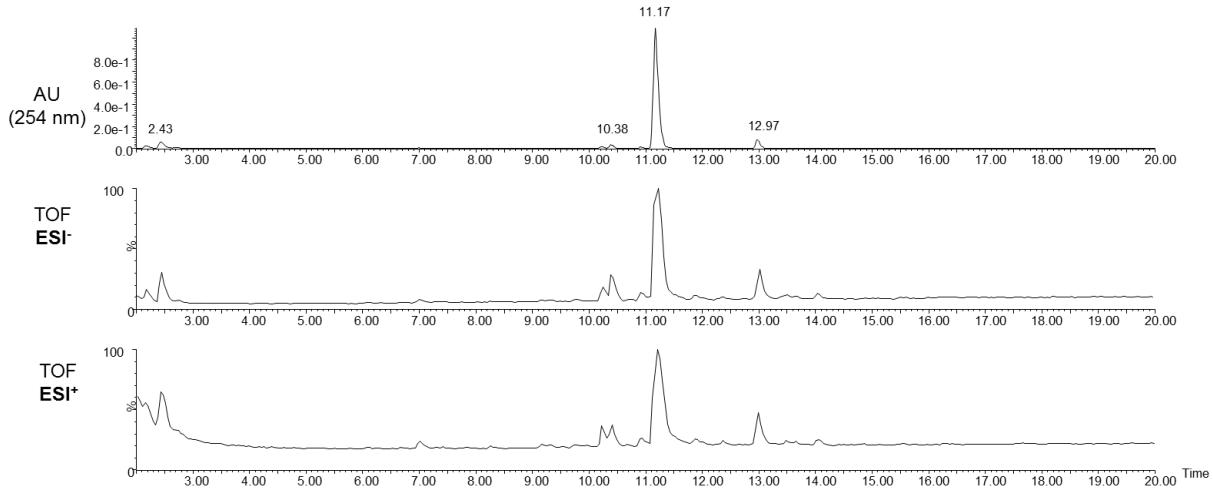
Identification of the products:

[**2a-3a$_2$**] (previously identified)
Retention time: 3.90 min.
Chemical Formula: $C_{24}H_{34}N_8O_{10}S_4$
Exact Mass: 722.1281

[**2a$_2$**] (previously identified)
Retention time: 11.50 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[**2a$_3$**] (previously identified)
Retention time: 12.23 min.
Chemical Formula: $C_{54}H_{66}N_{18}O_{18}S_6$
Exact Mass: 1446.3127

Mixture of **1a**+**2a**+**3a** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

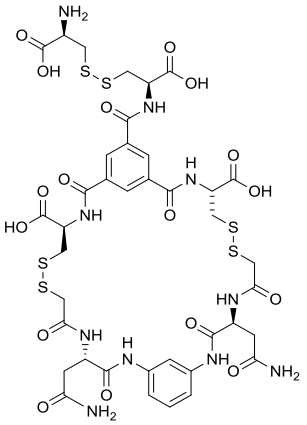


Identification of the products:

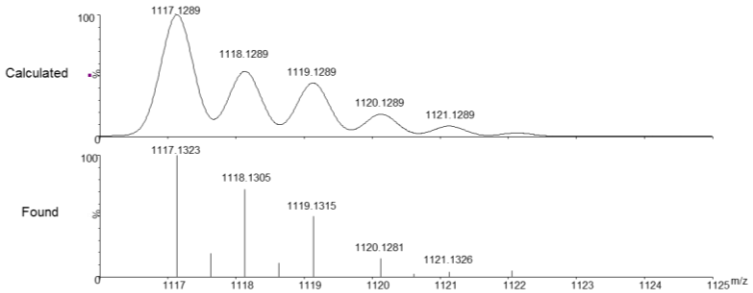
[**1a-3a$_3$**] (previously identified)
Retention time: 2.43 min.
Chemical Formula: $C_{27}H_{36}N_6O_{15}S_6$
Exact Mass: 876.0563

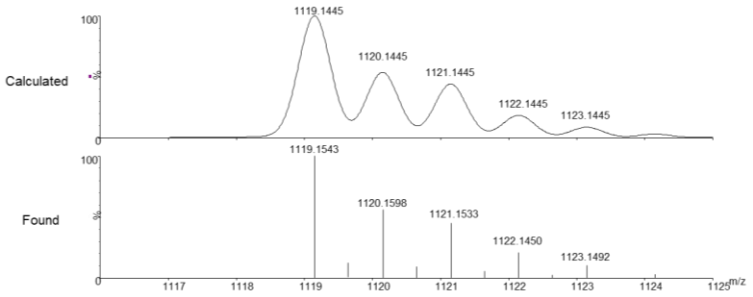
[**2a$_2$**] (previously identified)
Retention time: 10.38 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass:964.2084

[**1a-2a-3a**]



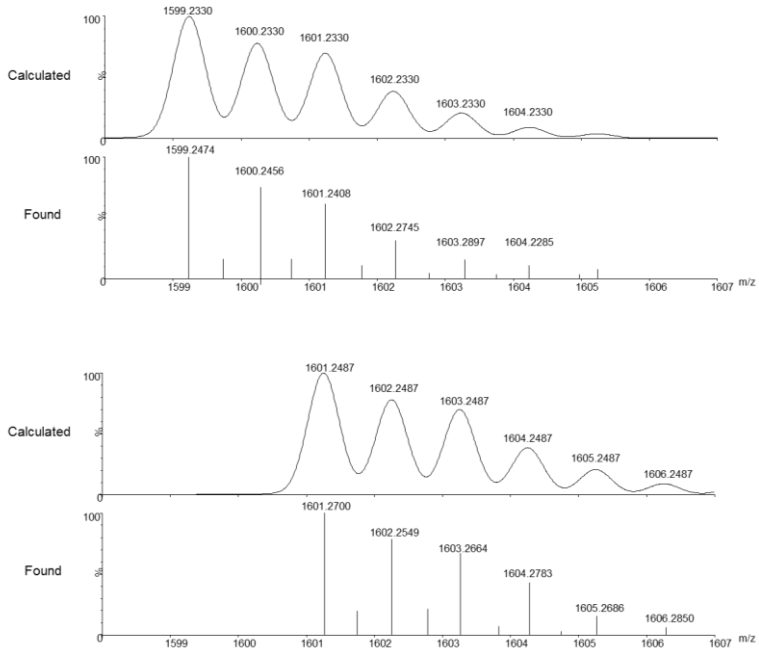
Retention time: 11.17 min.
Chemical Formula: $C_{39}H_{46}N_{10}O_{17}S_6$
Exact Mass: 1118.1367



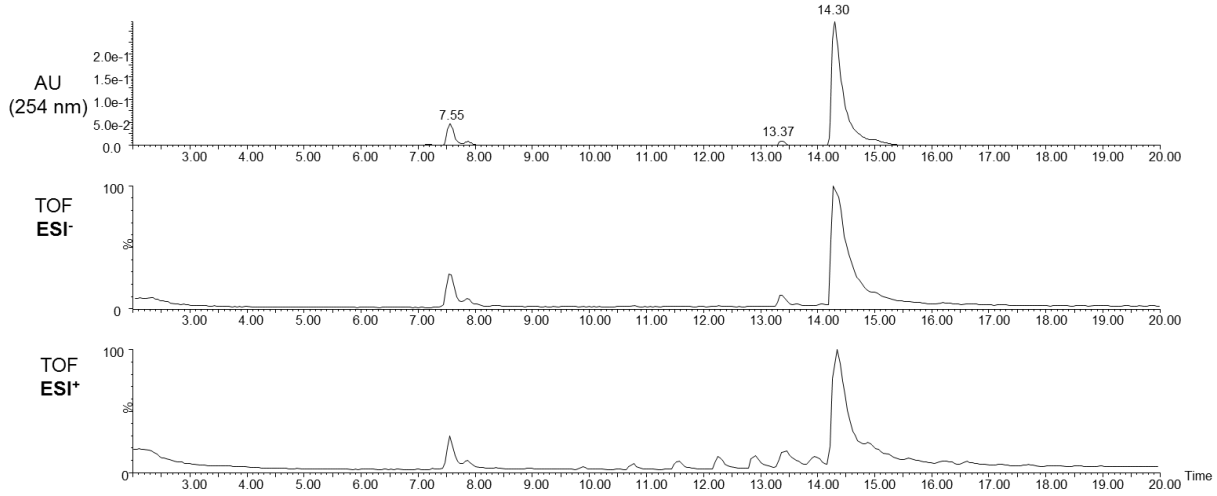


[**1a-2a$_2$-3a**]

Retention time: 12.37 min.
Chemical Formula: $C_{57}H_{68}N_{16}O_{23}S_8$
Exact Mass: 1600.2409

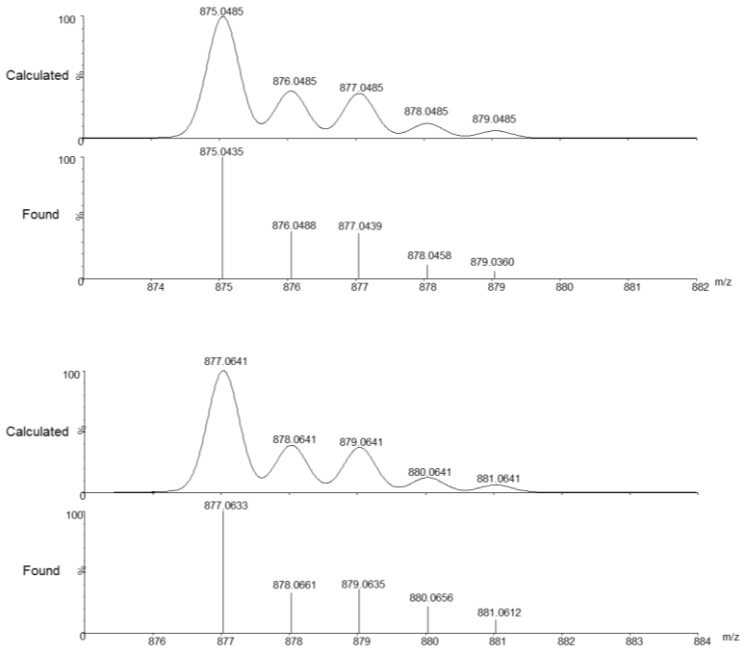


Mixture of **1a**+**2a**+**3b** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)



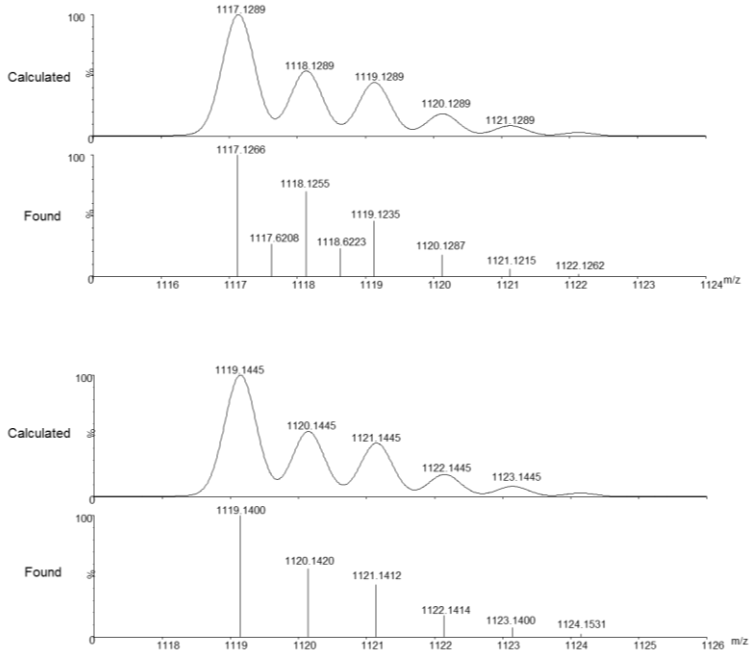
Identification of the products:

[**1a-3b$_3$**]
Retention time: 7.55 min.
Chemical Formula: $C_{27}H_{36}N_6O_{15}S_6$
Exact Mass: 876.0563

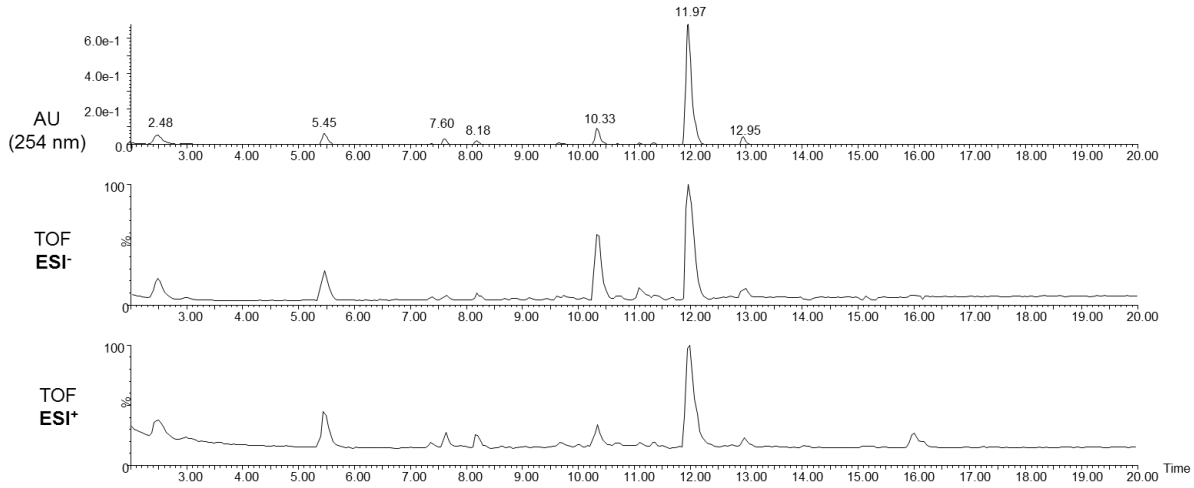


[**2a$_2$**] (previously identified)
Retention time: 13.37 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[**1a-2a-3b**]
Retention time: 14.30 min.
Chemical Formula: $C_{39}H_{46}N_{10}O_{17}S_6$
Exact Mass: 1118.1367

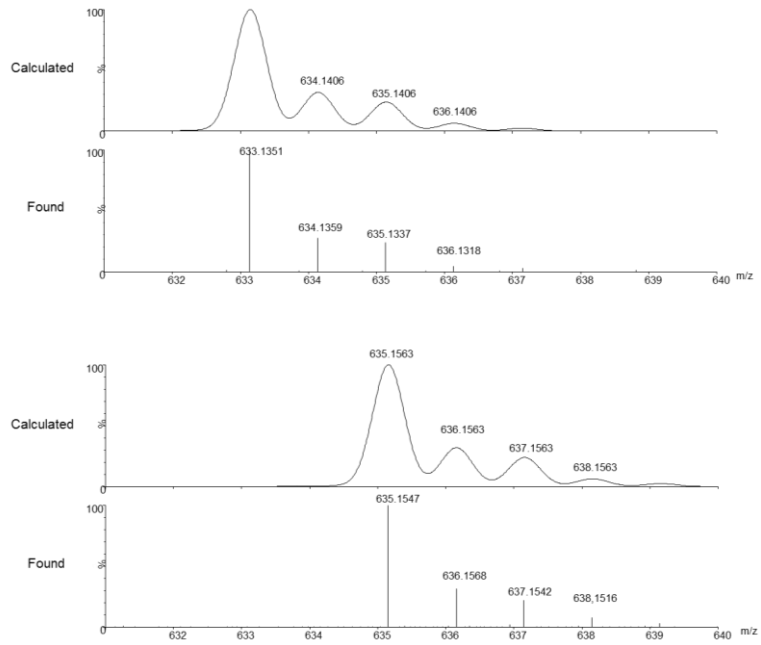


Mixture of **1a**+**2a**+**3c** (0.5, 0.5 and 2.5 mM, pH.6.5, *t* = 24 h)

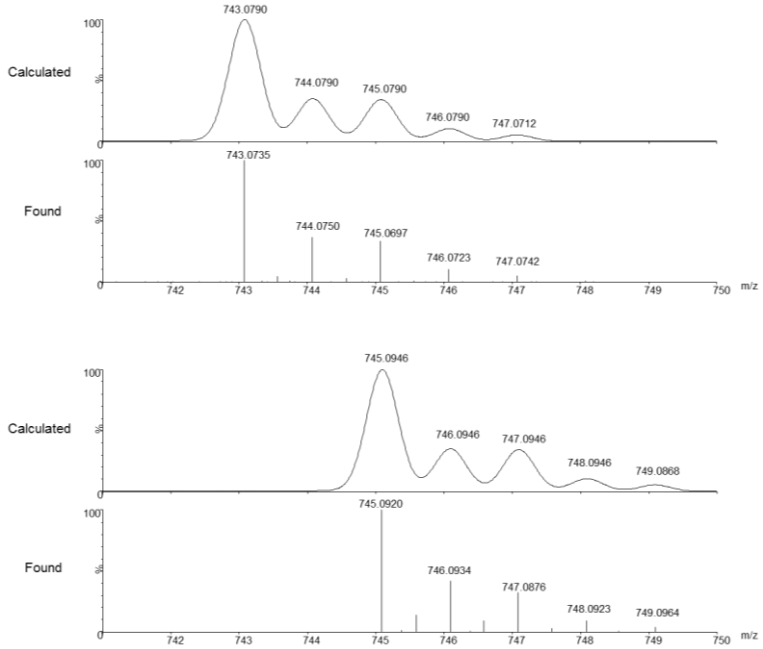


Identification of the products:

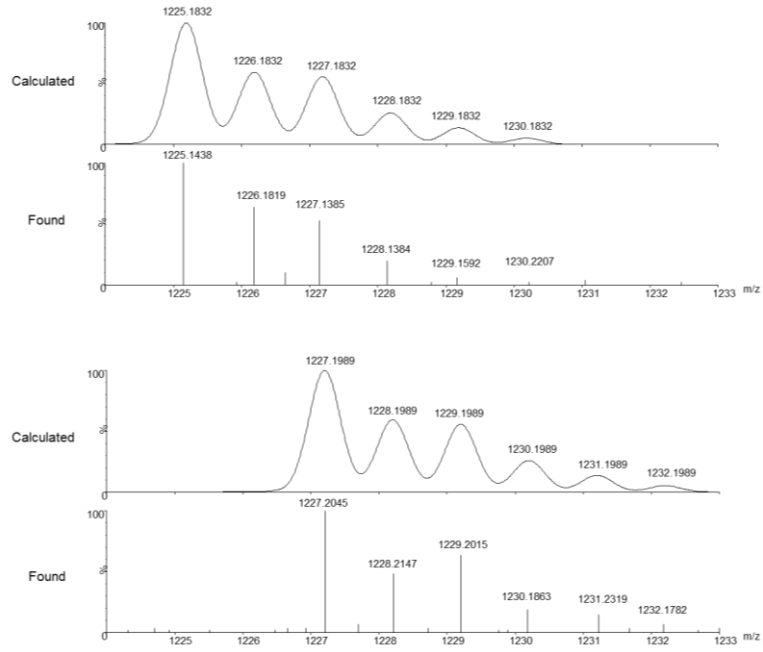
[**2a-3c$_2$**]
Retention time: 2.48 min.
Chemical Formula: $C_{22}H_{34}N_8O_6S_4$
Exact Mass: 634.1484



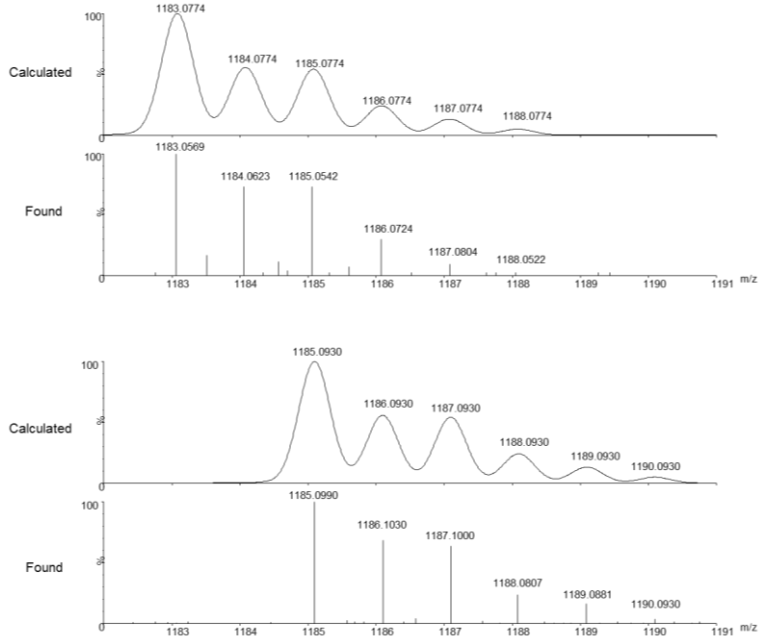
[**1a-3c$_3$**]
Retention time: 5.45 min.
Chemical Formula: $C_{24}H_{36}N_6O_9S_6$
Exact Mass: 744.0868



[**1a-2a-3c$_3$**]
Retention time: 7.60 min.
Chemical Formula: $C_{42}H_{58}N_{12}O_{15}S_8$
Exact Mass: 1226.1910

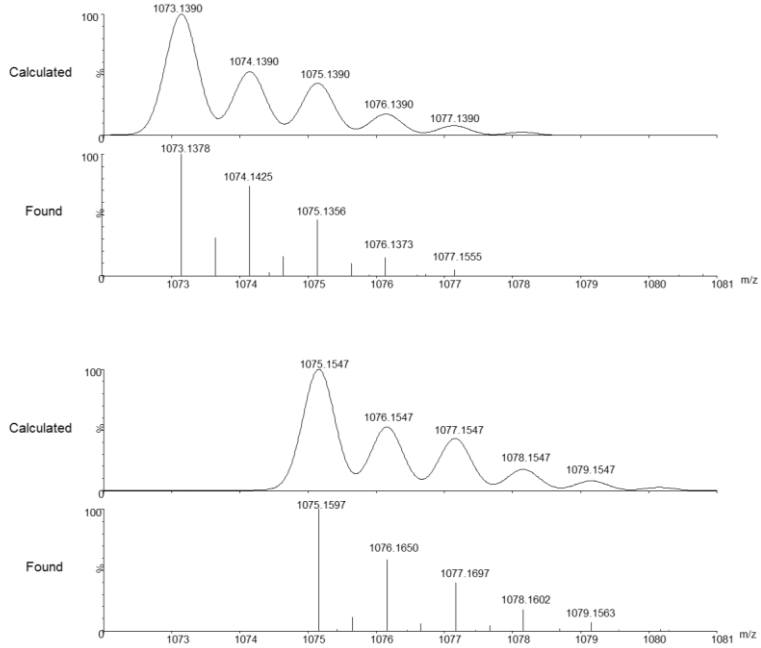


[**1a$_2$-3c$_2$**]
Retention time: 8.18 min.
Chemical Formula: $C_{40}H_{48}N_8O_{18}S_8$
Exact Mass: 1184.0852



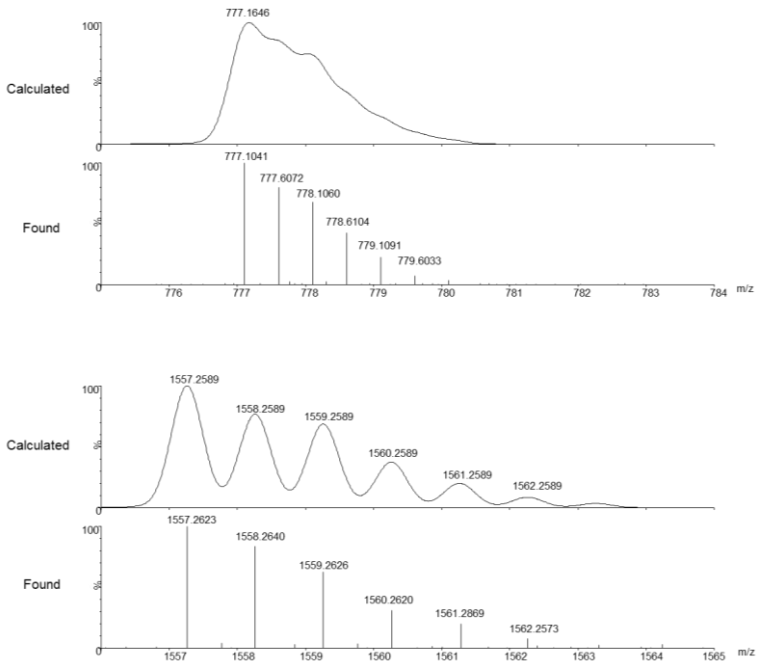
[**2a$_2$**] (previously identified)
Retention time: 10.32 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[**1a-2a-3c**]
Retention time: 11.97 min.
Chemical Formula: $C_{38}H_{46}N_{10}O_{15}S_6$
Exact Mass: 1074.1468

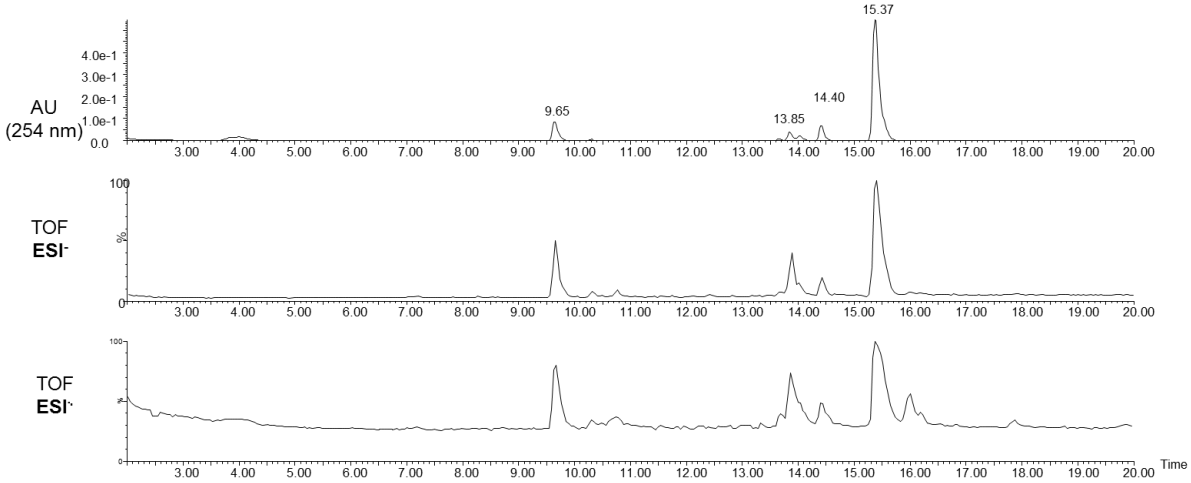


[**1a-2a$_2$-3c**]

Retention time: 12.95 min.
Chemical Formula: $C_{56}H_{68}N_{16}O_{21}S_8$
Exact Mass: 1556.2511

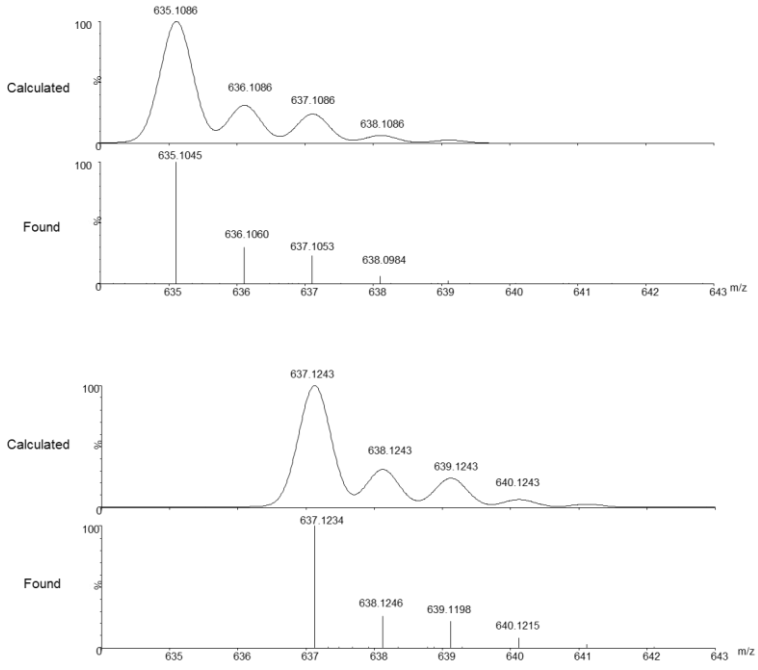


Mixture of **1a**+**2a**+**3d** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

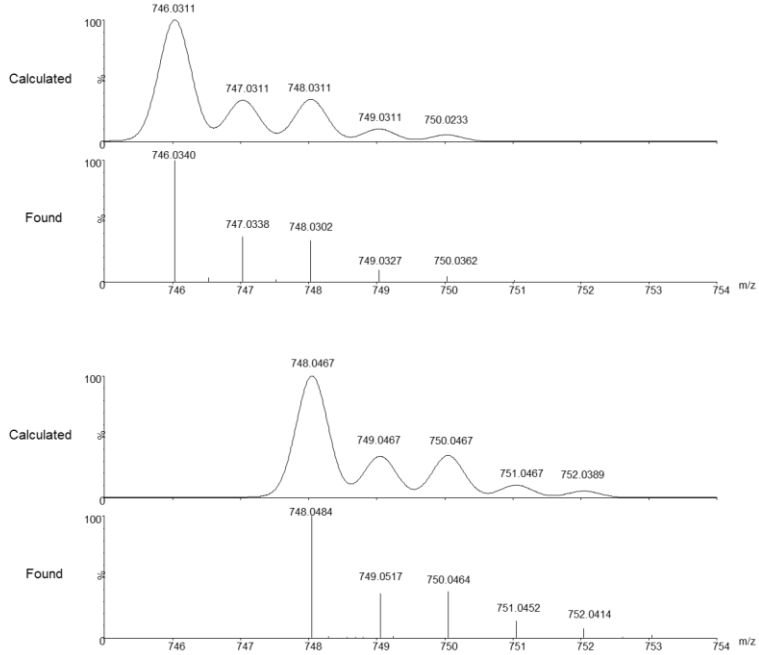


Identification of the products:

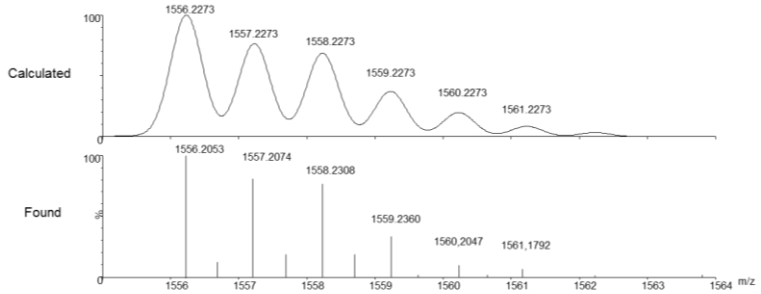
[**2a-3d$_2$**]
Retention time: 9.65 min.
Chemical Formula: $C_{22}H_{32}N_6O_8S_4$
Exact Mass: 636.1164

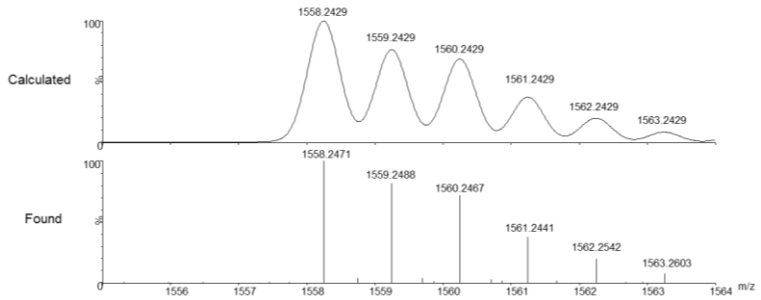


[**1a-3d$_3$**]
Retention time: 13.85 min.
Chemical Formula: $C_{24}H_{33}N_3O_{12}S_6$
Exact Mass: 747.0388

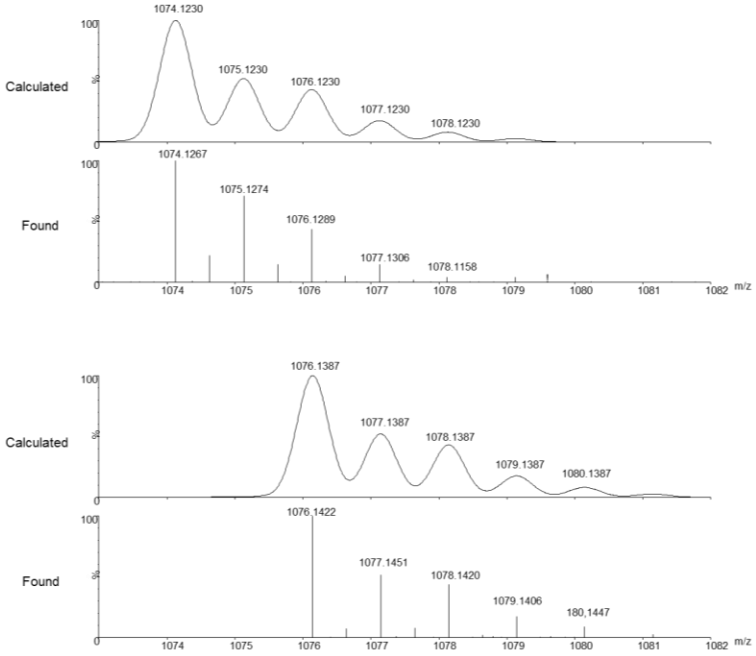


[**1a-2a$_2$-3d**]
Retention time: 14.40 min.
Chemical Formula:
$C_{56}H_{67}N_{15}O_{22}S_8$
Exact Mass:
1557.2351

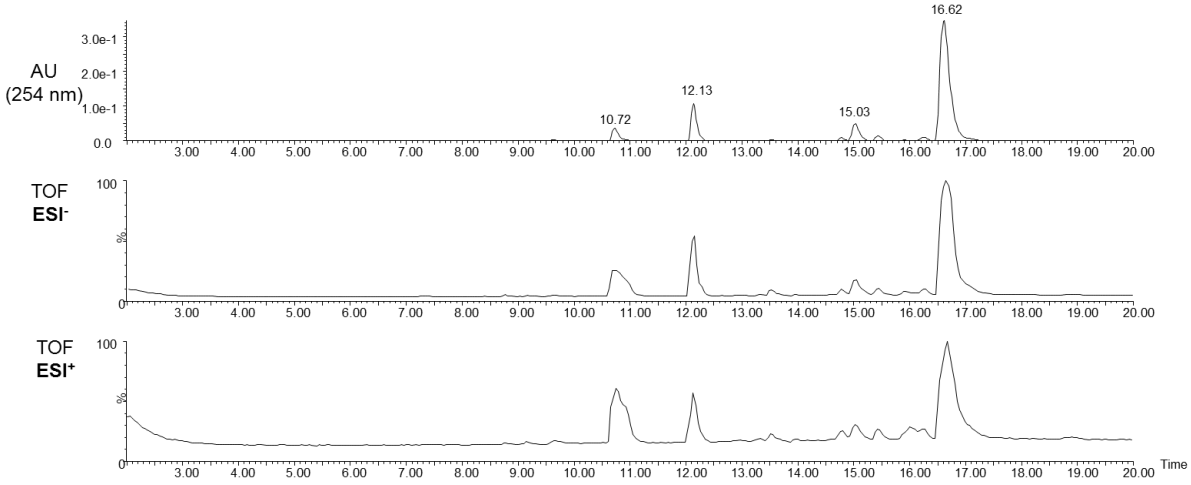




[**1a-2a-3d**]
Retention time: 15.37 min.
Chemical Formula: $C_{38}H_{45}N_9O_{16}S_6$
Exact Mass: 1075.1309

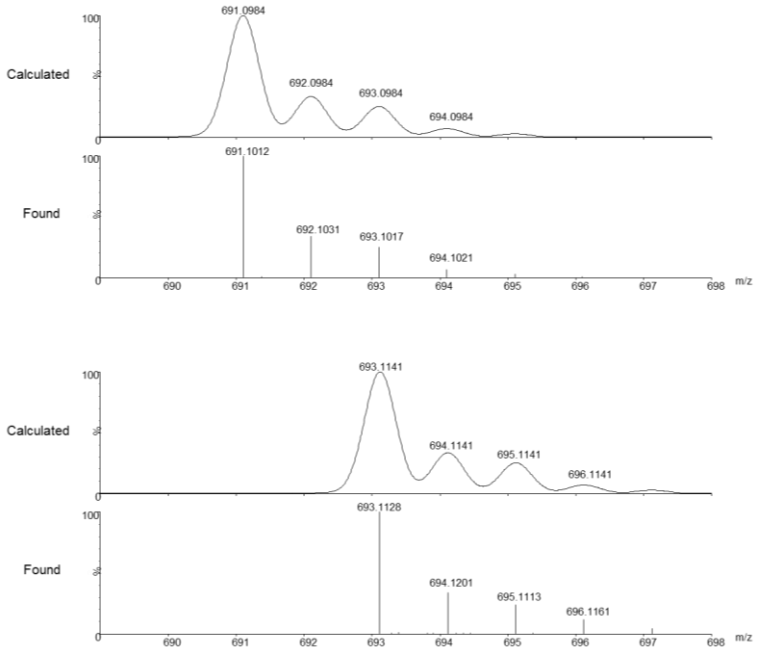


Mixture of **1a**+**2a**+**3e** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

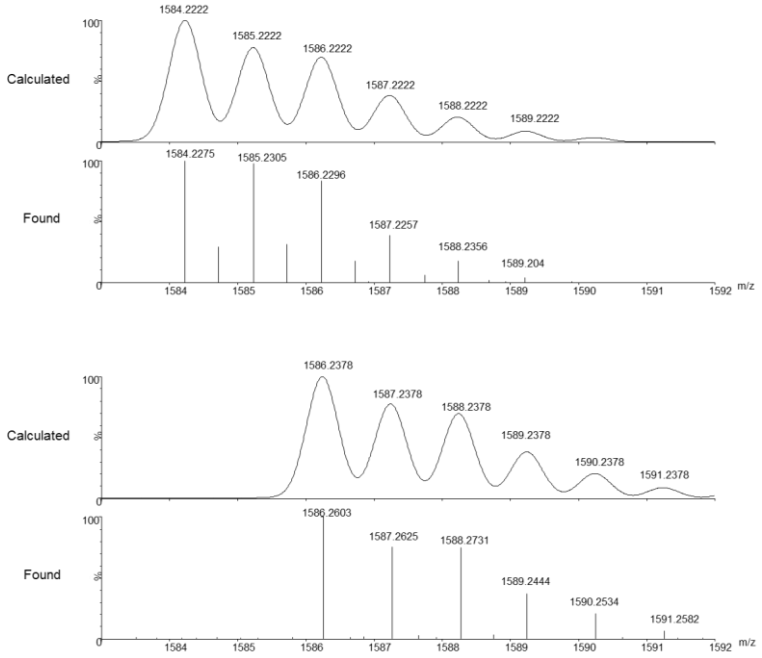


Identification of the products:

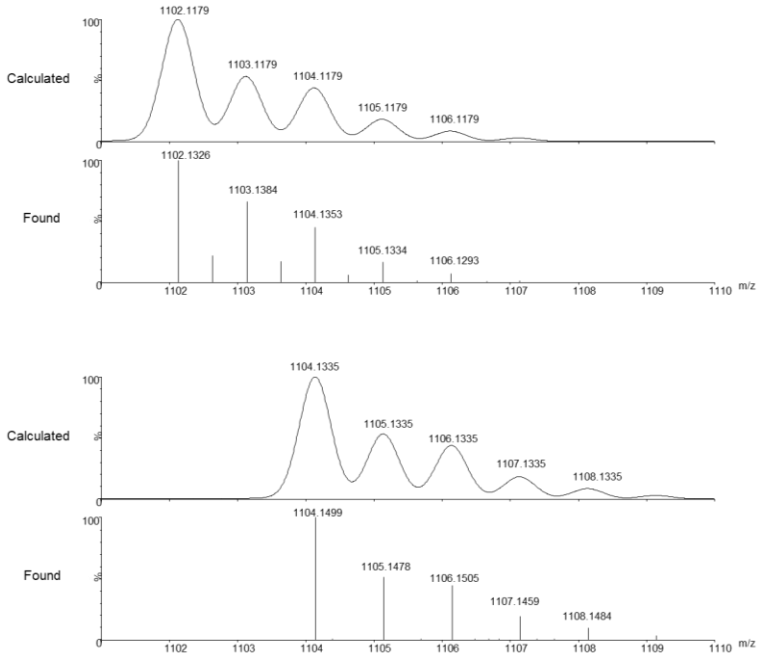
[**2a-3e$_2$**]
Retention time: 12.13 min.
Chemical Formula: $C_{24}H_{32}N_6O_{10}S_4$
Exact Mass: 692.1063



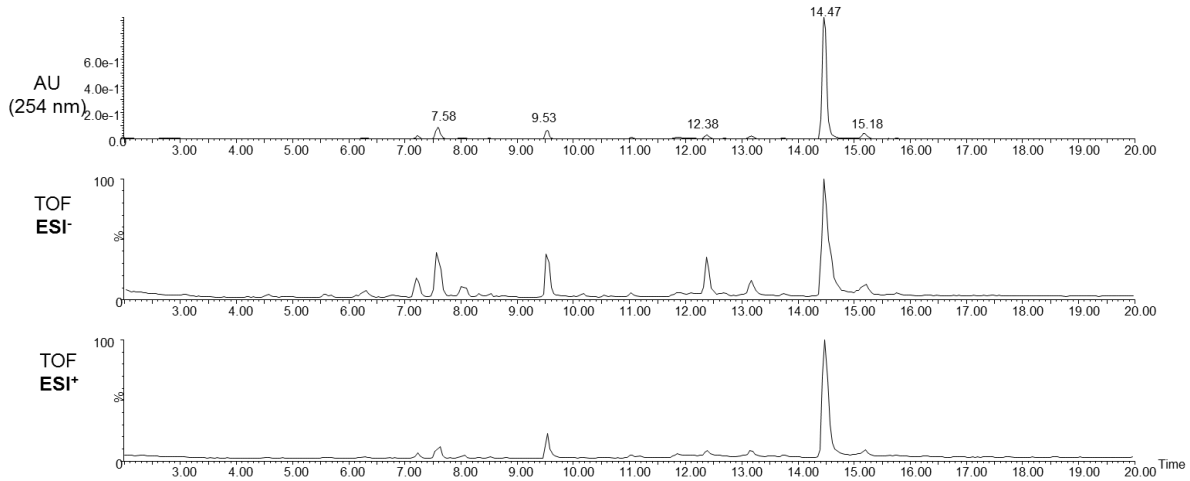
[**1a-2a$_2$-3e**]
Retention time: 16.62 min.
Chemical Formula: $C_{57}H_{67}N_{15}O_{23}S_8$
Exact Mass: 1585.2300



[**1a-2a-3e**]
Retention time: 15.03 min.
Chemical Formula: $C_{39}H_{45}N_9O_{17}S_6$
Exact Mass: 1103.1258

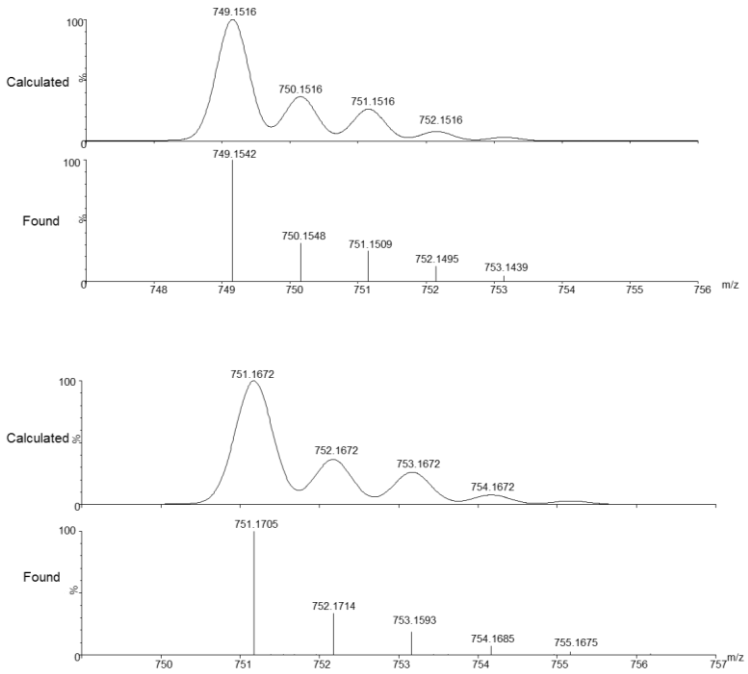


Mixture of **1a**+**2a**+**3f** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

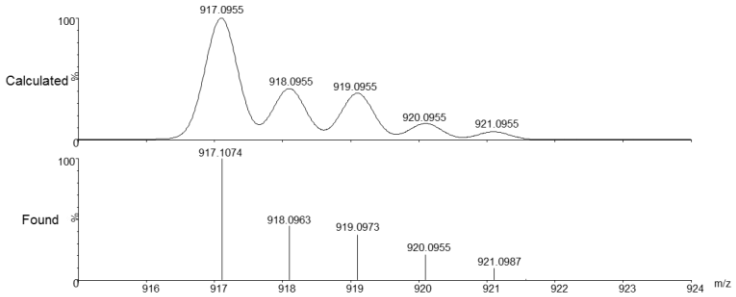


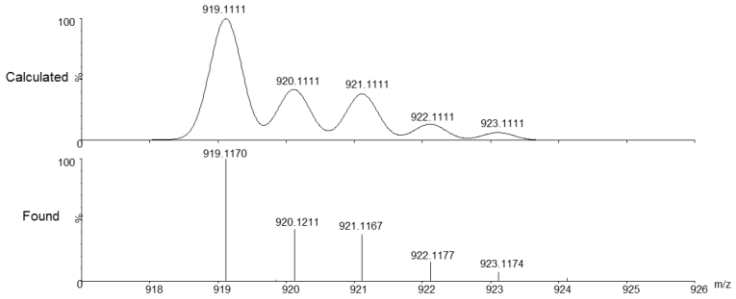
Identification of the products:

[**2a-3f$_2$**]
Retention time: 7.58 min.
Chemical Formula: $C_{26}H_{38}N_8O_{10}S_4$
Exact Mass: 750.1594



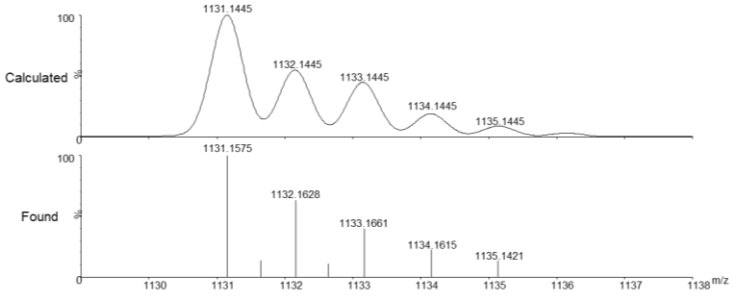
[**1a-3f$_3$**]
Retention time: 9.53 min.
Chemical Formula:$C_{30}H_{42}N_6O_{15}S_6$
Exact Mass: 918.1032

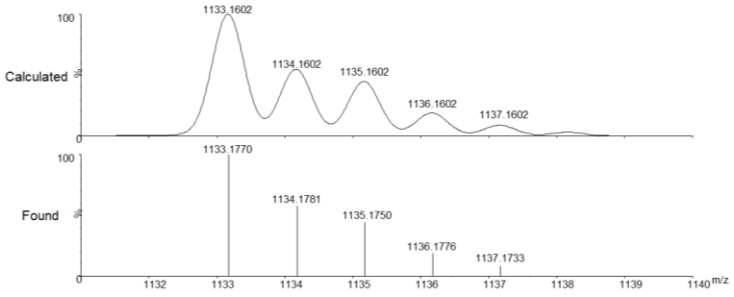




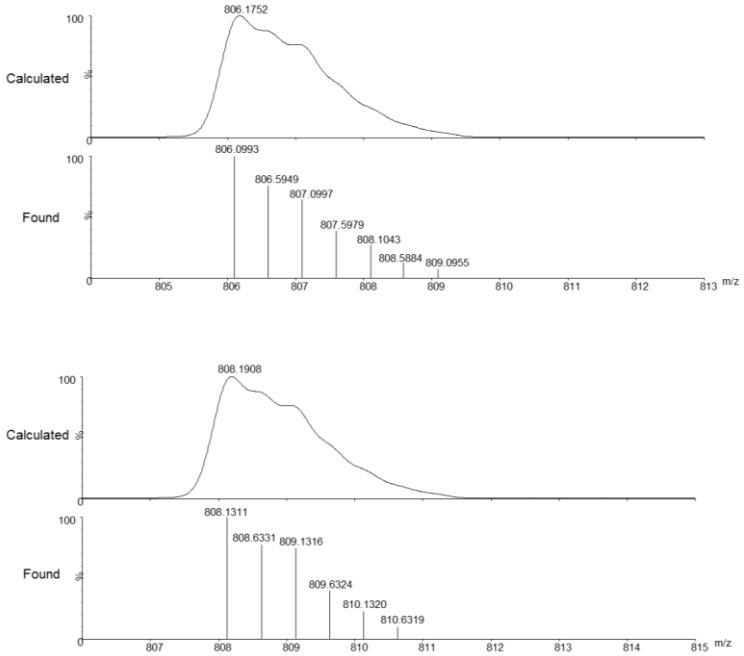
[**1a$_2$**] (previously identified)
Retention time: 12.38 min.
Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$
Exact Mass:1032.0410

[1a-2a-3f]
Retention time: 14.47 min.
Chemical Formula: $C_{40}H_{48}N_{10}O_{17}S_6$
Exact Mass: 1132.1523

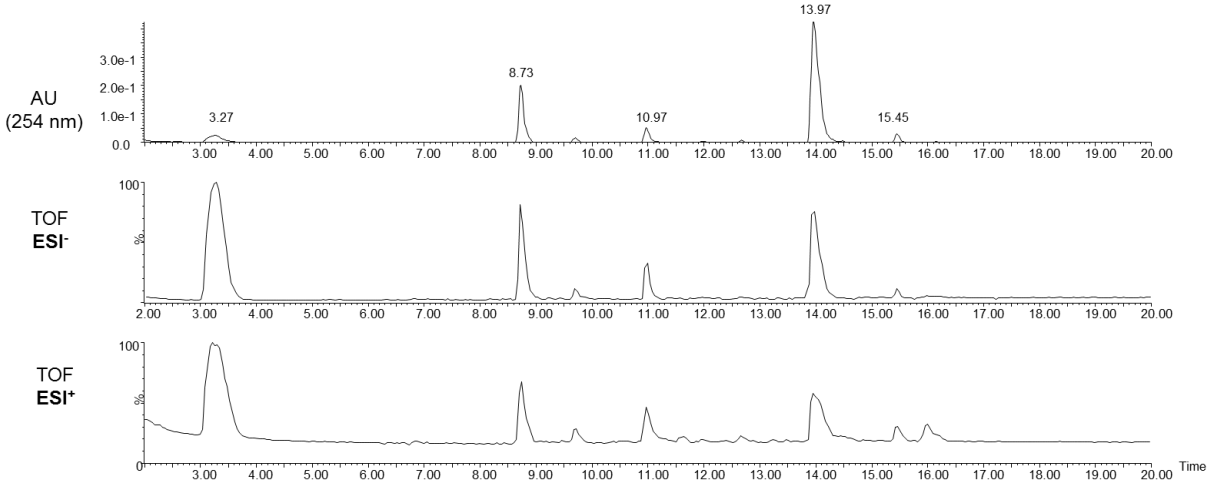




[**1a-2a$_2$-3f**]
Retention time: 15.18 min.
Chemical Formula: $C_{58}H_{70}N_{16}O_{23}S_8$
Exact Mass: 1614.2565

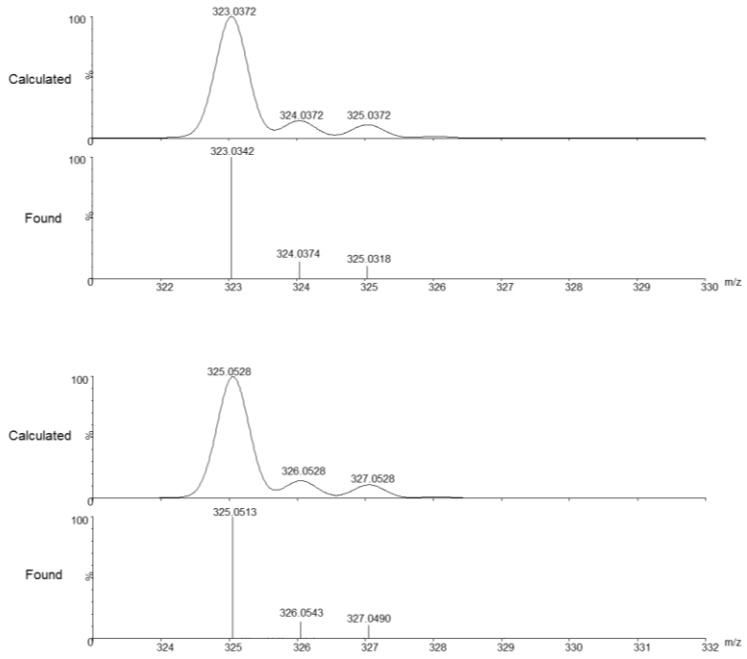


Mixture of **1a**+**2a**+**3g** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

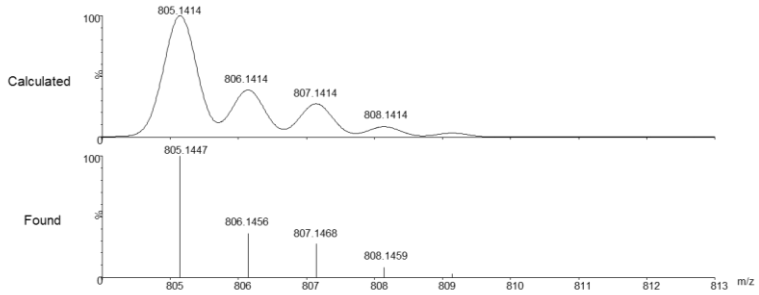


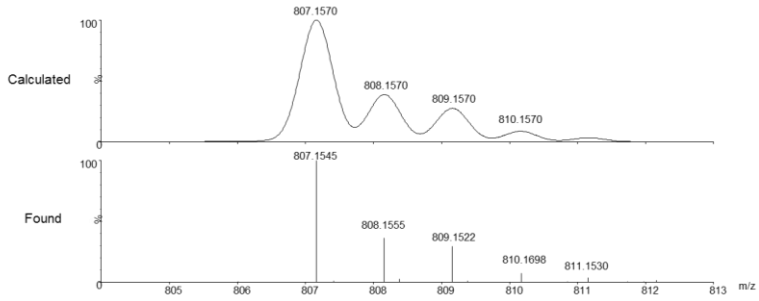
Identification of the products:

[**3g**$_2$]
Retention time: 3.27 min.
Chemical Formula:$C_{10}H_{16}N_2O_6S_2$
Exact Mass: 324.0450

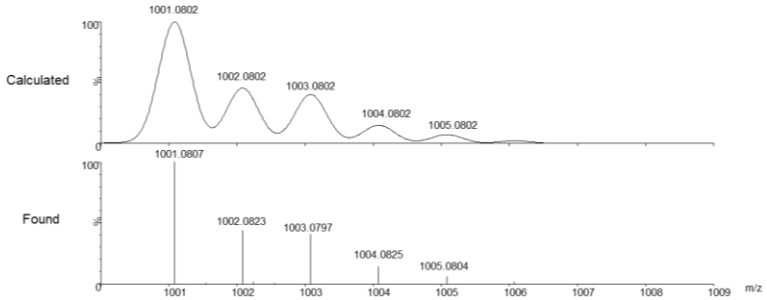


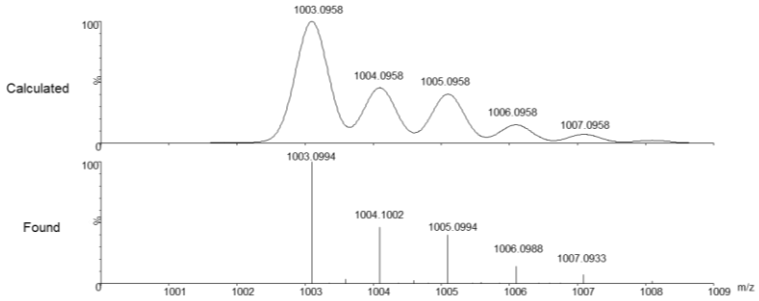
[**2a-3g$_2$**]
Retention time: 8.73 min.
Chemical Formula: $C_{28}H_{38}N_8O_{12}S_4$
Exact Mass: 806.1492



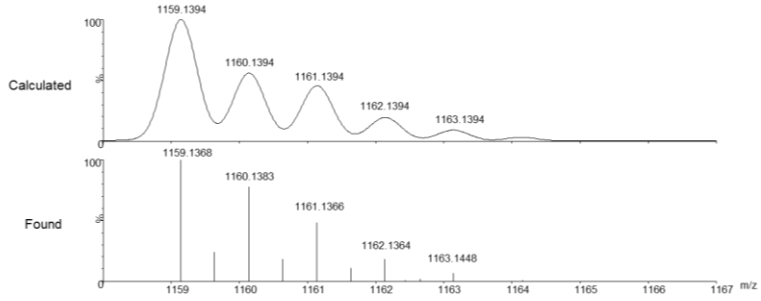


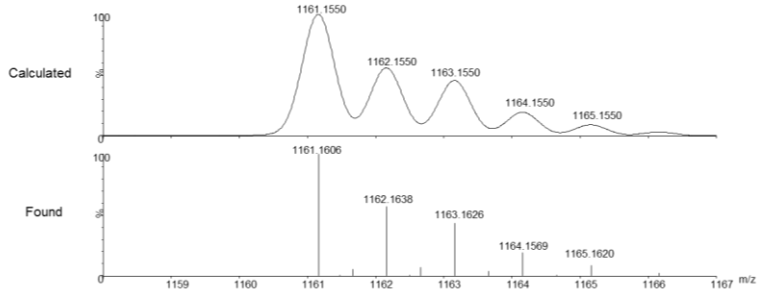
[**1a-3g$_3$**]
Retention time: 10.97 min.
Chemical Formula: $C_{33}H_{42}N_6O_{18}S_6$
Exact Mass: 1002.0880



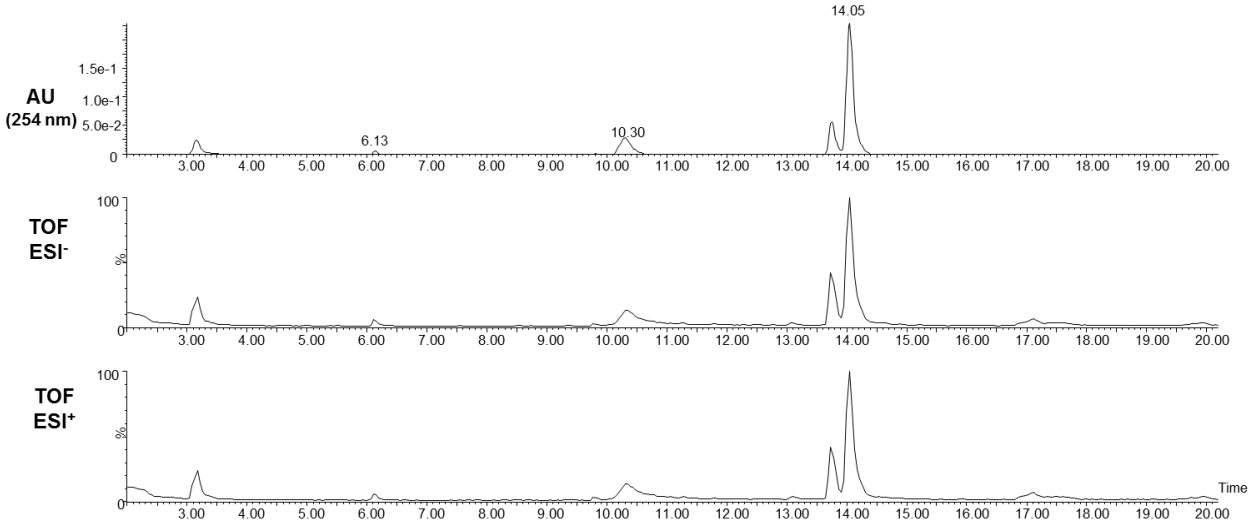


[**1a-2a-3g**]
Retention time: 13.97 min.
Chemical Formula: $C_{41}H_{48}N_{10}O_{18}S_6$
Exact Mass: 1160.1472





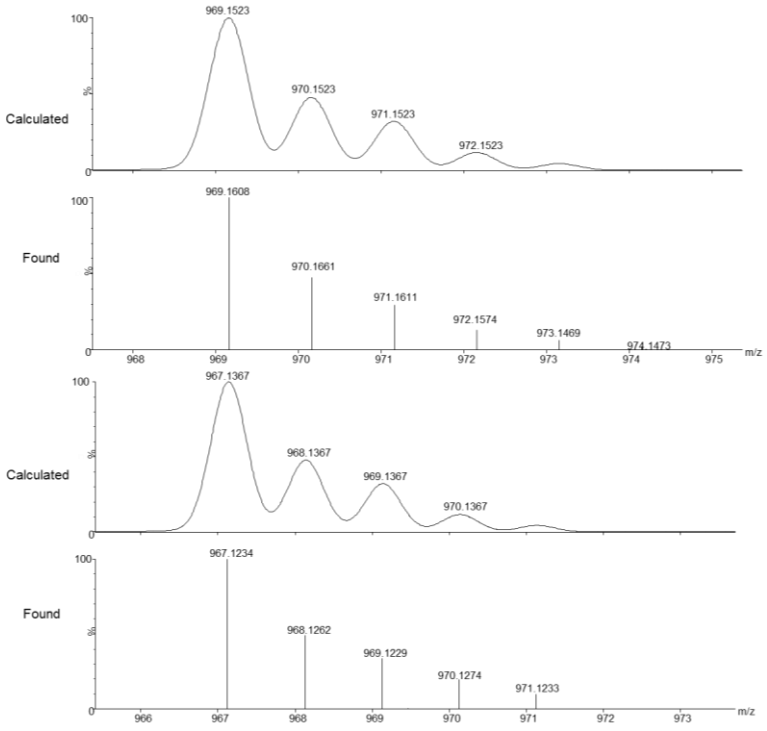
Mixture of **1a**+**2b** (0.5 mM each, pH.6.5, *t* = 24 h)



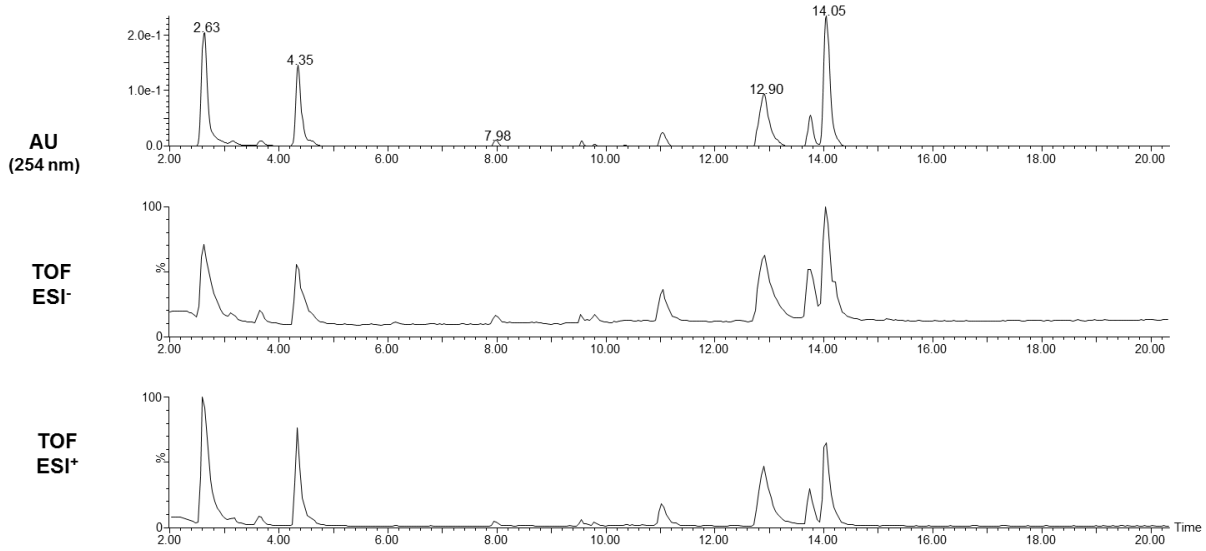
Identification of the products:

[**1a**$_2$] (previously identified)
Retention time: 10.30 min.
Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$
Exact Mass: 1032.0410

[**2b**$_2$]
Retention time: 6.13, 14.05 min.
Chemical Formula:
$C_{36}H_{40}N_8O_{16}S_4$
Exact Mass: 968.1445



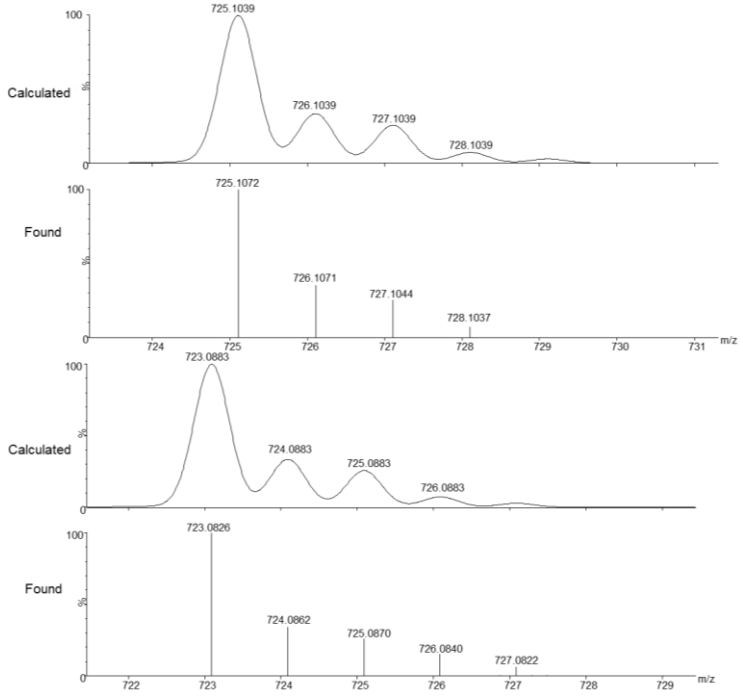
Mixture of **1a**+**2b**+**3a** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)



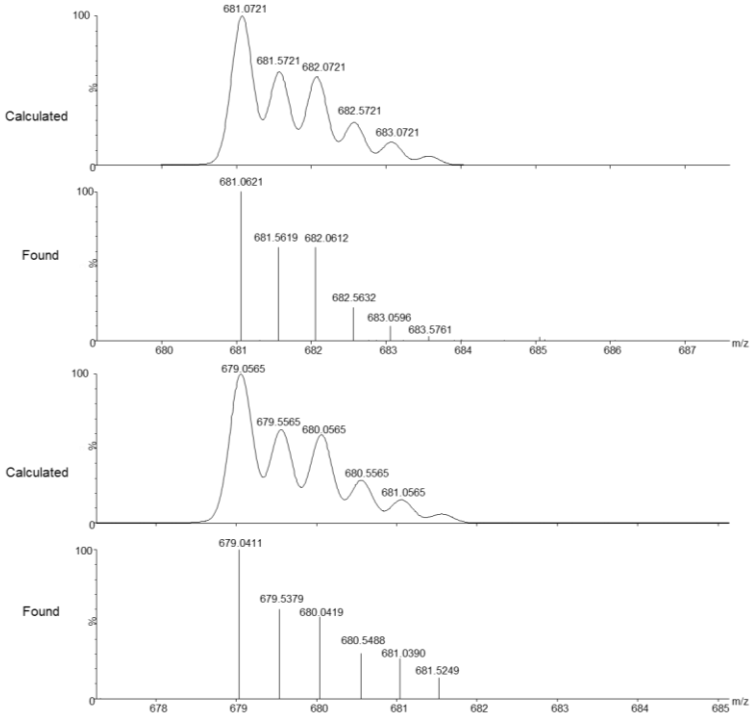
Identification of the products:

[**1a-3a**$_3$] (previously identified)
Retention time: 2.63 min.
Chemical Formula: $C_{27}H_{36}N_6O_{15}S_6$
Exact Mass: 876.0563

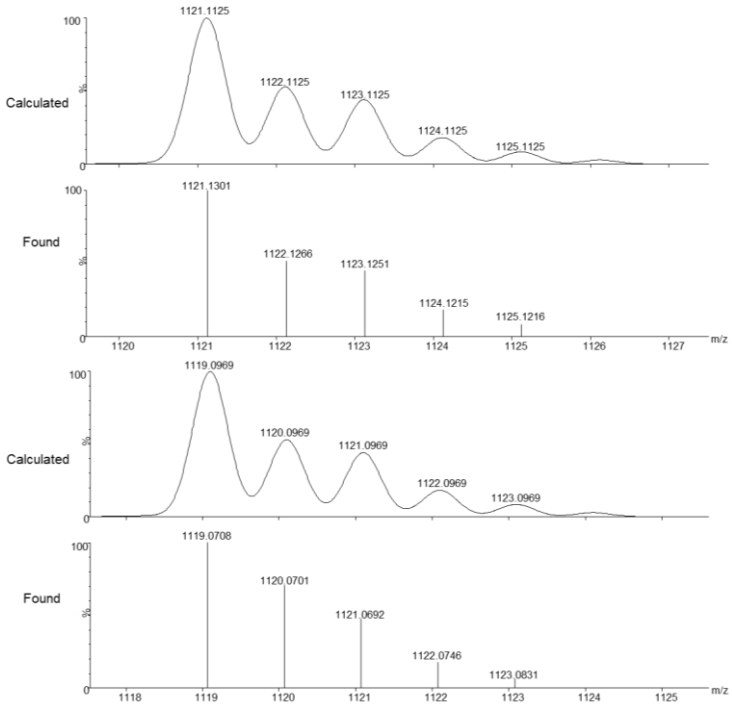
[**2b-3a**$_2$]
Retention time: 4.35 min.
Chemical Formula:$C_{24}H_{32}N_6O_{12}S_4$
Exact Mass: 724.0961



[1a-2b-3a$_3$]
Retention time: 7.98 min.
Chemical Formula:
$C_{45}H_{56}N_{10}O_{23}S_8$
Exact Mass: 1360.1285

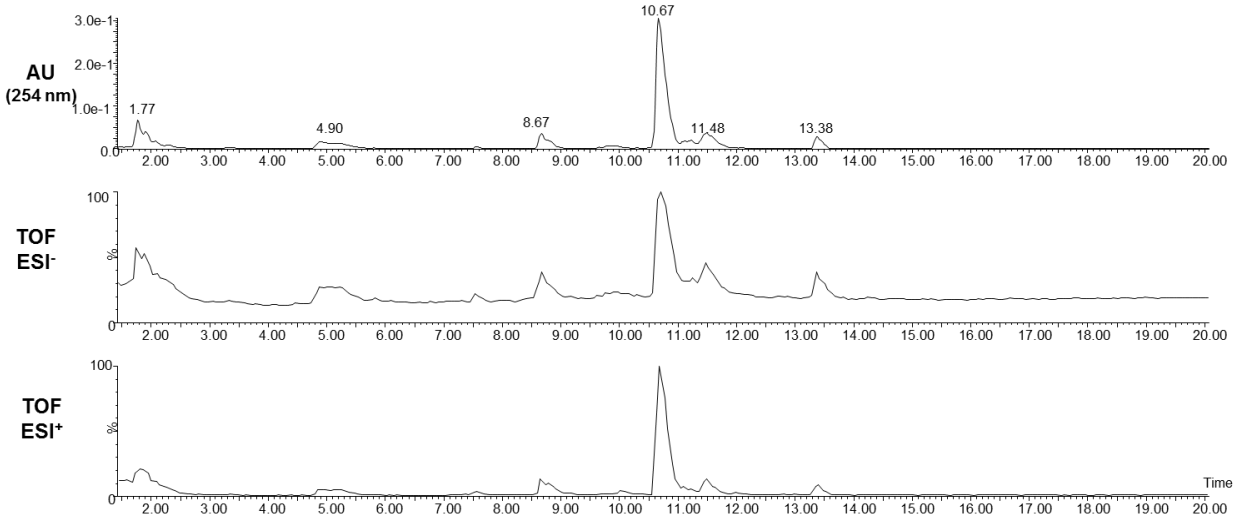


[1a-2b-3a]
Retention time: 12.90 min.
Chemical Formula: $C_{39}H_{44}N_8O_{19}S_6$
Exact Mass: 1120.1047



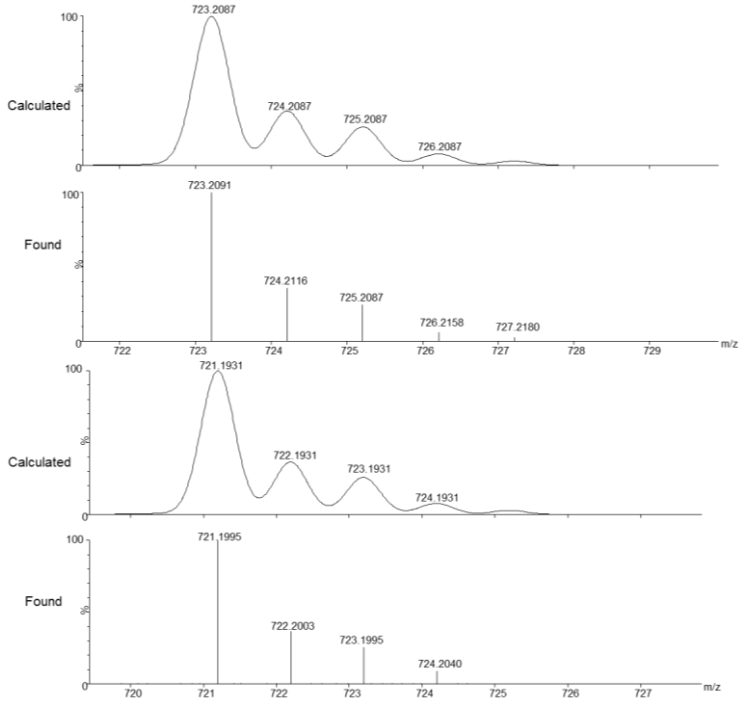
[**2b$_2$**] (previously identified)
Retention time: 14.05 min.
Chemical Formula: $C_{36}H_{40}N_8O_{16}S_4$
Exact Mass: 968.1445

Mixture of **1a**+**2c**+**3a** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

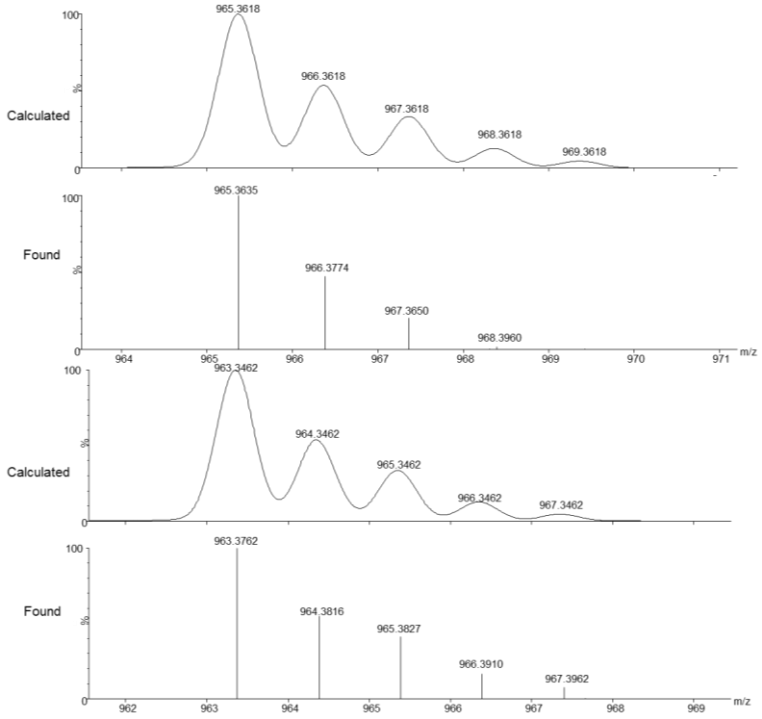


Identification of the products:

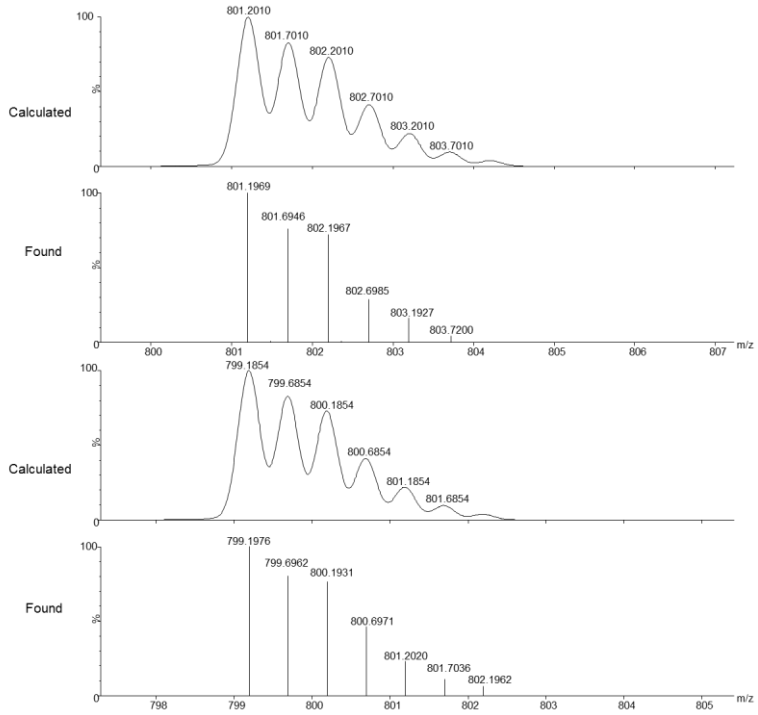
[**2c-3a$_2$**]
Retention time: 1.77 min.
Chemical Formula: $C_{26}H_{42}N_8O_8S_4$
Exact Mass: 722.2008



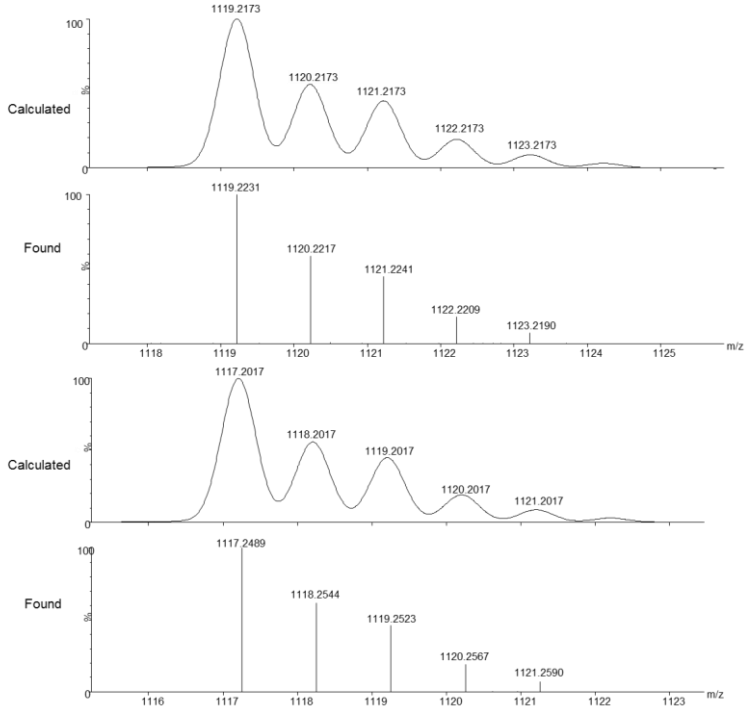
[**2c$_2$**]
Retention time: 4.90 min.
Chemical Formula: $C_{40}H_{60}N_{12}O_8S_4$
Exact Mass: 964.3540



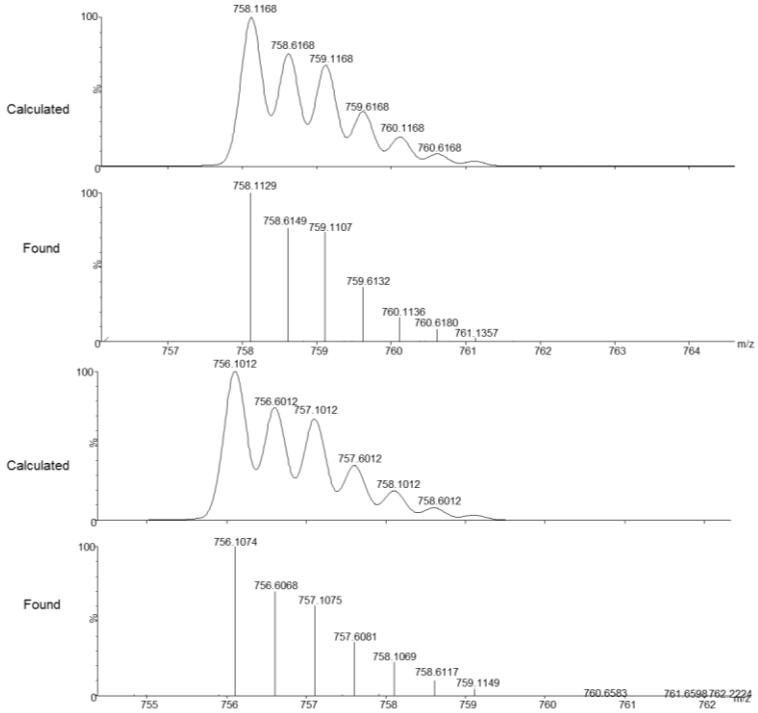
[**1a-2c$_2$-3a**]
Retention time: 8.67 min.
Chemical Formula: $C_{61}H_{84}N_{16}O_{19}S_8$
Exact Mass: 1600.3864



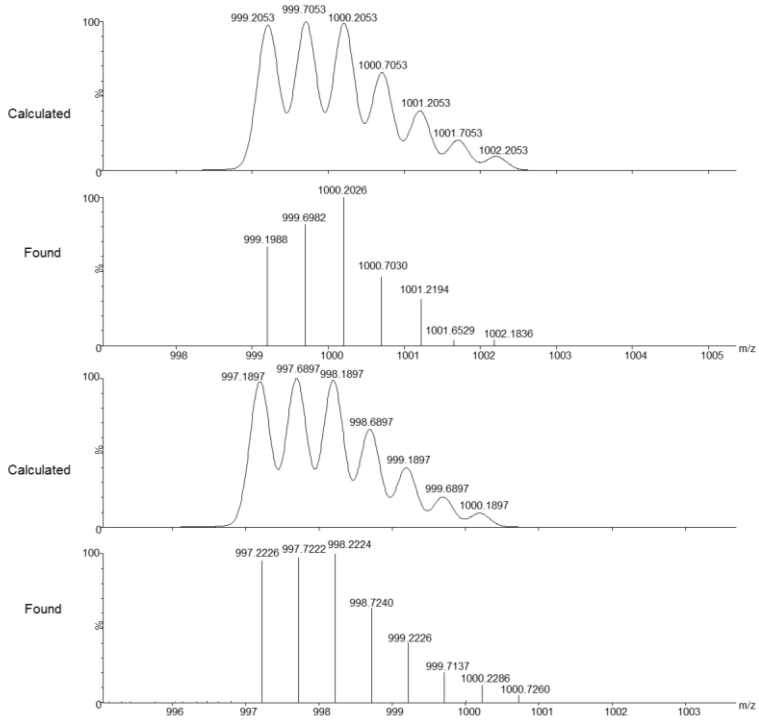
[**1a-2c-3a**]
Retention time: 10.67 min.
Chemical Formula: $C_{41}H_{54}N_{10}O_{15}S_6$
Exact Mass: 1118.2094



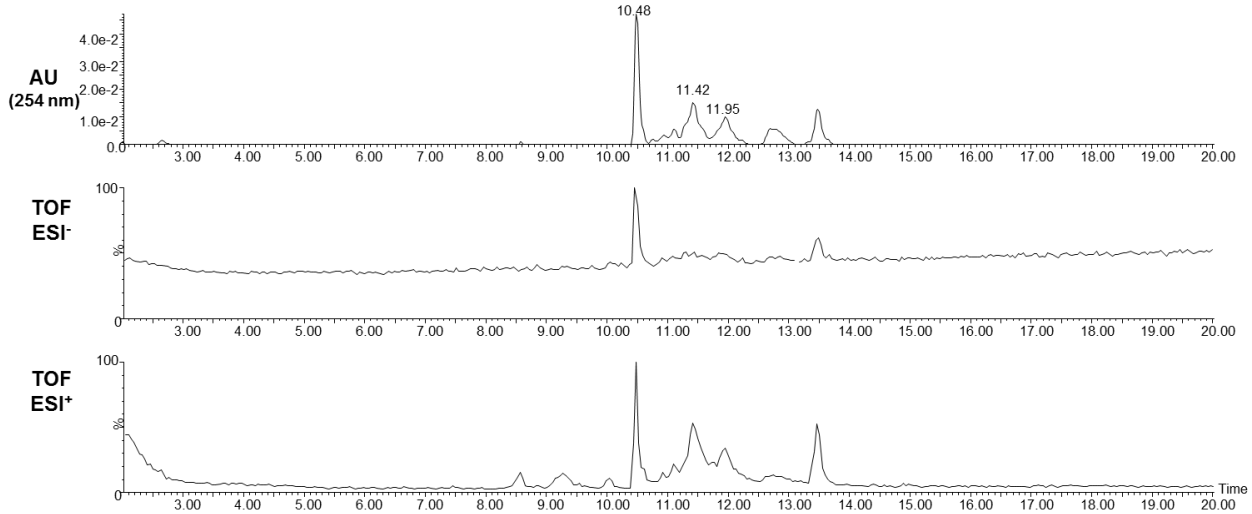
[**1a$_2$-2c**]
Retention time: 11.48 min.
Chemical Formula: $C_{56}H_{66}N_{12}O_{22}S_8$
Exact Mass: 1514.2180



[**1a**$_2$-**2c**$_2$]
Retention time: 13.38 min.
Chemical Formula: $C_{76}H_{96}N_{18}O_{26}S_{10}$
Exact Mass: 1996.3950



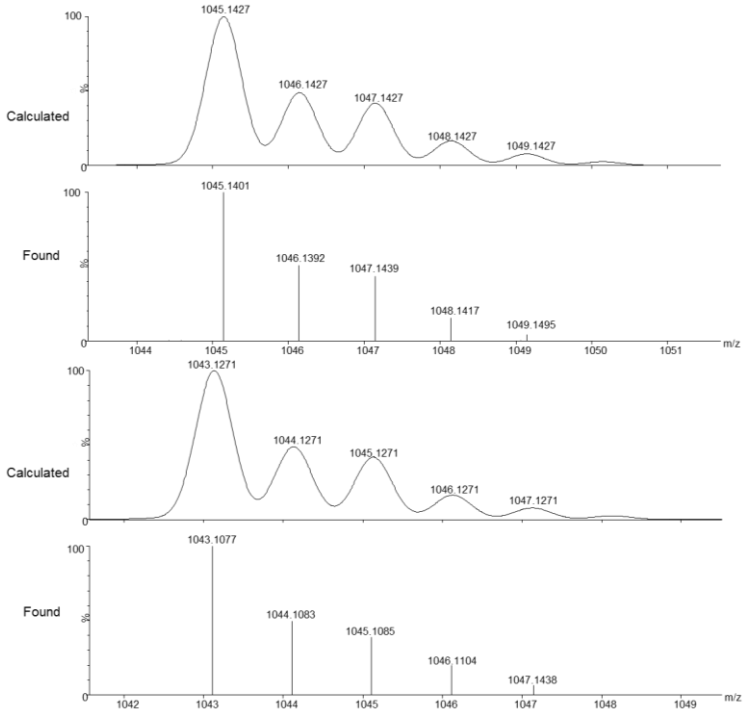
Mixture of **1b**+**2a** (0.5 each, pH.6.5, *t* = 24 h)



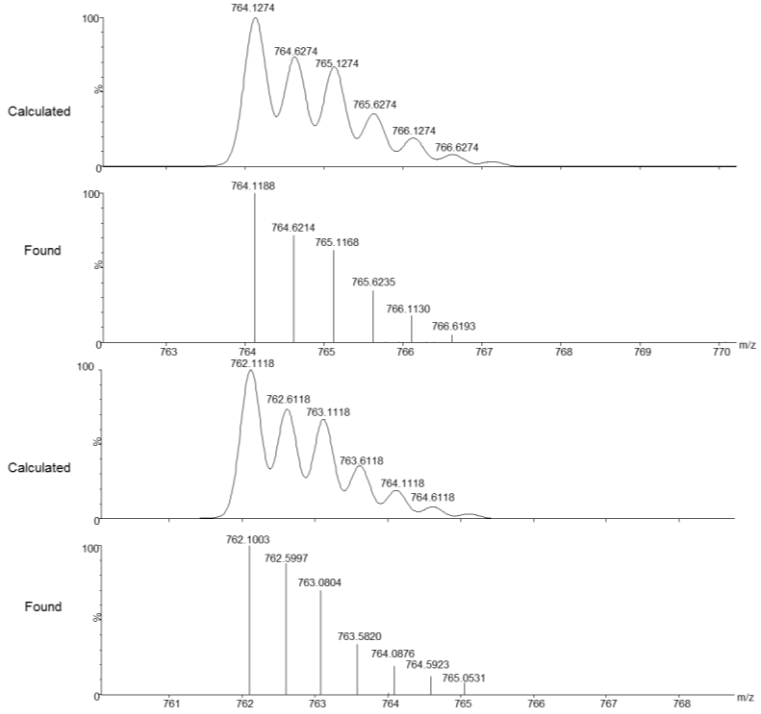
Identification of the products:

[**2a**$_2$] (previously identified)
Retention time: 10.48 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

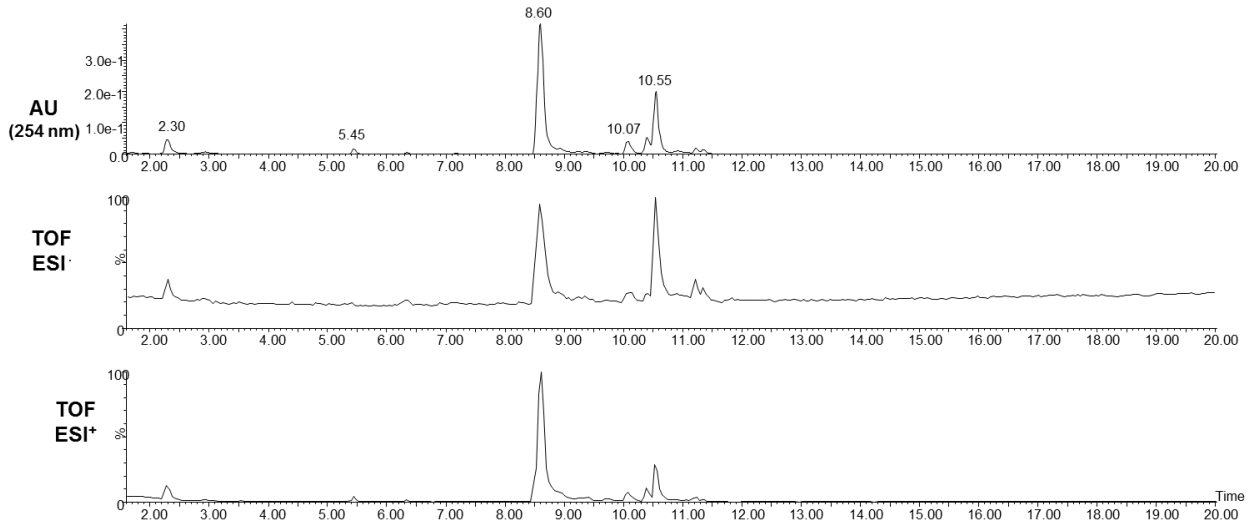
[**1b**$_2$]
Retention time: 11.42 min.
Chemical Formula: $C_{36}H_{48}N_6O_{18}S_6$
Exact Mass:1044.1349



[**1b$_2$-2a**]
Retention time: 11.95 min.
Chemical Formula: $C_{54}H_{70}N_{12}O_{24}S_8$
Exact Mass: 1526.2392

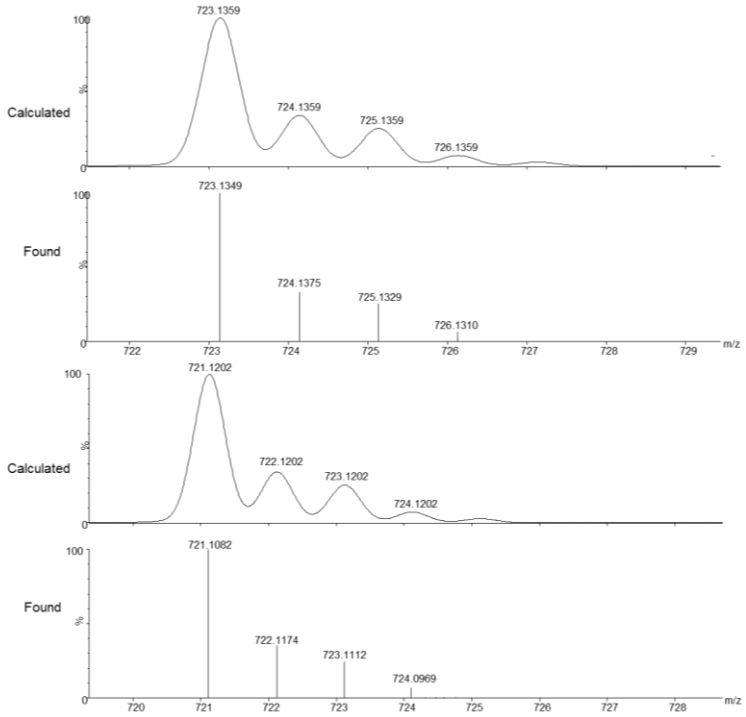


Mixture of **1b**+**2a**+**3a** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

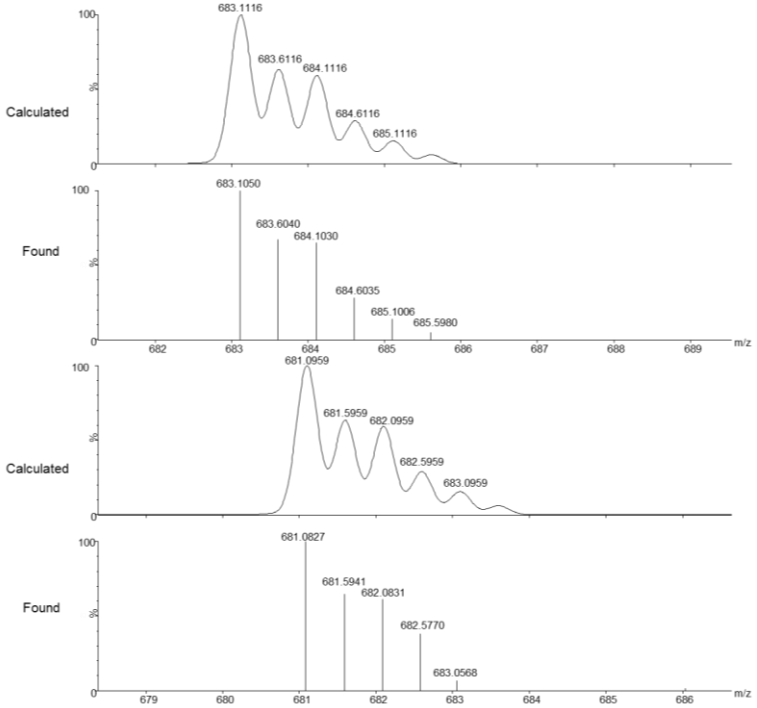


Identification of the products:

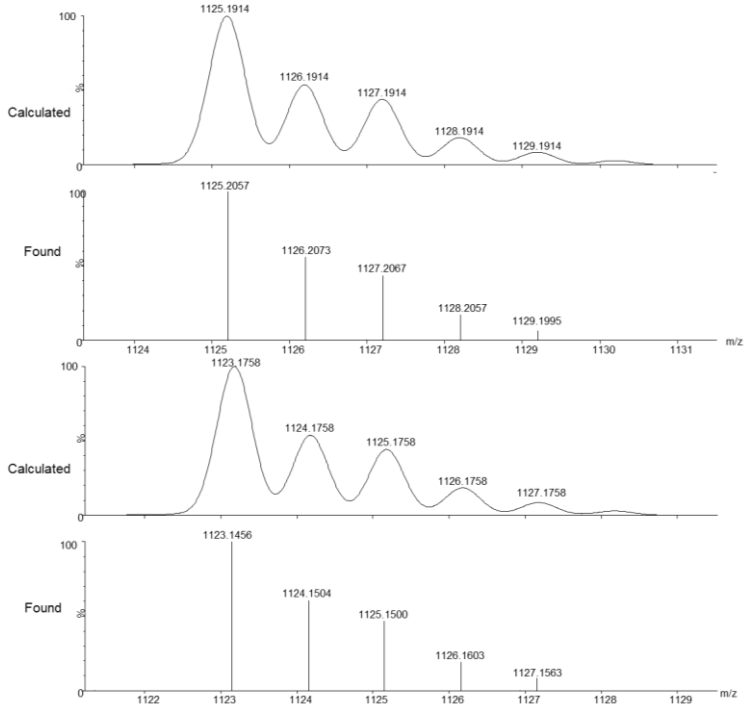
[**2a-3a$_2$**]
Retention time: 2.30 min.
Chemical Formula: $C_{24}H_{34}N_8O_{10}S_4$
Exact Mass: 722.1281



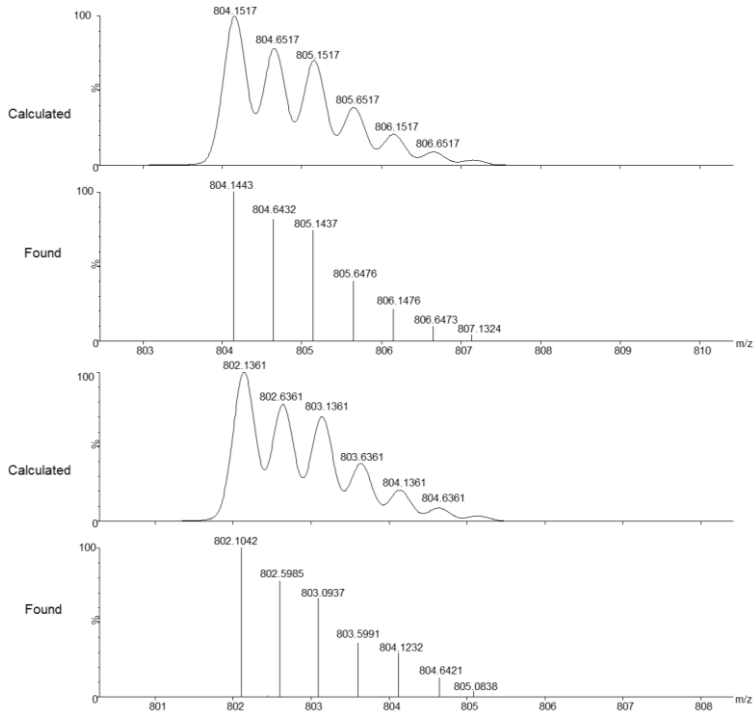
[1b-2a-3a$_3$]
Retention time: 5.45 min.
Chemical Formula: $C_{45}H_{64}N_{12}O_{21}S_8$
Exact Mass: 1364.2075



[1b-2a-3a]
Retention time: 8.60 min.
Chemical Formula: $C_{39}H_{52}N_{10}O_{17}S_6$
Exact Mass: 1124.1836

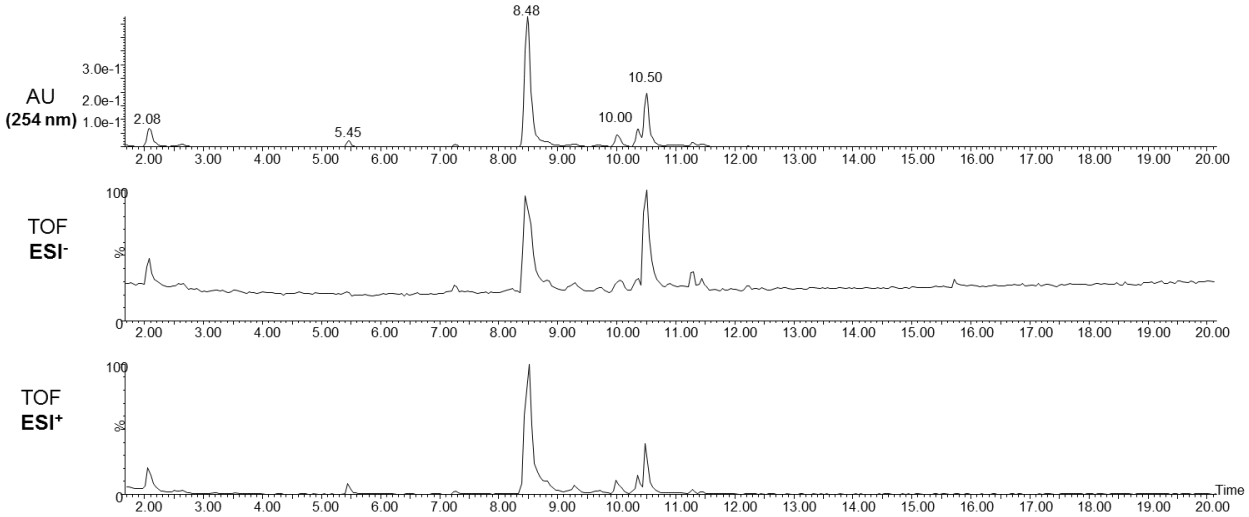


[1b-2a$_2$-3a]
Retention time: 10.40 min.
Chemical Formula: $C_{57}H_{74}N_{16}O_{23}S_8$
Exact Mass: 1606.2878



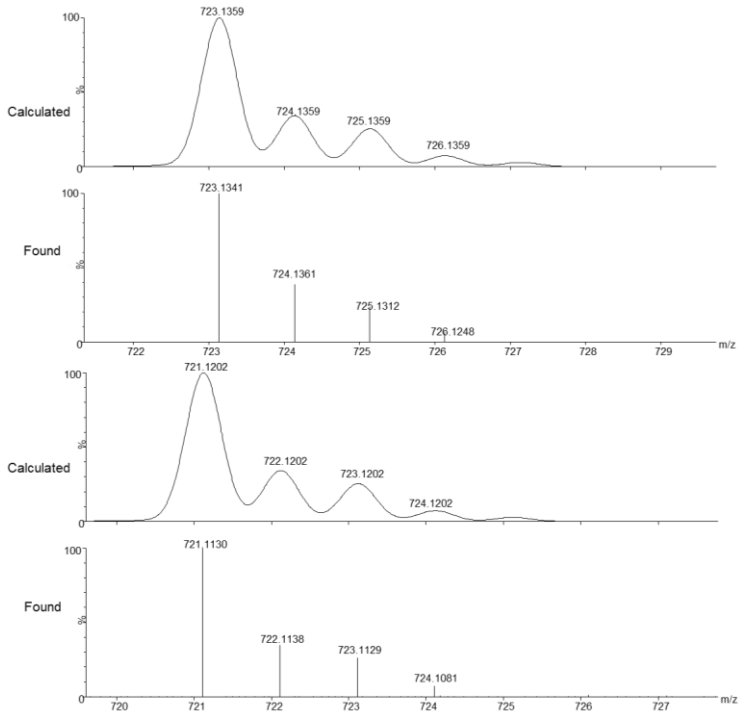
[**2a$_2$**] (previously identified)
Retention time: 10.55 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

Mixture of **1b**+**2a**+**3b** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

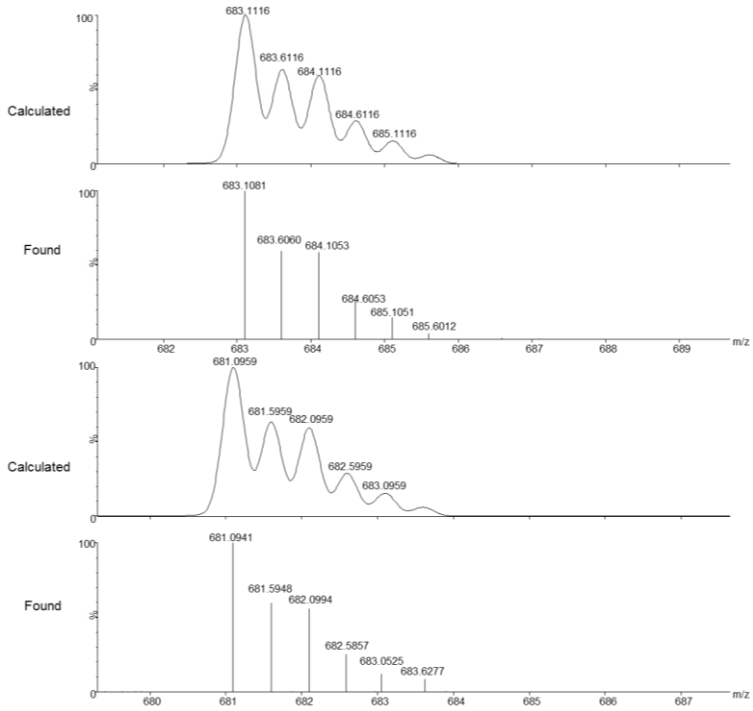


Identification of the products:

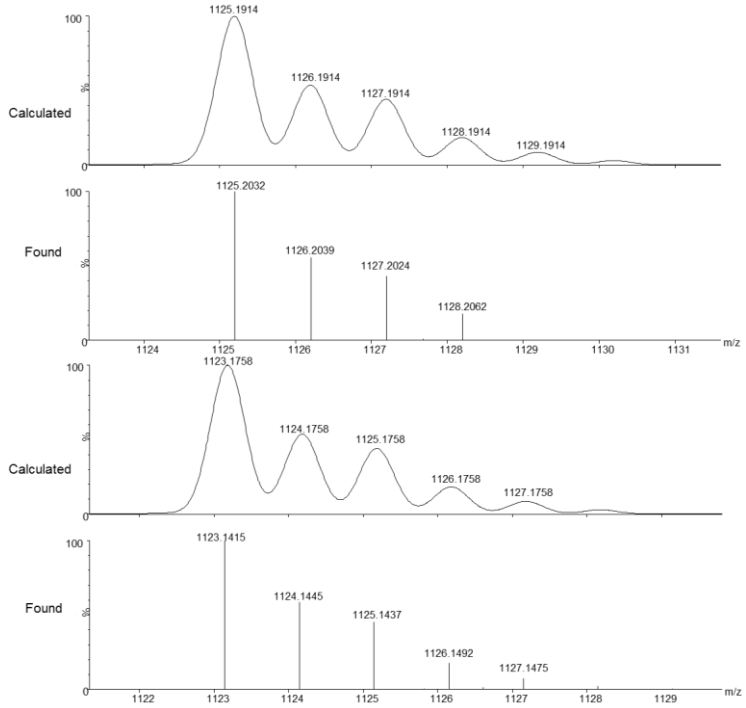
[**2a-3b$_2$**]
Retention time: 2.08 min.
Chemical Formula: $C_{24}H_{34}N_8O_{10}S_4$
Exact Mass: 722.1281



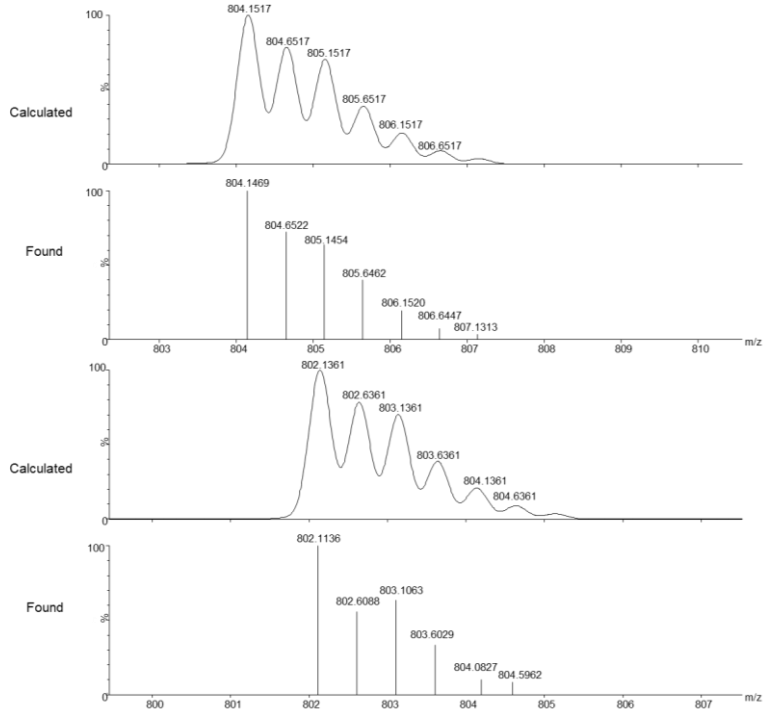
[**1b-2a$_2$-3b**]
Retention time: 5.45 min.
Chemical Formula: $C_{45}H_{64}N_{12}O_{21}S_8$
Exact Mass: 1364.2075



[**1b-2a-3b**]
Retention time: 8.48 min.
Chemical Formula: $C_{39}H_{52}N_{10}O_{17}S_6$
Exact Mass: 1124.1836

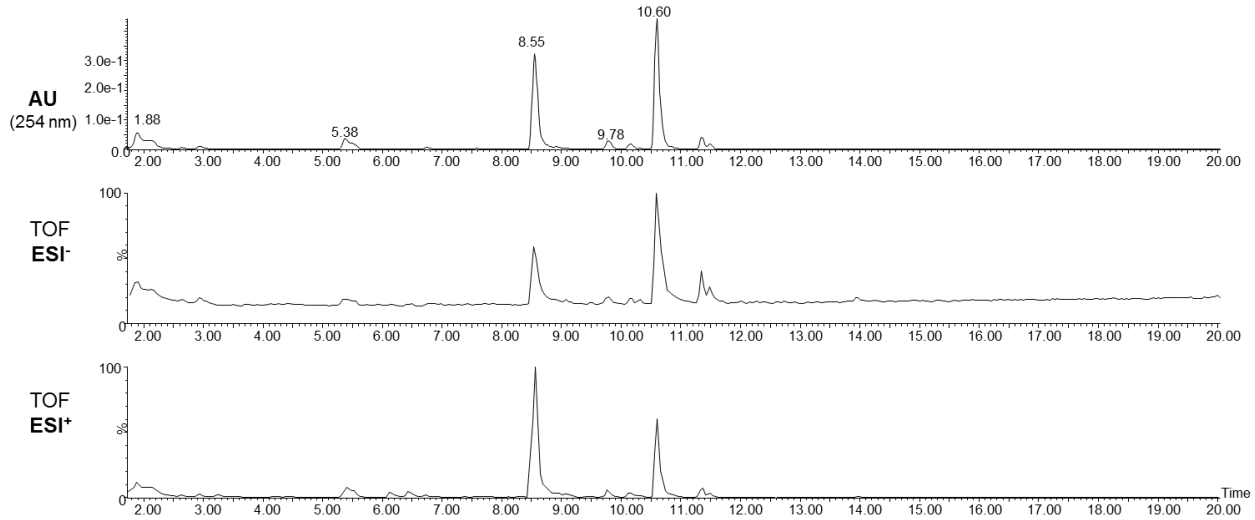


[**1b-2a$_2$-3b**]
Retention time: 10.00 min.
Chemical Formula: $C_{57}H_{74}N_{16}O_{23}S_8$
Exact Mass: 1606.2878



[**2a$_2$**] (previously identified)
Retention time: 10.50 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass:964.2084

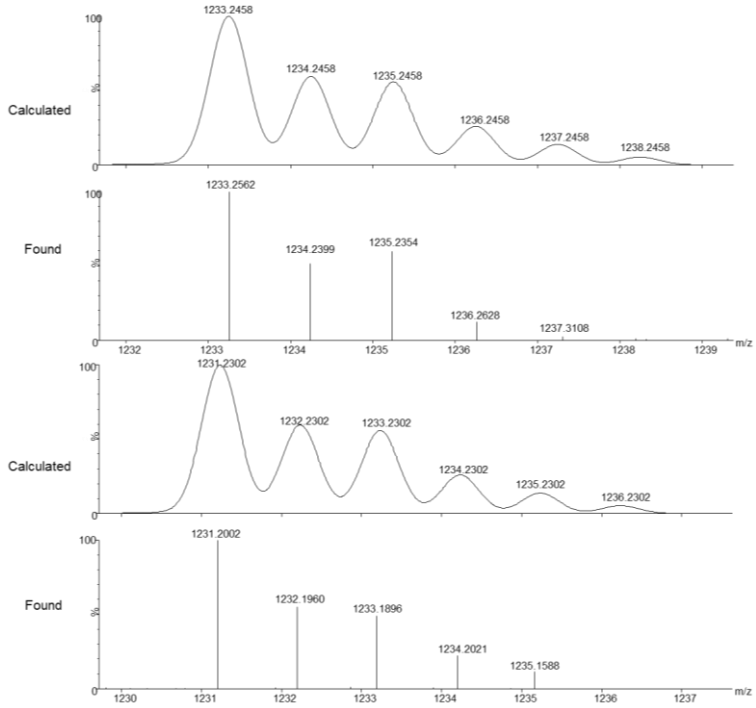
Mixture of **1b**+**2a**+**3c** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)



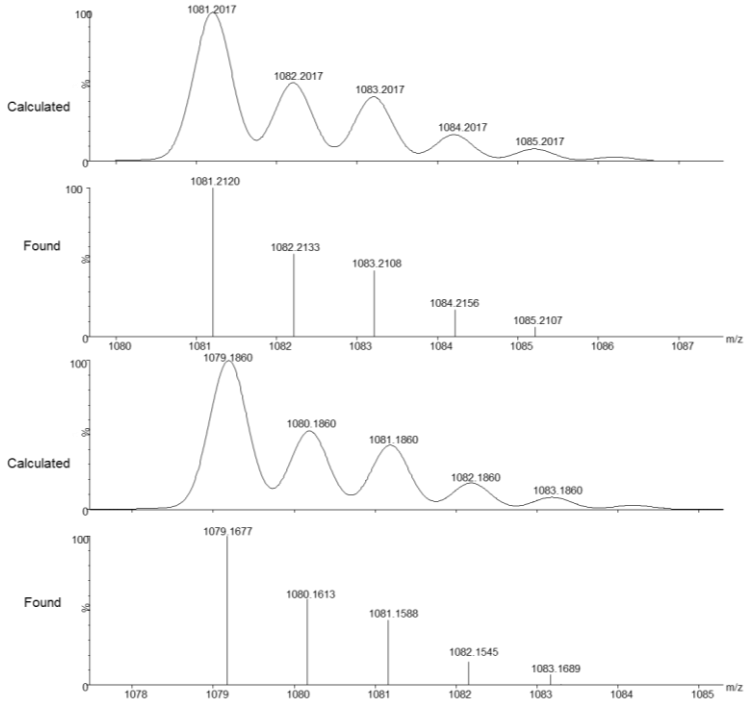
Identification of the products:

[**2a-3c$_2$**] (previously identified)
Retention time: 1.88min.
Chemical Formula: $C_{22}H_{34}N_8O_6S_4$
Exact Mass: 634.1484

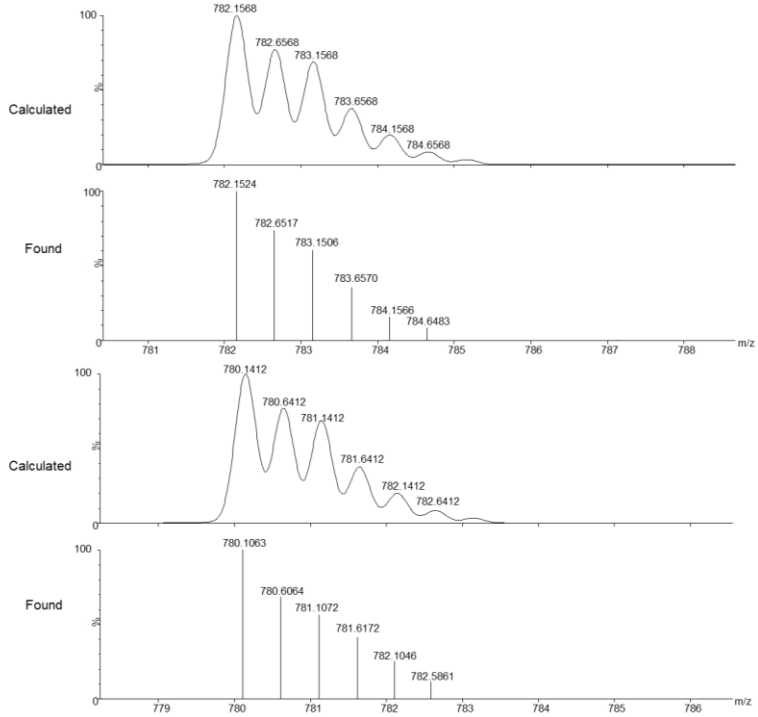
[**1b-2a-3c$_3$**]
Retention time: 5.38min.
Chemical Formula: $C_{42}H_{64}N_{12}O_{15}S_8$
Exact Mass: 1232.2380



[1b-2a-3c]
Retention time: 8.55min.
Chemical Formula: $C_{38}H_{52}N_{10}O_{15}S_6$
Exact Mass: 1080.1938

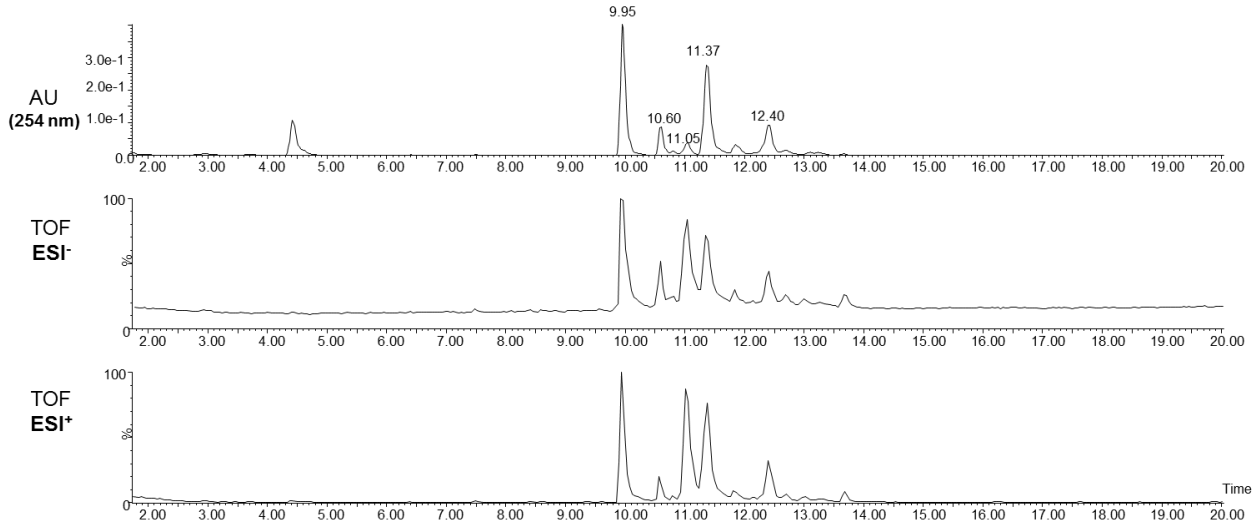


[**1b**-**2a$_2$**-**3c**]
Retention time: 9.78 min.
Chemical Formula: $C_{56}H_{74}N_{16}O_{21}S_8$
Exact Mass: 1562.2980



[**2a$_2$**] (previously identified)
Retention time: 10.60 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

Mixture of **1b**+**2a**+**3d** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

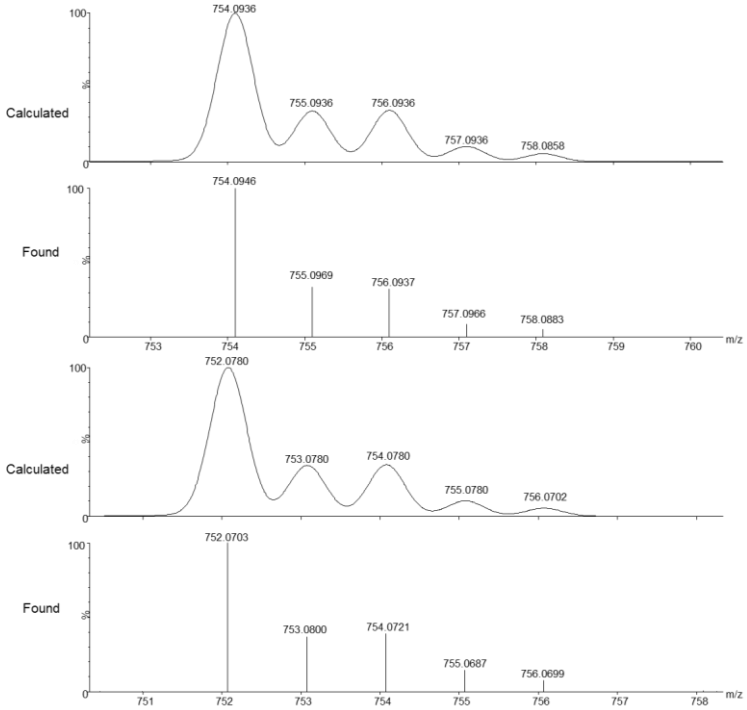


Identification of the products:

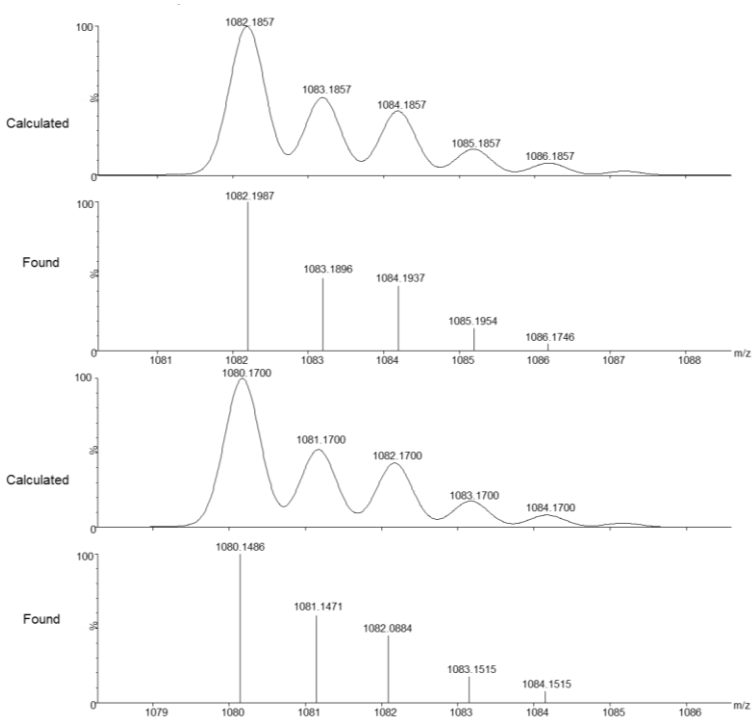
[**2a-3d$_2$**] (previously identified)
Retention time: 9.95 min.
Chemical Formula: $C_{22}H_{32}N_6O_8S_4$
Exact Mass: 636.1164

[**2a$_2$**] (previously identified)
Retention time: 10.60 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

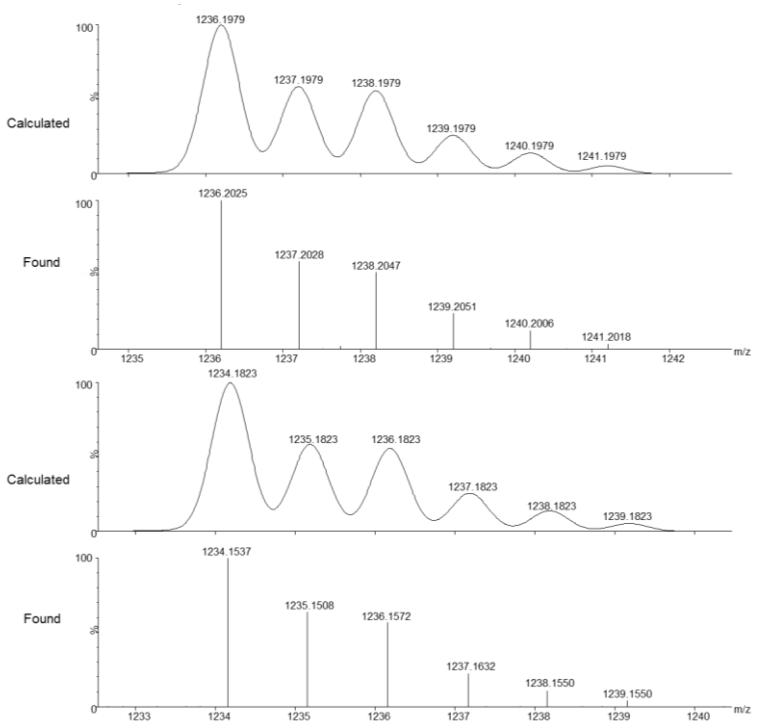
[**1b-3d$_3$**]
Retention time: 11.05 min.
Chemical Formula:
$C_{24}H_{39}N_3O_{12}S_6$
Exact Mass: 753.0858



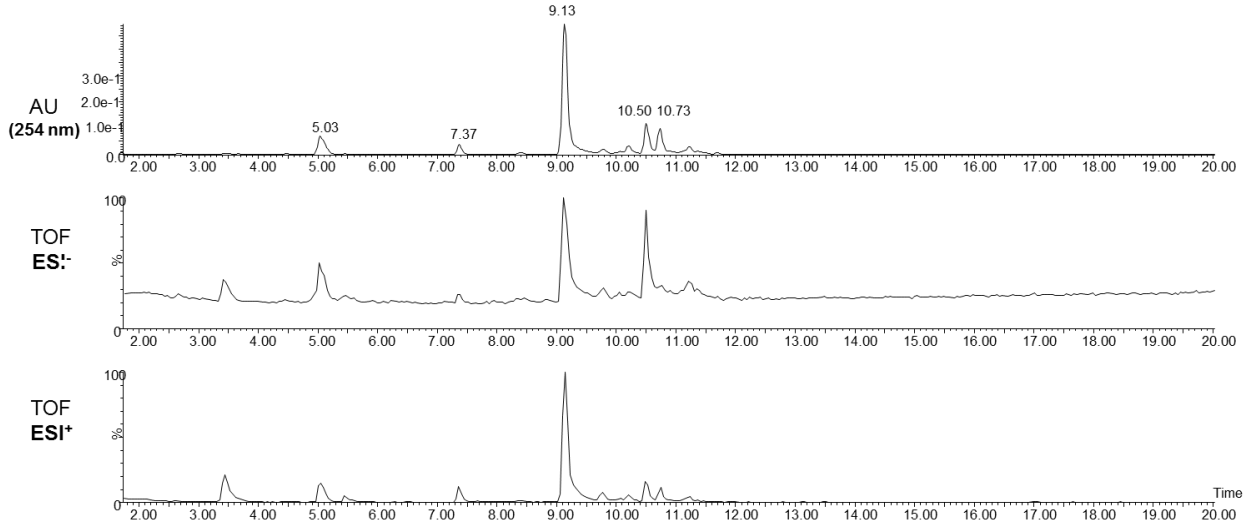
[**1b-2a-3d**]
Retention time: 11.37 min.
Chemical Formula: $C_{38}H_{51}N_9O_{16}S_6$
Exact Mass: 1081.1778



[1b-2a-3d$_3$]
Retention time: 12.40 min.
Chemical Formula: $C_{42}H_{61}N_9O_{18}S_8$
Exact Mass: 1235.1900



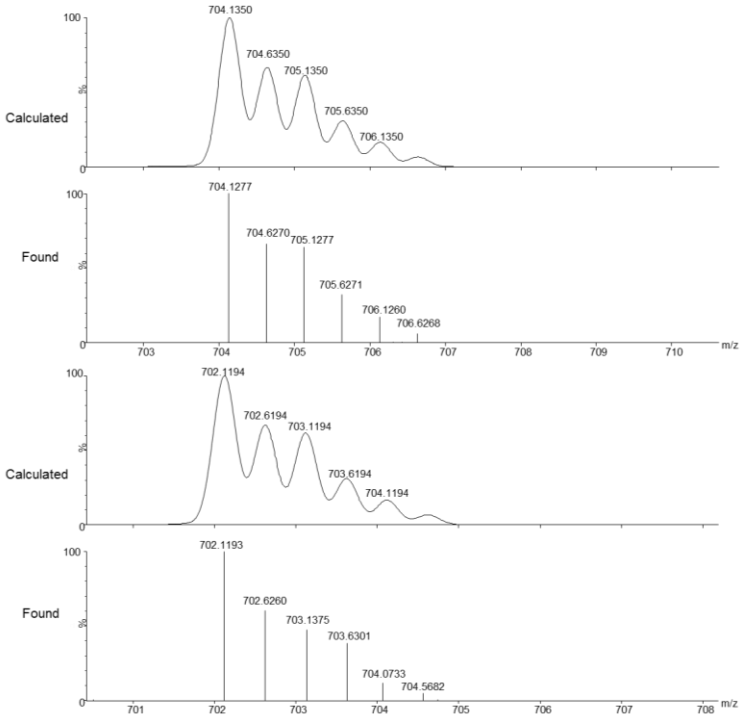
Mixture of **1b**+**2a**+**3f** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)



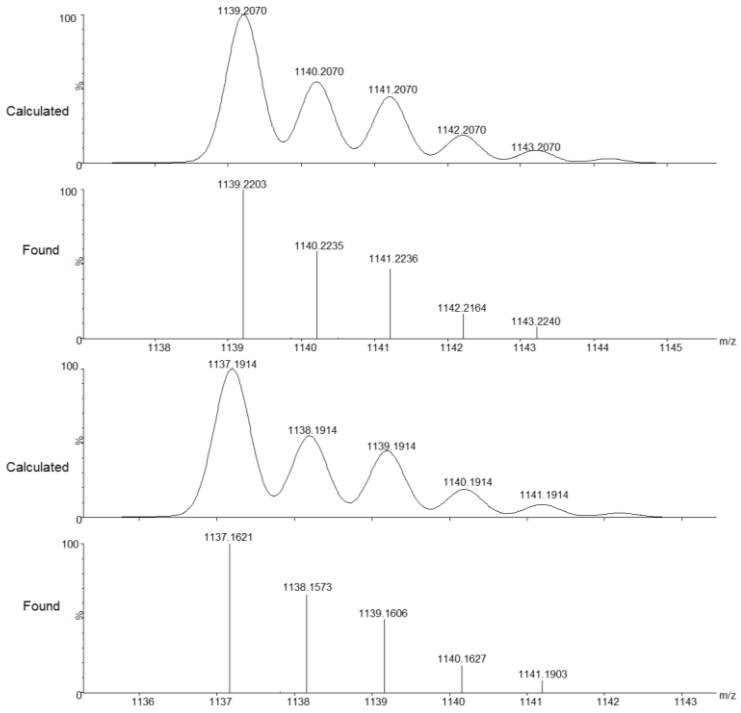
Identification of the products:

[2a-3f$_2$] (previously identified)
Retention time: 5.03 min.
Chemical Formula: $C_{26}H_{38}N_8O_{10}S_4$
Exact Mass: 750.1594

[1b-2a-3f$_3$]
Retention time: 7.37 min.
Chemical Formula: $C_{48}H_{70}N_{12}O_{21}S_8$
Exact Mass: 1406.2544

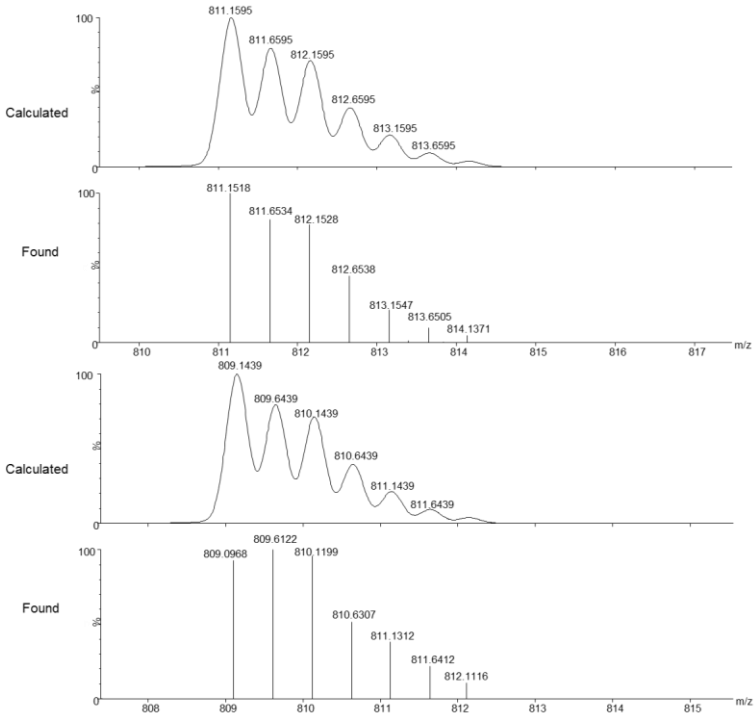


[1b-2a-3f]
Retention time: 9.13 min.
Chemical Formula: $C_{40}H_{54}N_{10}O_{17}S_6$
Exact Mass: 1138.1993

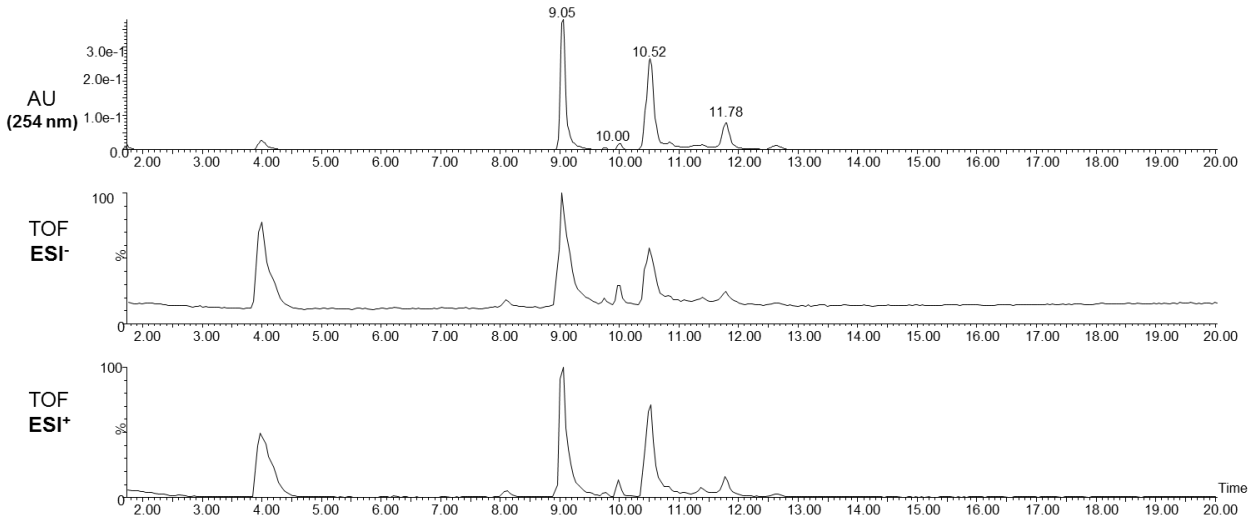


[$\mathbf{2a_2}$] (previously identified)
Retention time: 10.50 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[**1b-2a$_2$-3f**]
Retention time: 10.73 min.
Chemical Formula: $C_{58}H_{76}N_{16}O_{23}S_8$
Exact Mass: 1620.3035

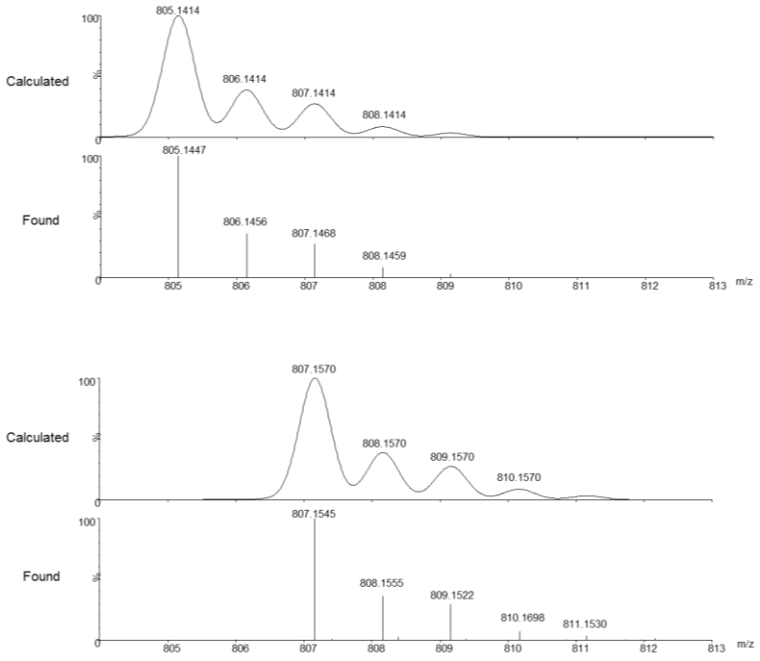


Mixture of **1b**+**2a**+**3g** (0.5, 0.5 and 2.5 mM respectively, pH.6.5, *t* = 24 h)

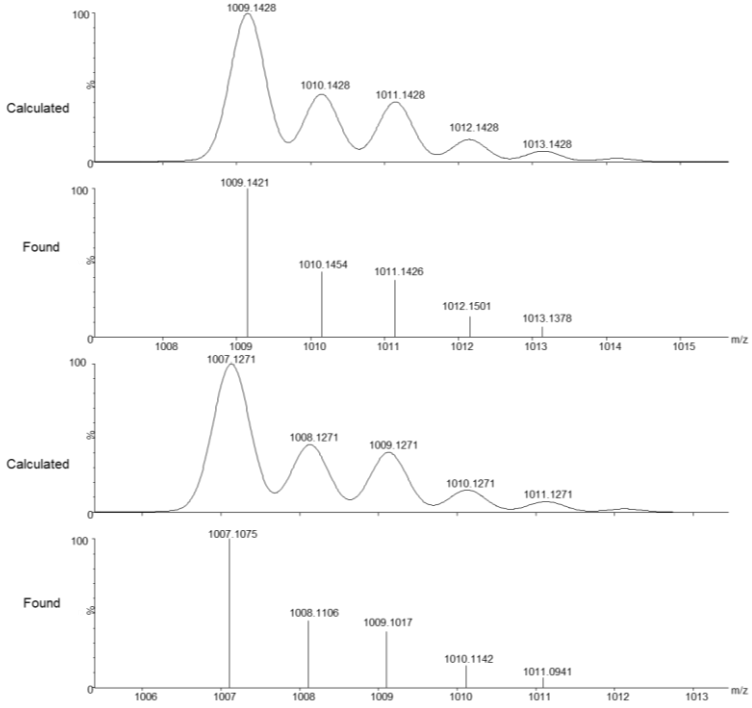


Identification of the products:

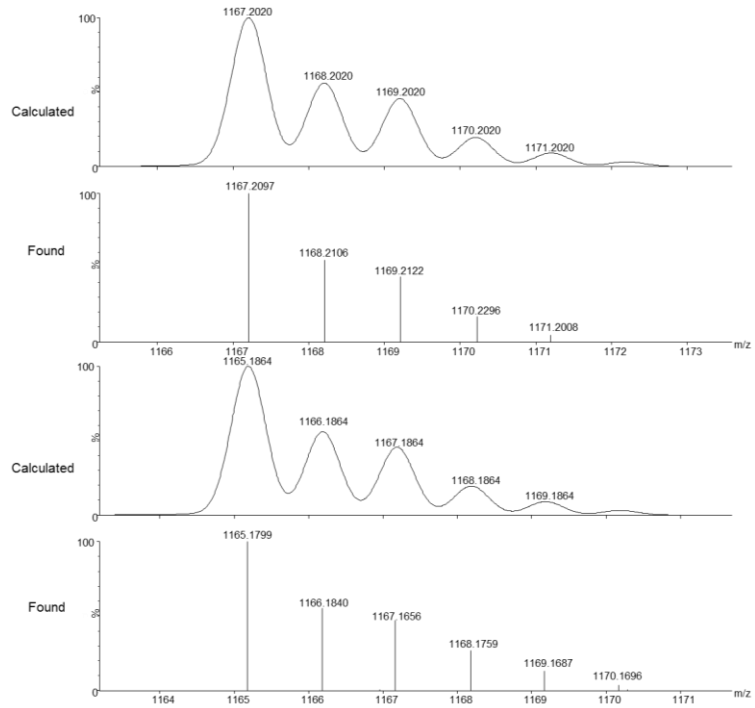
[**2a-3g$_2$**]
Retention time: 9.05 min.
Chemical Formula: $C_{28}H_{38}N_8O_{12}S_4$
Exact Mass: 806.1492



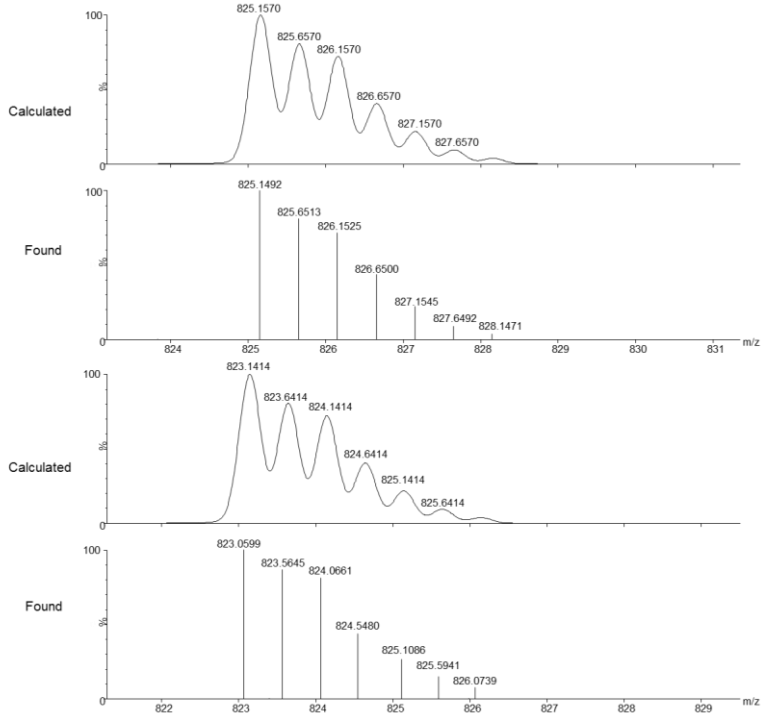
[1b-3g$_3$]
Retention time: 9.05 min.
Chemical Formula: $C_{33}H_{48}N_6O_{18}S_6$
Exact Mass: 1008.1349



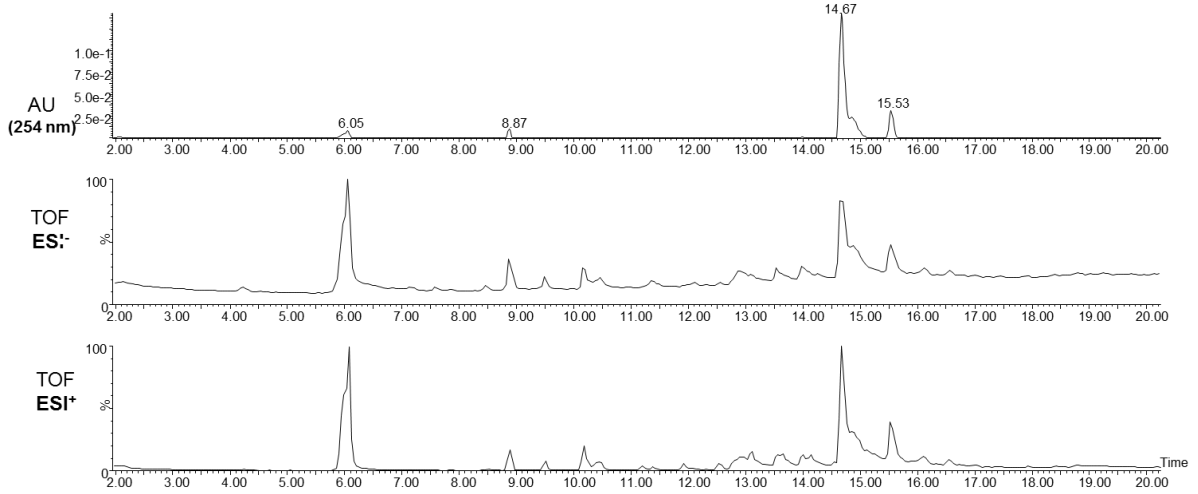
[1b-2a-3g]
Retention time: 10.52 min.
Chemical Formula: $C_{41}H_{54}N_{10}O_{18}S_6$
Exact Mass: 1166.1942



[**1b-2a$_2$-3g**]
Retention time: 11.78 min.
Chemical Formula: $C_{59}H_{76}N_{16}O_{24}S_8$
Exact Mass: 1648.2984



Mixture of **1a**+**2a**+**3h** (0.5 mM each, pH.6.5, *t* = 24 h)

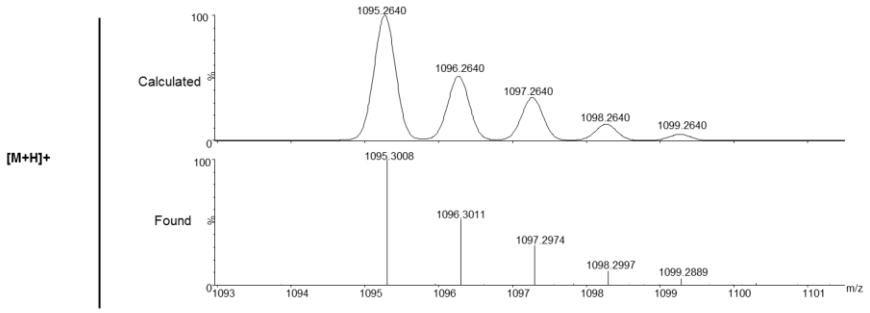


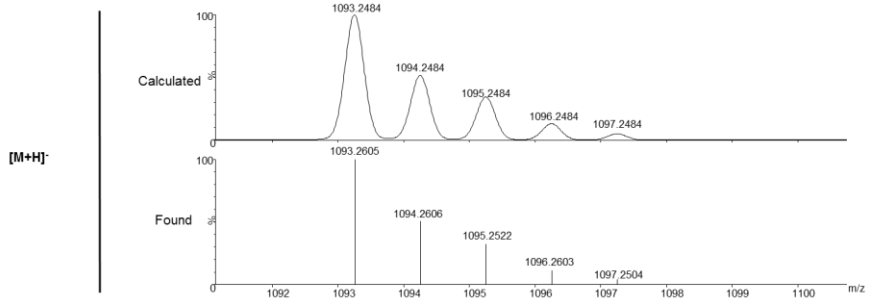
Identification of the products:

[**2a-3h$_2$**]
Retention time: 6.05 min.

Chemical Formula:
$C_{38}H_{54}N_{12}O_{18}S_4$

Exact Mass:
1094.2562



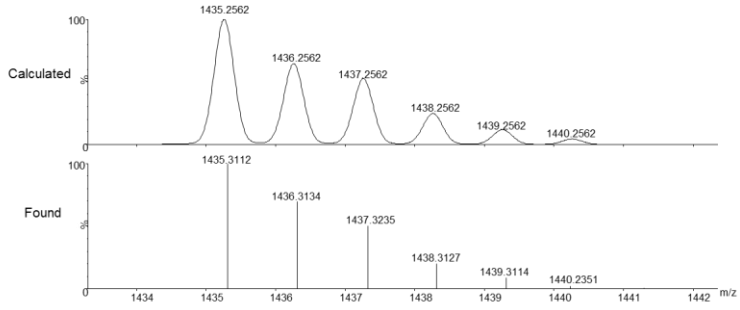


[**1a-3h$_3$**]
Retention time: 8.87 min.

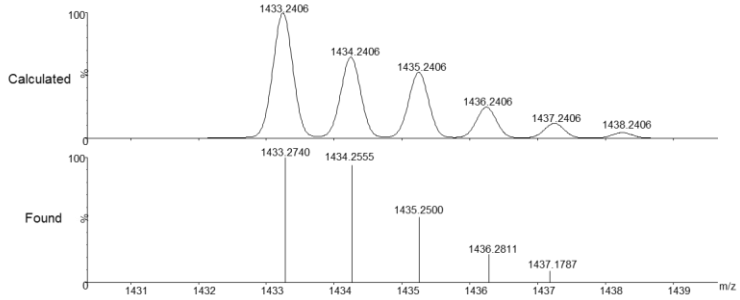
Chemical Formula:
$C_{48}H_{66}N_{12}O_{27}S_6$

Exact Mass:
1434.2485

[M+H]+



[M+H]-

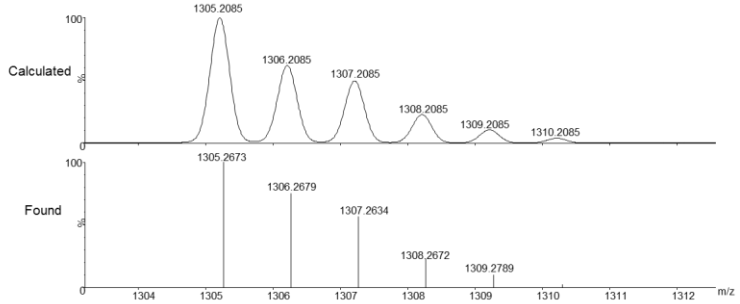


[**1a-2a-3h**]
Retention time: 14.67 min.

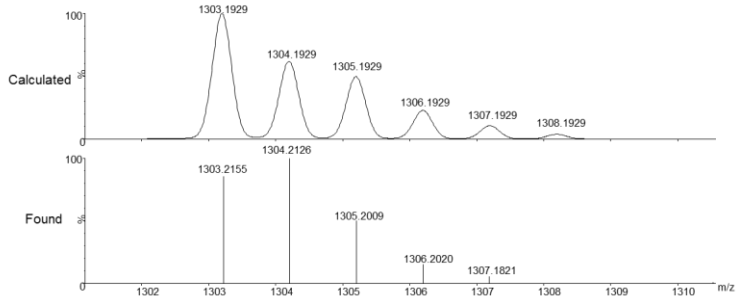
Chemical Formula:
$C_{46}H_{56}N_{12}O_{21}S_6$

Exact Mass:
1304.2007

[M+H]+



[M+H]-

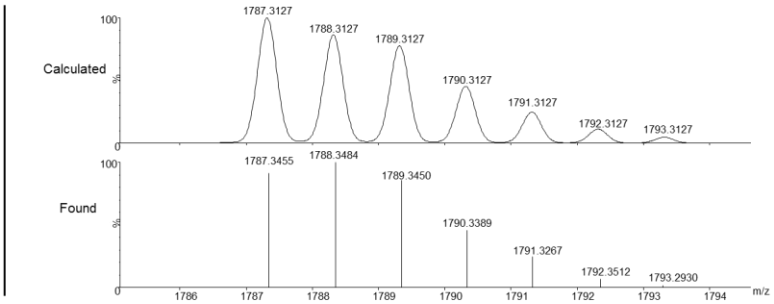


[**1a-2a-3h**]
Retention time: 15.53 min.

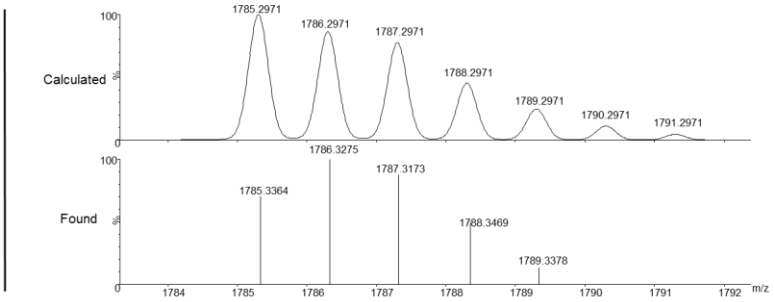
Chemical Formula:
$C_{64}H_{78}N_{18}O_{27}S_8$

Exact Mass:
1786.3049

[M+H]+

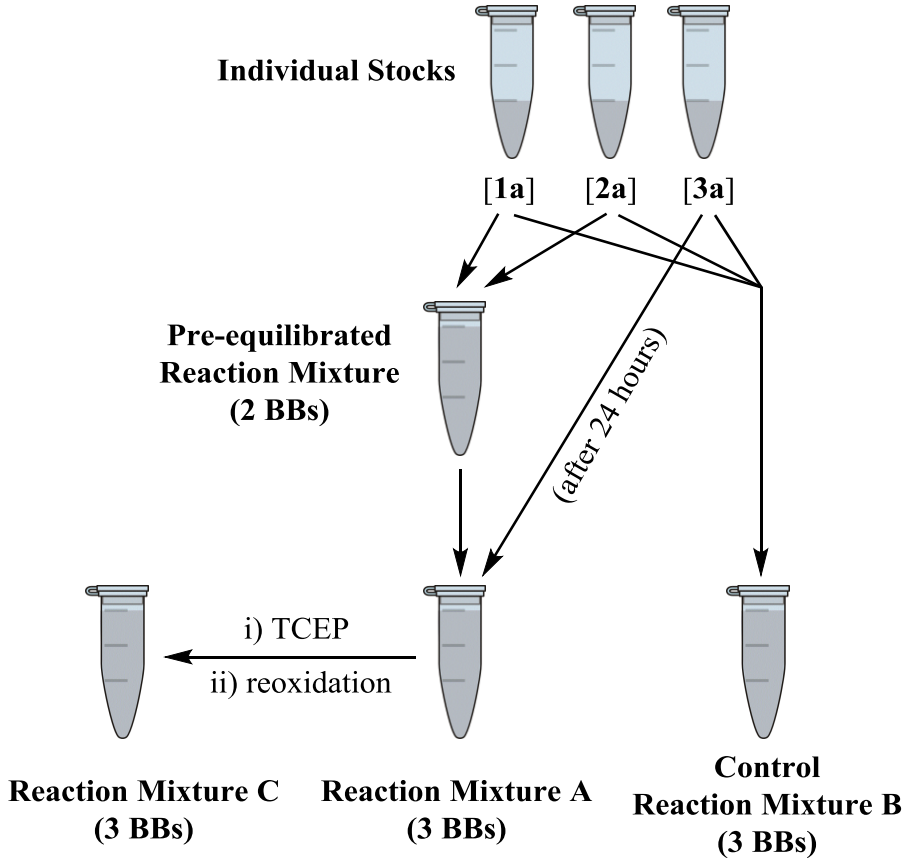


[M+H]-



REVERSIBILITY TEST

Individual stocks of [**1a**], [**2a**] and [**3a**] were prepared (4 mM in DMSO for **1a** and **2a**, and 20 mM in milli-Q water for **3a**) (see Scheme S1). From these, a **pre-equilibrated reaction mixture** was prepared by adding 40 μL of [**1a**] and [**2a**] to 200 μL of a 66.7 mM BIS-Tris methane buffer (pH 6.5). The **individual stock** of [**3a**] was stored at -80 ºC. After 24 hours, the **reaction mixture A** was prepared by adding 30 μL of the **individual stock** of [**3a**] to 210 μL of the **pre-equilibrated reaction mixture**.



Scheme S1: Preparation of the reaction mixtures of the reversibility test.

Simultaneously, the **control reaction mixture B** was prepared by mixing 30 μL of each **individual stock** with 150 μL of a 66.7 mM BIS-Tris methane buffer (pH 6.5). After 24 hours, the **reaction mixture A**, the **control reaction mixture B** and the **pre-equilibrated reaction mixture** were analysed by HPLC (see Figure S28).

Finally, the **reaction mixture C** was prepared by adding 0.35 equivalents of Tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl)[4] 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, 24 hours later, the **reaction mixture C** was analysed by HPLC (see Figure S28).

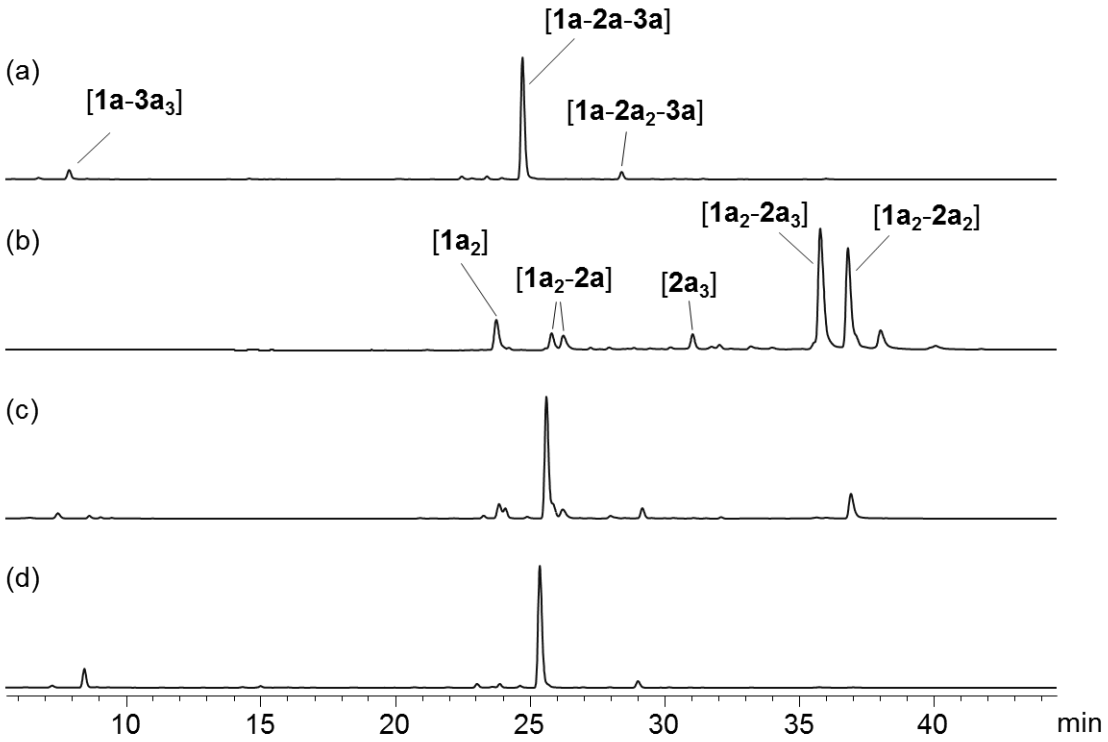
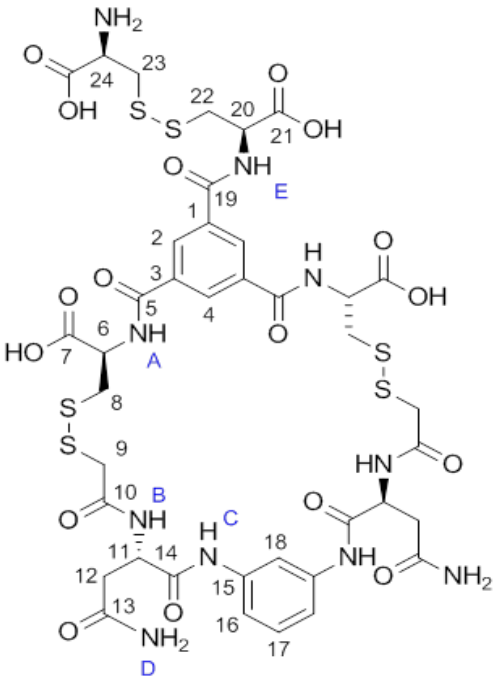


Figure S28: HPLC traces of the **control reaction mixture B** (a), the **pre-equilibrated reaction mixture** (b), the **reaction mixture A** (c) and the **reaction mixture C** (d).

SYNTHESIS AND NMR CHARACTERIZATION OF **1a2a3a**



For the synthesis of the trimer **1a2a3a**, [**1a**] (5.2 mg, 0.010 mmol), [**2a**] (4.8 mg, 0.010 mmol), and [**3a**] (2.4 mg, 0.20 mmol) were dissolved in 2 mL of MeOH and 1 mL of milli-Q water. Then, the pH of the solution was adjusted to 7.0 by the addition of NaOH (aq). After 6 days at room temperature, complete oxidation of the starting reagents was observed by HPLC. The MeOH was evaporated *in vacuo* and the residue was lyophilized. The resulting white solid was purified by reversed-face flash chromatography (gradient: from 5% to 15% CH_3CN in H_2O) and 3.8 mg of pure **1a2a3a** (34% yield) were obtained as a white solid. HRMS (ESI-) calcd. for $C_{39}H_{46}N_{10}O_{17}S_6$ $[M-H]^-$ (m/z): 1117.1289, found: 1117.1323. ^{1}H NMR (500 MHz, H_2O:DMSO-d_6 (85:15)): δ = 8.69 (s, 2H, H_C), 8.62 (s, 1H, H_4), 8.54 (d, *J* = 8.4 Hz, 2H, H_B), 8.46 (d, *J* = 8.8 Hz, 2H, H_A), 8.38 (d, *J* = 8.2 Hz, 1H, H_E), 8.29 (d, *J* = 1.7 Hz, 2H, H_2), 7.58 (s, 2H, H_D), 7.20 (d, *J* = 1.9 Hz, 1H, H_{18}), 6.94 (t, *J* = 8.2 Hz, 1H, H_{17}), 6.89 (d, *J* = 8.2 Hz, 2H, H_{16}), 6.76 (s, 2H, $H_{D'}$), 4.05 – 4.00 (m, 1H, H_{24}), 4.01 (d, *J* = 14.8 Hz, 2H, H_9), 3.47 (dd, *J* = 15.2 Hz, *J* = 2.4 Hz, 1H, H_8), 3.45 (d, *J* = 2.2 Hz, 1H, $H_{9'}$), 3.35 (dd, *J* = 9.9 Hz, *J* = 3.9 Hz, 1H, H_{22}), 3.32 (dd, *J* = 10.7 Hz, *J* = 4.3 Hz, 1H, H_{23}), 3.19 (dd, *J* = 15.2, 11.0 Hz, 1H, H_8), 3.04 (dd, *J* = 15.2, 9.3 Hz, 1H, $H_{22'}$), 2.99 (dd, *J* = 15.8, 9.9 Hz, 1H, $H_{23'}$), 2.93 – 2.82 (m, 2H, H_{12}). ^{13}C NMR (125 MHz, D_2O:DMSO-d_6 (85:15)): δ = 174.6 (CO, C_{13}), 172.1 (CO, C_{10}), 169.6 (CO, C_{14}), 167.2 (CO, C_{19}), 166.8 (CO, C_5), 137.0 (C, C_{15}), 134.0 (C, C3), 133.4 (C, C1), 129.7 (CH, C_4), 129.65 (CH, C_{16}), 129.60 (CH, C_2), 116.3 (CH, C_{17}), 111.5 (CH, C_{18}), 55.8 (CH, C_6), 54.4 (CH, C_{20}), 51.0 (CH, C_{11}), 53.4 (CH, C_{24}), 45.0 (CH_2, C_8), 41.2 (CH_2, C_9), 39.5 (CH_2, C_{22}), 37.9 (CH_2, C_{23}), 35.9 (CH_2, C_{12}).

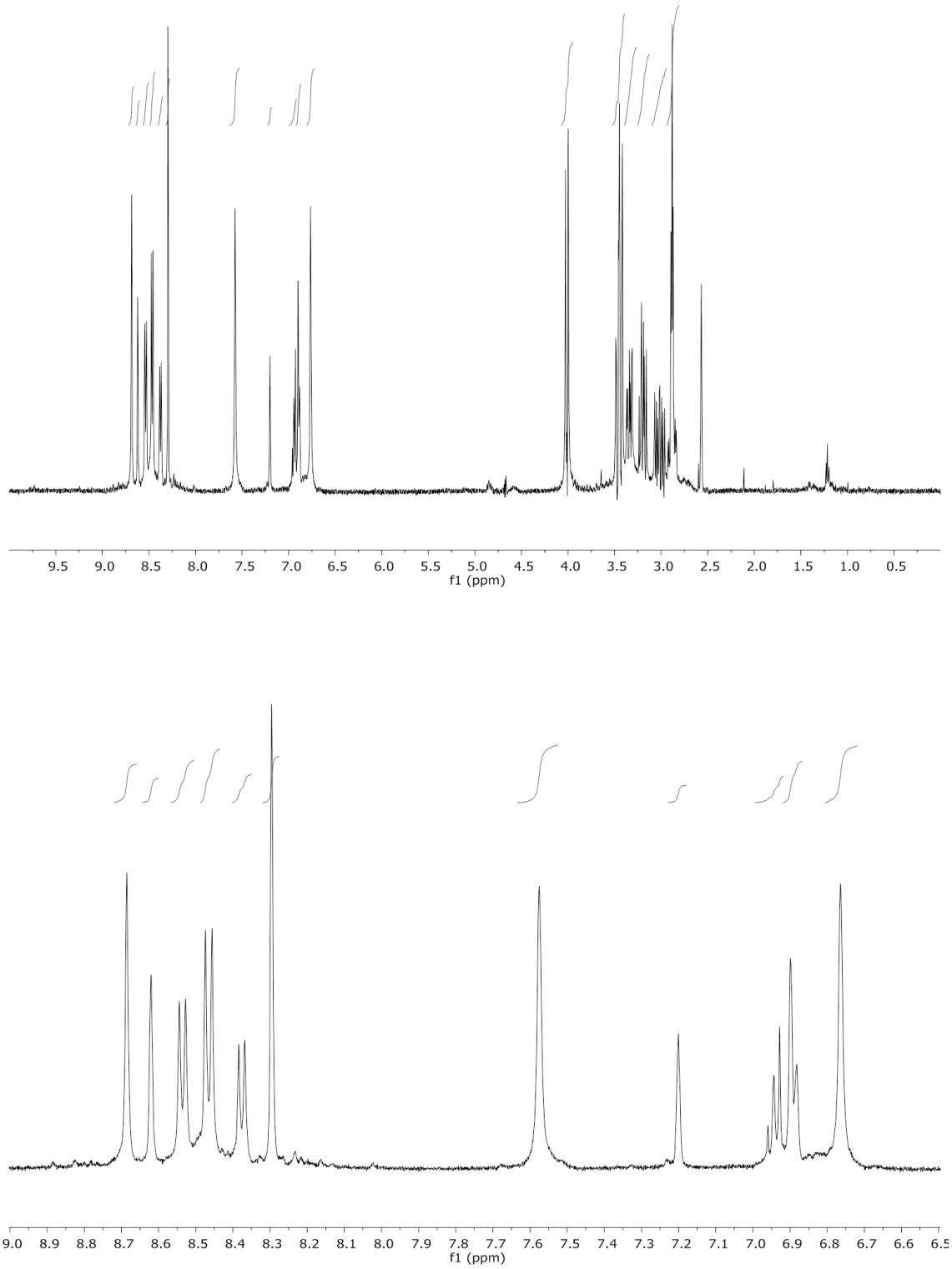


Figure S29: ^{1}H (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectrum of **1a2a3a** (8.5 mM phosphate buffer, pH 6.5), and expansion of the amide region (9.0 – 6.5 ppm).

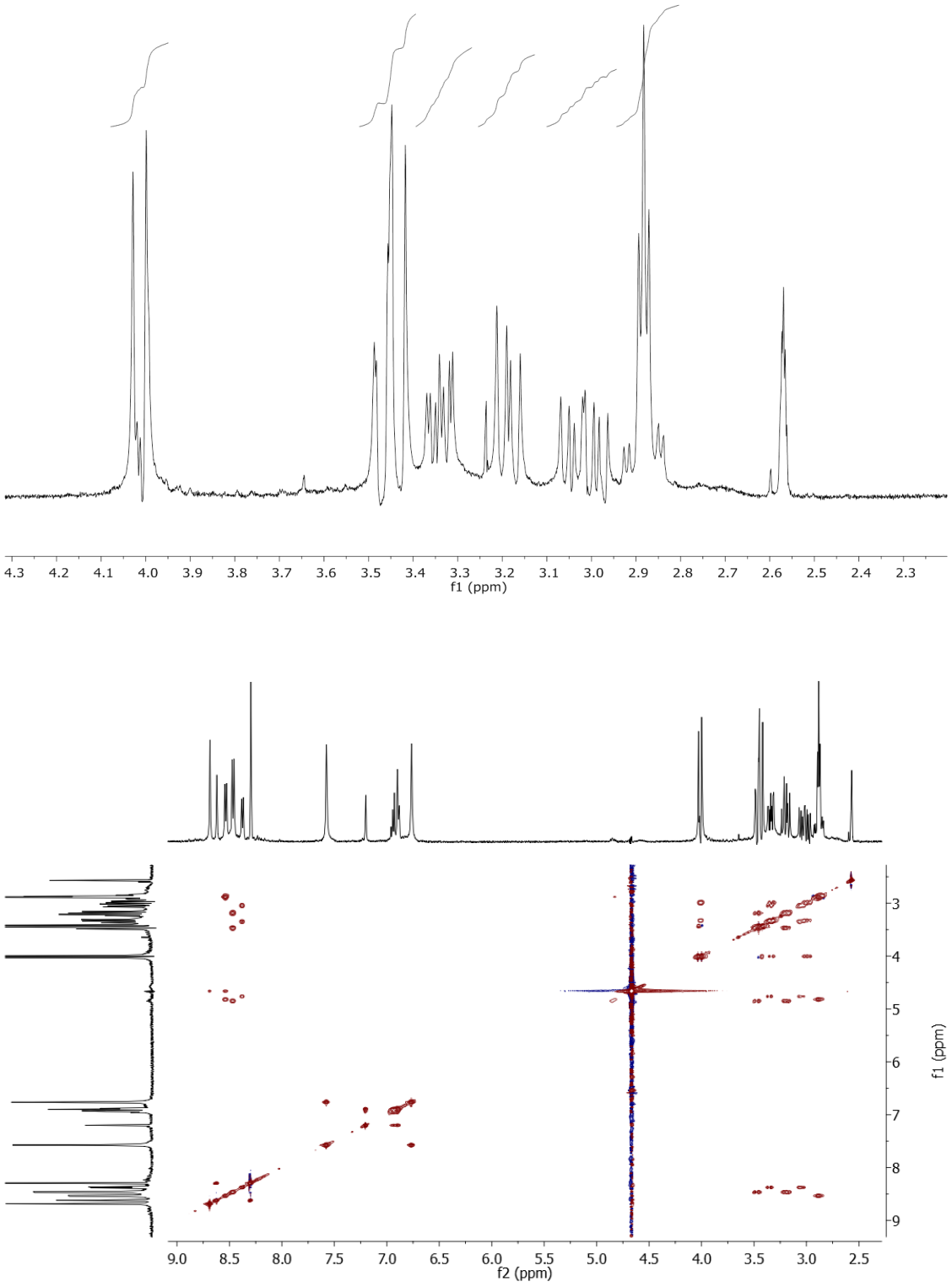


Figure S30: Expansion of the aliphatic region (4.3 – 2.3 ppm) of ^{1}H (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) and ^{1}H 2D TOCSY (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectra of **1a2a3a** (8.5 mM phosphate buffer, pH 6.5).

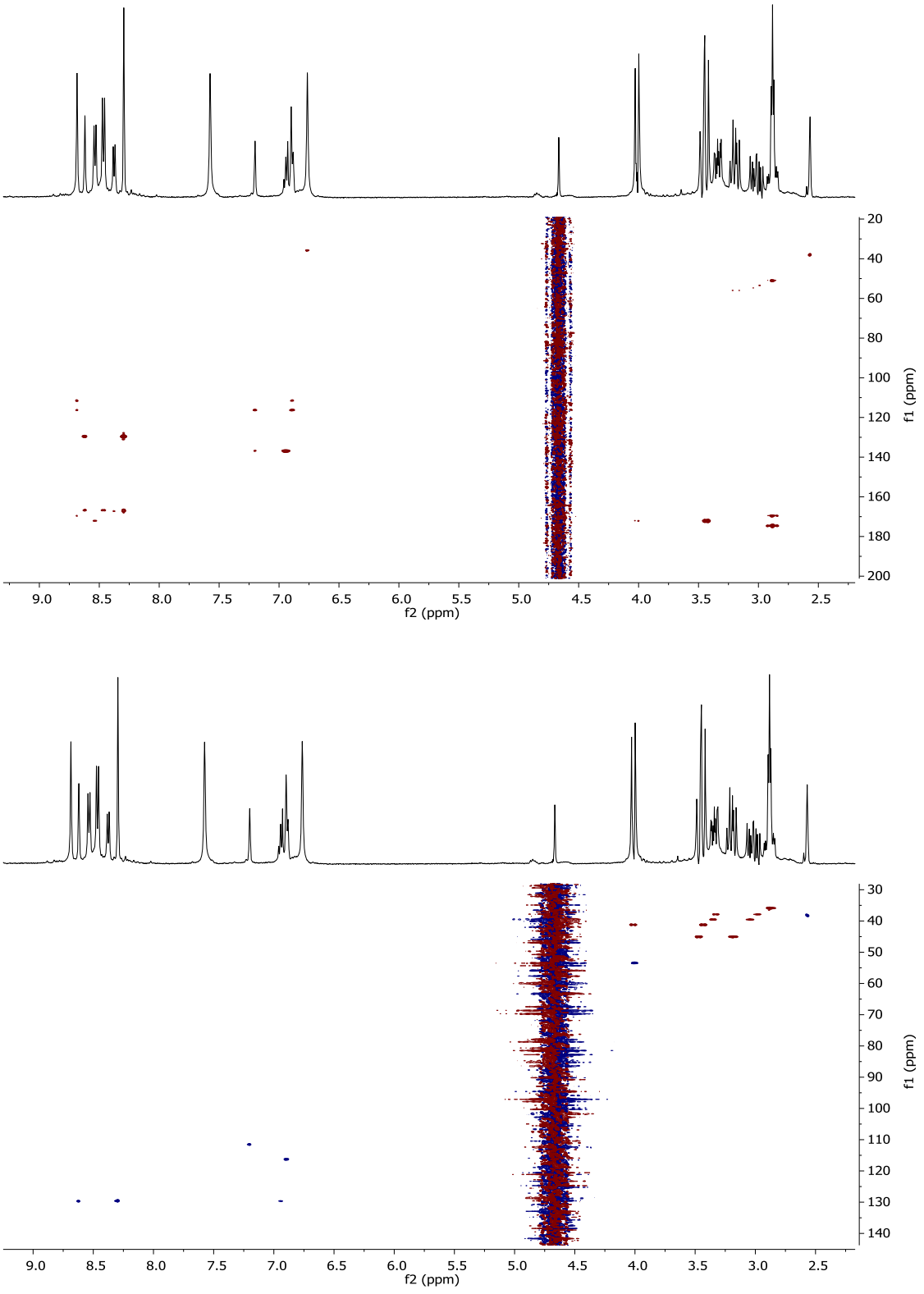


Figure S31: ^{1}H-^{13}C gHMBC (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) and ^{1}H-^{13}C gHSQC (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) of **1a2a3a** (8.5 mM phosphate buffer, pH 6.5).

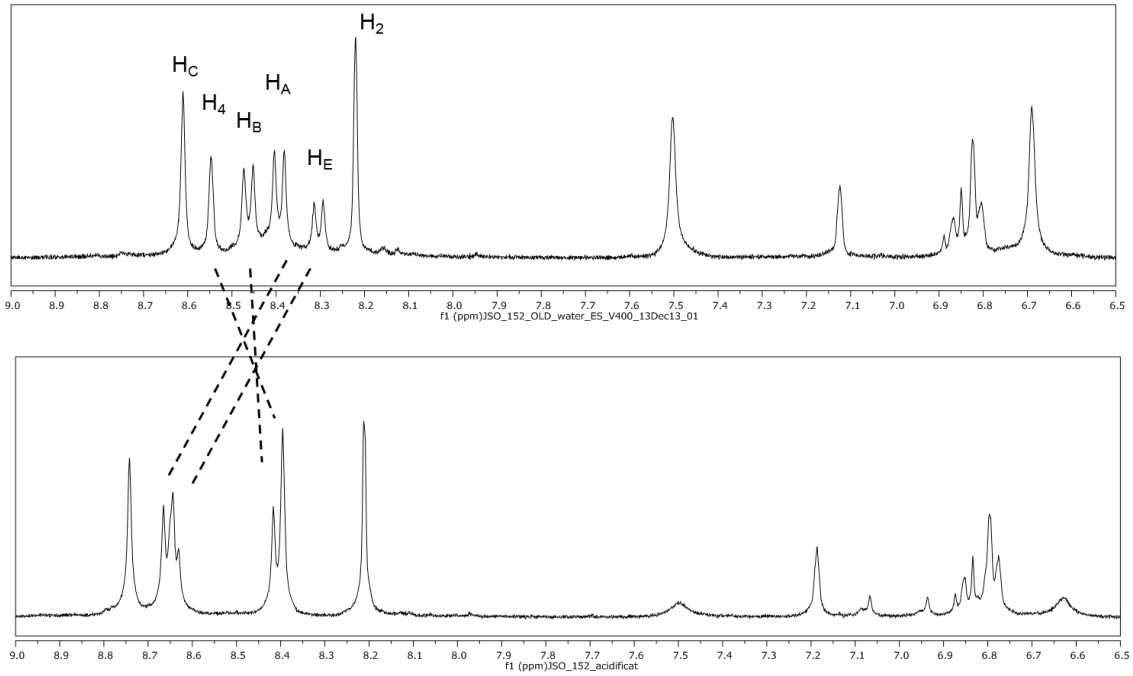


Figure S32: Expansion of the amide region (9.0 – 6.5 ppm) of ^{1}H (400 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectrum of **1a2a3a**. Top: 8.5 mM phosphate buffer, pH 6.5. Bottom: after acidification with TFA, pH ~2. The differences in chemical shift for protons H_A, H_E (moving downfield) and H_4 (moving upfield) are consistent with an unfolding of the proposed conformation for **1a2a3a**.

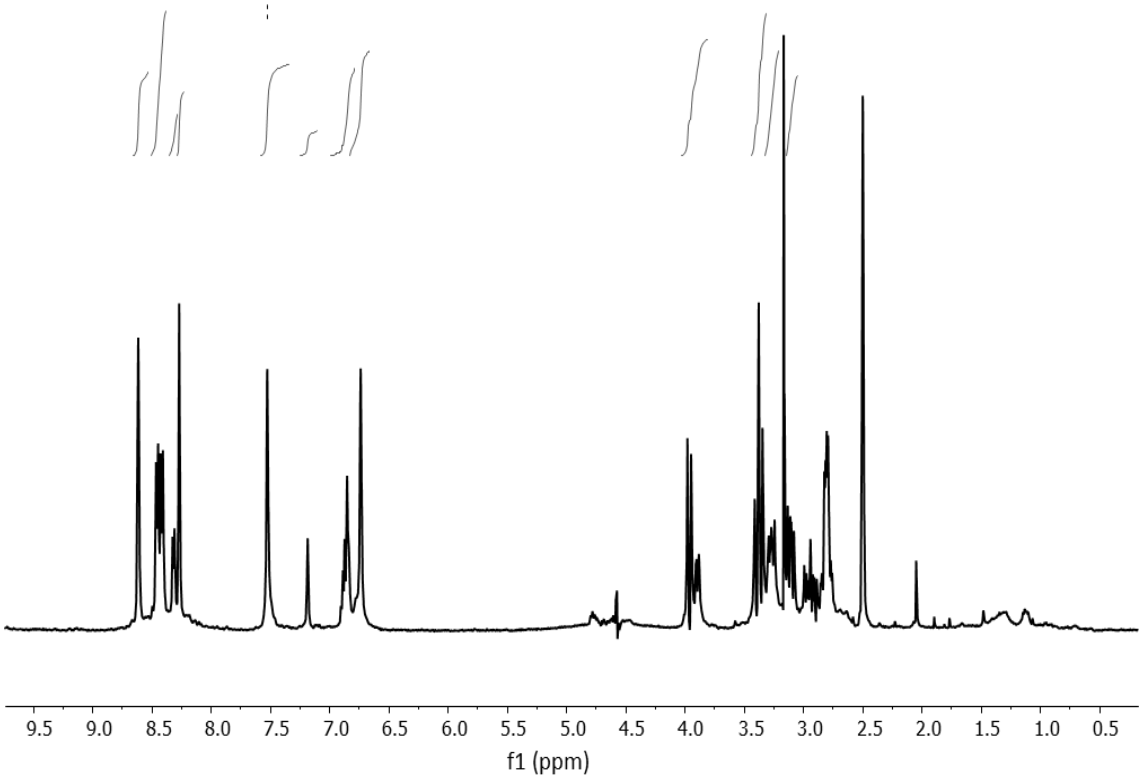
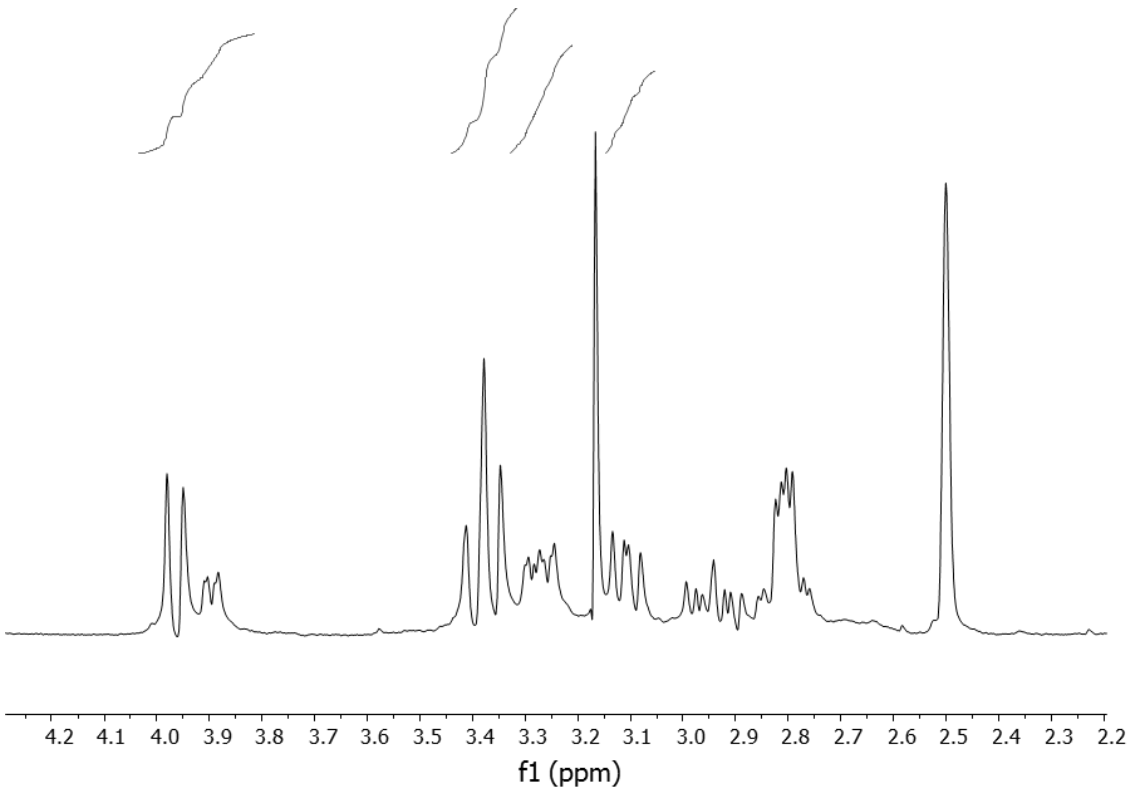


Figure S33: ^{1}H (500 MHz, 298 K in H_2O:DMSO-d_6 (85:15), with water suppression by excitation sculpting scheme) spectrum of **1a2a3a** (8.5 mM phosphate buffer, pH 6.5).



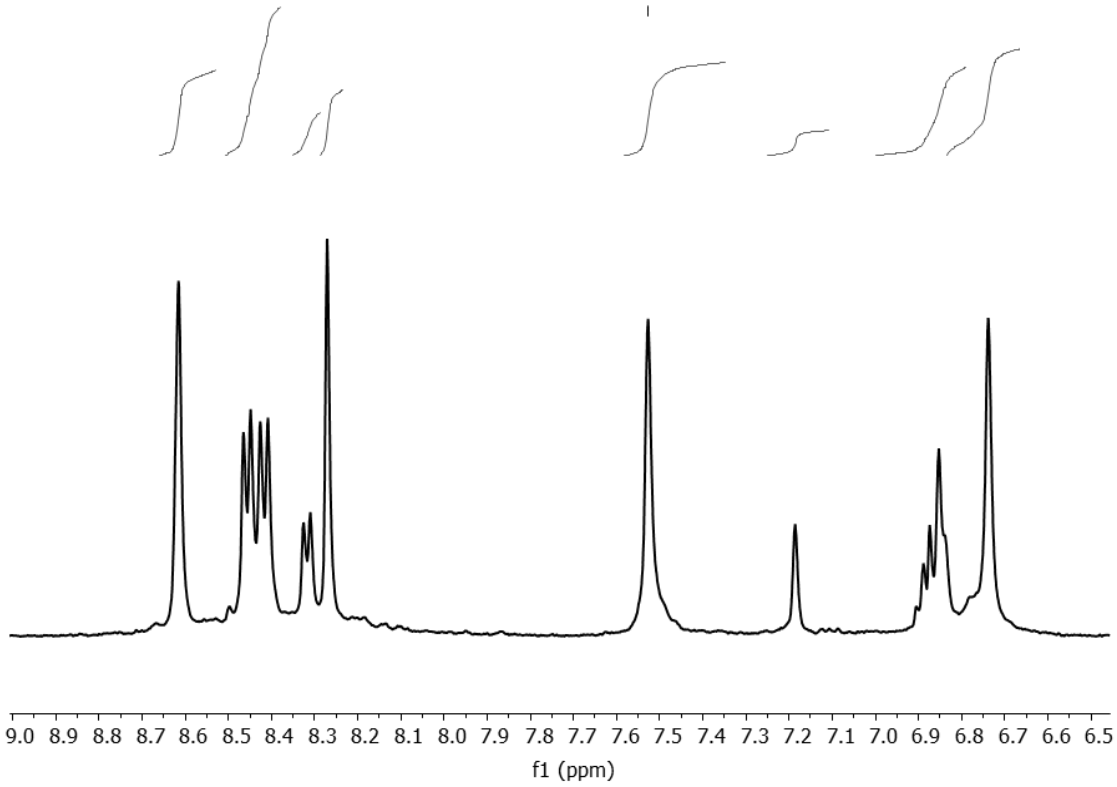


Figure S34: Expansion of the aliphatic region (4.2 – 2.2 ppm) of ^{1}H (500 MHz, 298 K in H_2O:DMSO-d_6 (75:25), with water suppression by excitation sculpting scheme) and and expansion of the amide region (9.0 – 6.5 ppm).

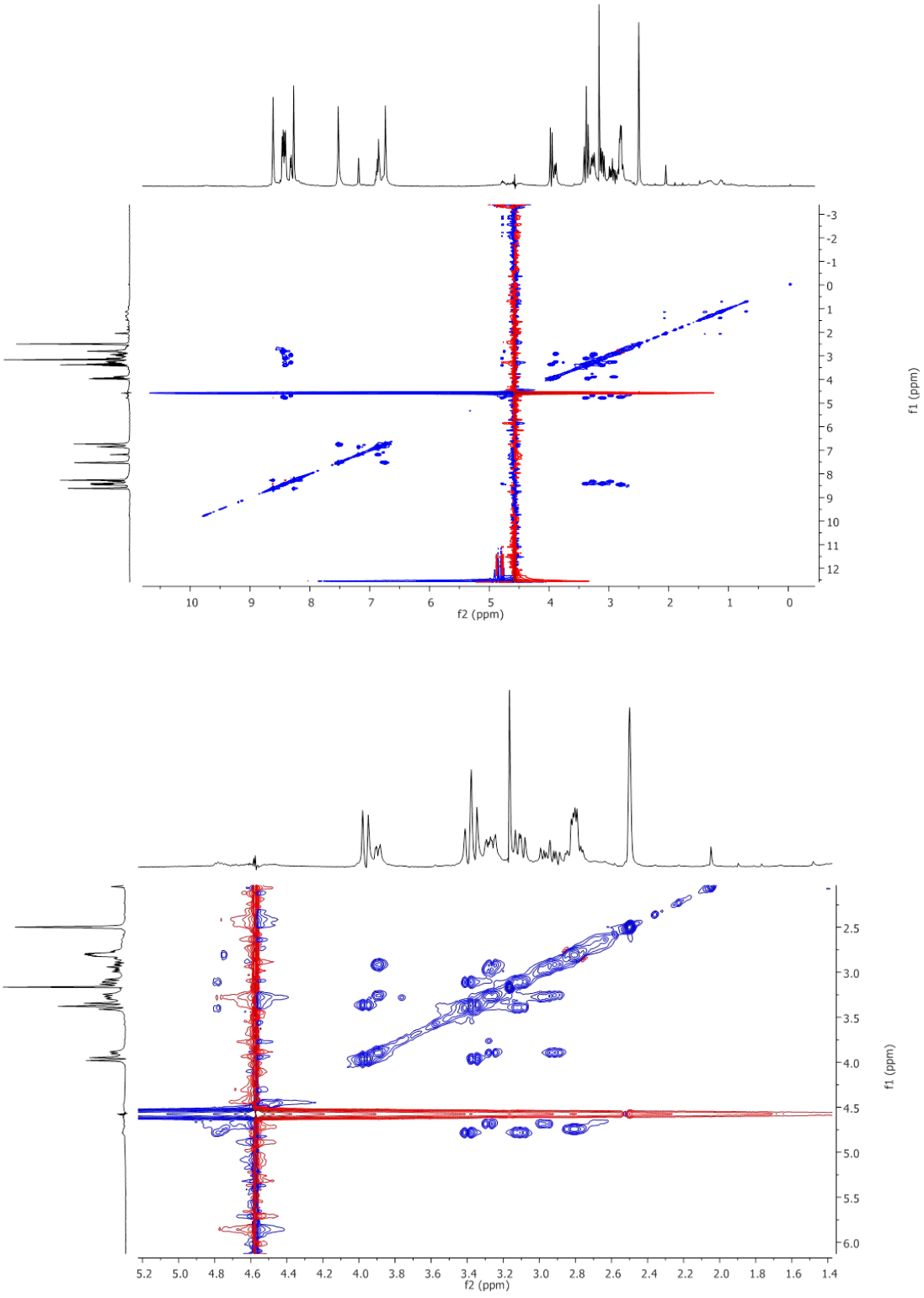


Figure S35: ^{1}H 2D TOCSY (500 MHz, 298 K in H_2O:DMSO-d_6 (75:25), with water suppression by excitation sculpting scheme) spectra of **1a2a3a** (8.5 mM phosphate buffer, pH 6.5) and expansion of the aliphatic region (5.2 – 1.4 ppm).

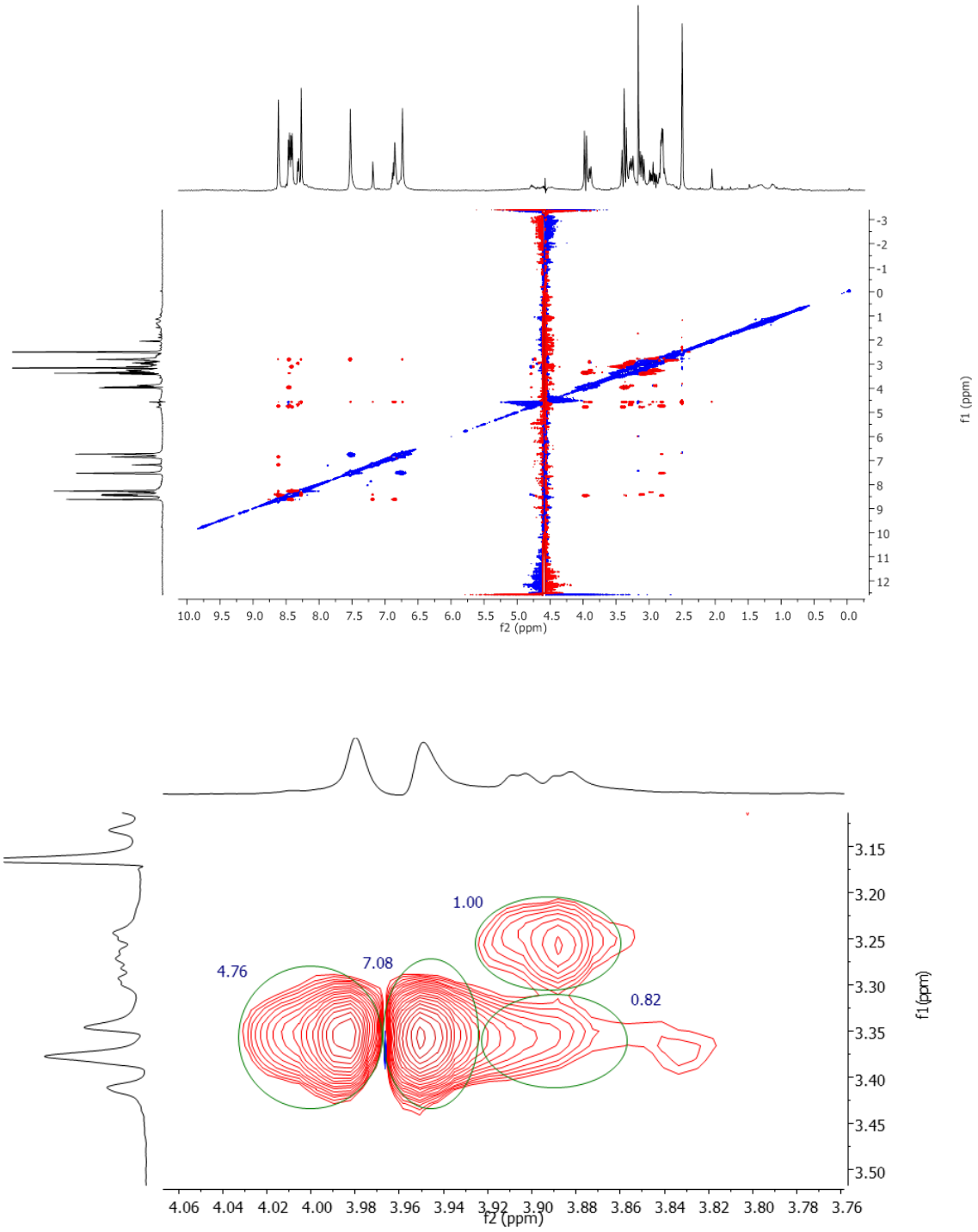


Figure S36: ^{1}H ROESY (500 MHz, 298 K in H_2O:DMSO-d_6 (75:25), with water suppression by excitation sculpting scheme) spectra of **1a2a3a** (8.5 mM phosphate buffer, pH 6.5), and expansion the region for proton H20 showing the critical contacts in the folded structure with H9.

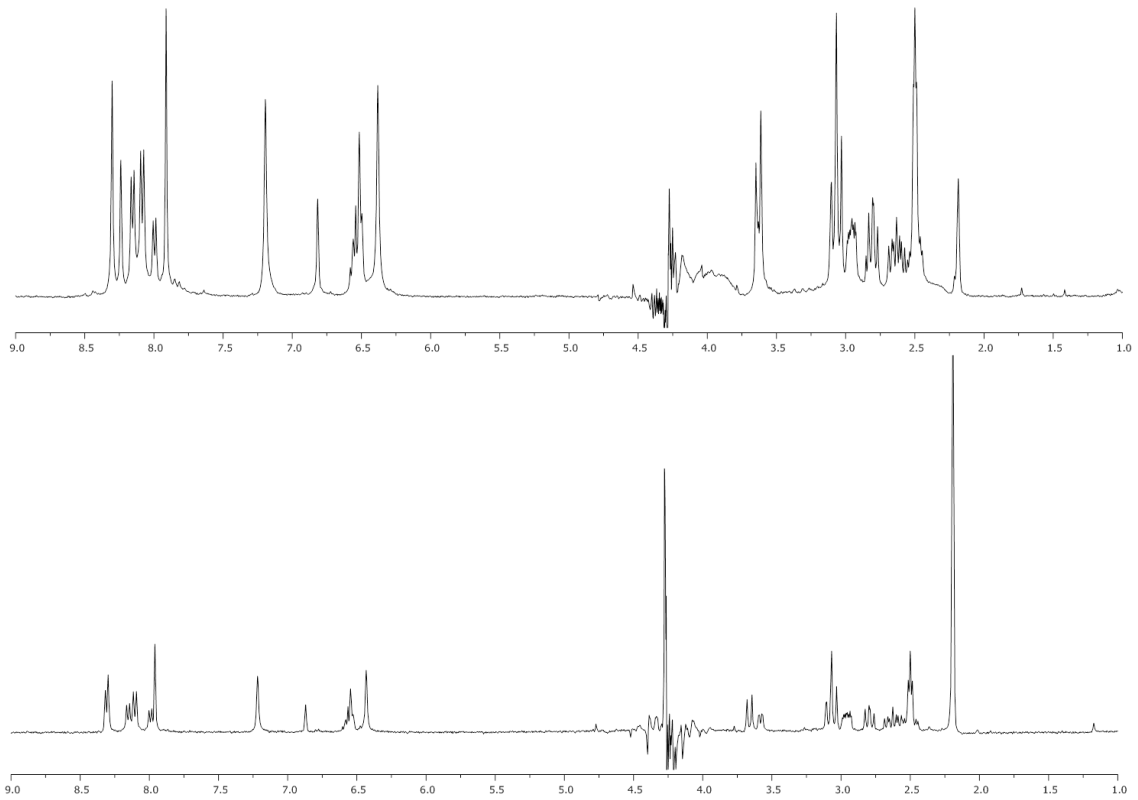


Figure S37: ^{1}H NMR (400 MHz, 298K) with water suppression by excitation sculpting scheme for **1a2a3a**. Top: H_2O:DMSO-d_6 (85:15, 8.5 mM phosphate buffer, pH 6.5). Bottom: H_2O:DMSO-d_6 (75:25, 8.5 mM phosphate buffer, pH 6.5). Subtle changes are appreciated, in a more organic solvent the folded structure of **1a2a3a** is more favored as can be seen by downfield shift for protons H_4 and H_2.

REFERENCES

[1] **Related references:** (a) W. Brandt, T. Herberg, L. Wessjohann, *Biopolymers (Protein Science)*, **2011**, *96*, 651-667; (b) C. F. Rodriquez, G. Orlova, Y. Guo, X. Li, C.-K. Siu, A. C. Hopkinson, K. W. M.Siu, *J. Phys. Chem. B* **2006,** *110*, 7528-7537; (c) E. F. Strittmatter, E. R. Williams, *J. Phys. Chem. A* **2000,** *104*, 6069-6076; (d) M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, R. A. Friesner, *J. Am. Chem. Soc.* **1997,** *119*, 5908-5920.

[2] K. R. West, K. D. Bake and S. Otto, *Org. Lett.*, **2005**, *7*, 2615-2618

[3] (a) J. Atcher, A. Moure and I. Alfonso, *Chem. Commun.*, 2013, **49**, 487-489. (b) J. Atcher and I. Alfonso, *RSC Adv.*, 2013, **3**, 25605-25608. (c) J. Solà, M. Lafuente, J. Atcher and I. Alfonso, *Chem. Commun.*, 2014, **50**, 4564-4566.

[4] 5. J. A. Burns, J. C. Butler, J. Moran, and G. M. Whitesides, *J. Org. Chem.*, **1991**, *56*, 2648-2650.

2.5. PUBLICATION C:

Structurally selective assembly of a specific macrobicycles from a dynamic library of pseudopeptidic disulfides.

To be submitted

Structurally selective assembly of a specific macrobicycle from a dynamic library of pseudopeptidic disulfides

Abstract:

Molecular recognition is essential in many chemical and biological processes. Studying the behavior of pseudopeptides using dynamic covalent chemistry allows exploring a wide range of structural components and molecular interactions with a minimal synthetic effort. Here, we describe how non-covalent attractive forces in pseudopeptidic building blocks can succeed in guiding the product distribution in a dynamic library towards topologically more complex compounds which are in principle not expected. The interactions described herein are highly dependent on molecular architecture and media so effective recognition can be altered by just minimal structural or environmental changes. Thus, the chemical and constitutional information contained in the respective building blocks is decoded and expressed through dynamic covalent and non-covalent bonds in the assembling of either a single macrostructure or an ensemble of components with larger structural diversity. The understanding of supramolecular forces responsible for the component assembly in minimalistic systems can help to comprehend more complex bio-related processes such as protein folding or protein-protein interactions.

Introduction

Pseudopeptidic compounds are intriguing molecules[1] that have shown very remarkable applications in various fields, such as supramolecular chemistry,[2] biological chemistry[3] and in the preparation of nanomaterials.[4] The combination of abiotic fragments that can confer enzymatic stability with the chemical information that derive from the amino acid functionalities make them especially attractive fragments to study molecular interactions in simplified models with biological significance.[5] The variety of functionalities present in natural amino acids[6] allows to generate diversity and to explore a plethora of non-covalent contacts. In this context, dynamic constitutional (covalent) chemistry (DCC)[7] provides a valuable tool to study intramolecular and intermolecular interactions. DCC allows generating mixtures of compounds using reversible bond formation which allows the exchange between different subspecies (different building blocks) giving rise to diverse libraries.[8] As the composition of the mixture is determined by the stability of each member and the global energy of the system, the predominance of unexpected compounds or deviation from 'statistical' libraries allows to identify different self-recognition[9] or even autocatalytic processes, being the emergence of self-replicating systems[10] maybe the most relevant example for its significance in the origin of life.[11] Thanks to the ability of DCC to generate complexity, the multiple potential applications and the advance in analytical tools which now enable dealing with mixtures, the study of dynamic systems is receiving increasing interest. In the last years, we have focused on the study of covalent dynamic libraries based in pseudopeptidic compounds in aqueous media.[12] We found that disulfide chemistry is especially suited for this purpose as we developed conditions that allow us to rapidly form libraries in a wide range of pHs.[13] In addition, we showed that some of these libraries respond to external stimuli such as the increase of ionic strength or the presence of biological macromolecules.[12a-d] This methodology had also been employed by different research groups to obtain receptors for relevant biomolecules,[14] but also has been proved a valid tool for the synthesis of topologically complex molecules like catenanes[15] and knots. [9a]

During our investigations we became particularly interested in increasing molecular and topological diversity to generate more complex systems. Commonly, most of the libraries described earlier were based in C2 symmetrical bifunctional components and, therefore, they combine to assemble more or less complex structures that can always close one or several links forming macrocycles.[9a,14,15,16] Alternatively, although less explored, the combination of building blocks having different geometries and different number of functionalities that form

dynamic bonds can give rise to other final structures.[17] We decided to study dynamic libraries that could potentially form structures other than macrocycles in aqueous environments. To this aim, we began to investigate libraries obtained by mixing pseudopeptides with a different number of thiol functional groups, in other words, different 'valence' in the DCC process. In these libraries we could eventually observe self-recognition processes able to drive the library towards the exclusive formation of a single structure.[18] Later on, this knowledge allowed us to rationally design a dynamic chemical network that reorganizes in the presence of an analyte (cysteine or cystine) liberating a fluorescent compound as readable output. Thus, by knowing the exact nature of the interactions we were able to generate a system that acts as a sensor for a biologically relevant molecule.[19] This example illustrates that understanding the recognition processes that guide virtually complex libraries is very appealing and allows to reveal structural and chemical complementarities that can be further used in the design of molecular devices.[20] In this article, we show two dissimilar behaviors in the final output of chemical libraries that involve analogous compounds which may only differ only in one methylene unit. In the absence of complementarity, a wide structural diversity of compounds is formed, as expected. However, in the case of a perfect chemical and structural match between a tripodal building block (containing 3 thiol functionalities) and a bipodal moiety (a member that contains two SH) an unexpected [2+2] cage-like compound emerges selectively from the potential complex mixture as a single product. Our results illustrate how subtle changes in the structure of one compound result in big changes in the final outcome of the whole system. Moreover, the delicate structural and functional information contained in the BBs define the final composition of the library, in the way of either a unique macro-assembly or a complex mixture.

Results and Discussion

The pseudopeptidic building blocks necessary for the generation of the different libraries under study are easily accessible by straightforward peptide synthesis protocols in solution (Figure 1). Tripodal compound **1** and bipodal unit **3** are obtained in two steps from H-Cys(Trt)OtBu and the corresponding trimesic[21] and isopthalic[22] acids respectively. Bipodal components **2a-e** are obtained in a 4 step-procedure using the methodology described earlier in our group (Figure 1, see ESI for detailed synthetic procedures).[12c]

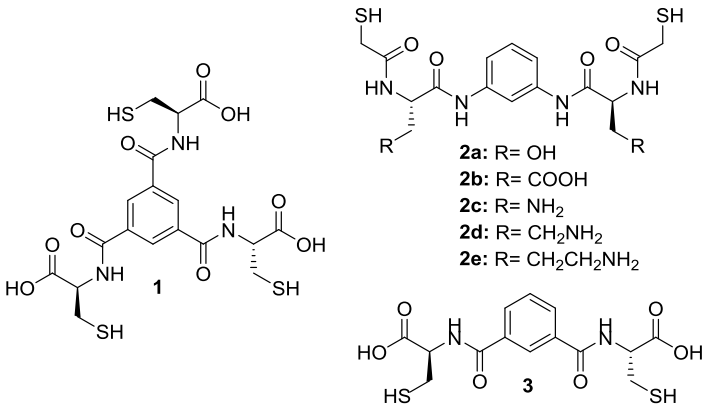


Figure 1. Chemical structures of the pseudopeptidic building blocks studied in this article.

Compound distribution of a DCL containing different functional groups. We started by analyzing the products formed in a dynamic library resulting from mixing the tripodal scaffold **1** and different bipodal units with different functionalities and charge (**2a-c**) in an individual and equimolar manner. The reactions were performed in aqueous media containing 25% of DMSO at pH 6.5 using a bis-tris buffer (50 mM); 0.5 mM concentrations for each of the compontents were used in each experiment. Not surprisingly, the mixtures that contained **1** and either **2a** or **2b** revealed the formation of several products with different architectures as can be detected by HPLC-MS analysis. As expected, the negatively charged compound **1** efficiently combined with neutral **2a**, the major compounds joined two units of **1** with either two or three molecules of **2a** (**[1_2-$2a_2$]**, **[1_2-$2a_3$]**). In addition, the corresponding homodimers (**[1_2]**, **[$2a_2$]**) were also detected as minor components (Figure 1). Interestingly, when **1** was combined with a negatively charged bipodal compound **2c** the system was driven towards the formation of the homodimeric species that are presumably more stable than the mixed compounds, which would bring together more negative charges. However, when the same experiment was performed with the negatively charged **1** and a positively charged bipodal component, **2c**, an exclusive single product was formed, corresponding to a heterotetramer **[1_2-$2c_2$]** containing two tripodal units (**1**) and two bipodal builging blocks **2c** (Figure 2C). The newly formed species showed a remarkable stability as the library showed no alteration when 2.5 mM of L-cysteine (enough to saturate all the library thiols) was added to the mixture (Figure S19). We reasoned that this stability would be due to the establishment of highly stabilizing forces between the counterparts that would find a perfect fit in charge and space.

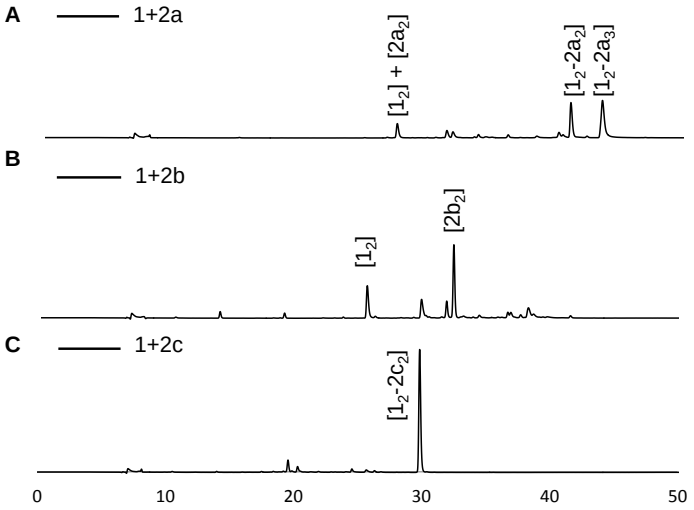


Figure 2. HPLCs traces (UV detection at 254 nm) for the DCLs formed by **1** and **2a, 2b or 2c** (0.5mM each) in aqueous solution (25%DMSO, pH 6.5).

This huge stability was further confirmed by other control experiments: increasing amounts of cysteamine (2-mercaptoethylamine) were added in an attempt disrupt the formation of the tetramer; however it was necessary to add up to 50 times more cysteamine with respect to the other BBs to start forming other species (Figure S24).

Effect of the reaction media. At this point we reasoned that the divergence in the libraries behavior must be due to the establishment of strong non-covalent stabilizing interactions between the functional groups of the amino acid side-chains present in each of the counterparts. To prove the polar nature of these interactions we varied the reaction conditions (polarity and pH) and evaluated the outcome of the reaction. As expected, the selectivity towards the formation the selected tetramer compound was greatly affected by the polarity of the environment: an increase of the ionic strength of the medium led to a diminished predominance of the [2+2] component. The addition of NaCl prevents the formation of **[1$_2$-2a$_2$]** and its area in the HPLC-UV dropped from more than 95% in the absence of salt to less than 30% when 1.0 M NaCl was added to the reaction; the addition of more salt had then a limited effect (Figure 3, blue trace). Non-surprisingly, the pH also plays a key role in the nature of the species formed. The optimal pH for the formation of **[1$_2$-2a$_2$]** within the building blocks during the assembly process is around 6.5; at this pH amino side-chains are positively charged whereas the carboxylic acid functionalities are in their carboxylate form. Moving away from this value changes the protonation state of either basic or acidic side-chains and as a result it is detrimental for self-recognition (Figure 3, red trace). These combined results highlight the

nature of polar interactions (salt bridges and hydrogen bonds) which are responsible for the emergence of a single species. Interestingly, the high selectivity towards **[1_2-$2a_2$]** formation was retained down to 1% DMSO (Figure S30) while lower concentrations of DMSO resulted in partial precipitation, which precluded a reliable analysis of the composition of the libraries.

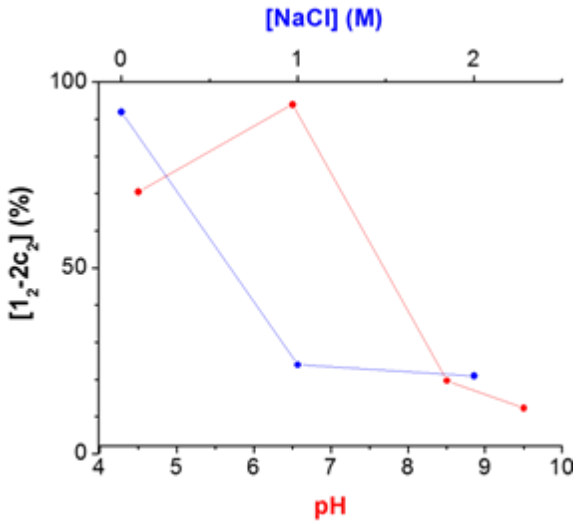


Figure 3. Abundance of **[1_2-$2a_2$]** expressed as percentage of the compound in the HPLC-UV traces (non-corrected by the specific absorbance of each component). Red: effect of pH. Blue: effect of ionic strength.

Structural characterization of the predominant species and study of the factors that caused its emergence. The high selectivity observed for the formation of the heterotetramer between **1** and **2a** is quite intriguing and, accordingly, we attempted to determine the exact structure and the topology of the self-selected species. For **[1_2-$2a_2$]** there are two possible symmetric (see below) constitutional isomers that can be described as a bis(macrocyclic) structure **Ia** (Figure 4A) and a macrobicyclic cage-like structure **Ib** (Figure 4B). In order to determine the connectivity between each counterpart, we synthesized compound **[1_2-$2a_2$]** on a preparative scale and, once isolated, we analyzed its structure by different 1D and 2D NMR experiments in d_6-DMSO (Figures S93-S97). The spectra showed a highly symmetric structure in solution (D_2 symmetry) and all the signals could be successfully assigned by the combination of the corresponding NMR experiments (Figure 4D). Unfortunately, at least in theory, the two possible regioisomers could show this D_2 average symmetry and therefore an unambiguous assignment was not possible at this point. Consequently, we looked for further indications on the nature of the product formed. We reasoned that the comparison of the ^{1}H-NMR spectrum with that of a simpler but related compound would bring some light into the discussion. To this aim, we analyzed the library formed by reacting **2c** and compound **3** which is a structural

analogue of tripodal unit **1a** but lacking one of the peptidic arms. In this occasion, the simple monomacrocycle **[2c-3]** arose as the main compound of the mixture, while evidences of larger species were not found (Figure S23); **[2c-3]** only forms one particular macrocycle with a defined size (Figure 4C). In order to compare the NMR spectra of **[1_2-$2c_2$]** and **[2c-3]** the latter was also synthesized on a preparative scale. The ^{1}H-NMR spectra of **[1_2-$2a_2$]** and **[2a-3]** showed significant differences; the ^{1}H-NMR spectrum of **[2a-3]** in d_6-DMSO presents a much larger number of signals with different intensities suggesting the presence of several conformations of the macrocyclic ring that interconvert slowly in the NMR timescale. Moreover, for each set of the ^{1}H NMR signals, the expected C_2 symmetry was broken in solution (Figure 4D). This observation clearly contrasted with the solution behavior of **[1_2-$2a_2$]** that showed a full D_2 symmetry in solution. The putative bis(macrocyclic) heterotetramer **Ia** can be described as two **[2c-3]** macrocycles connected by a cystine bridge, and thus a similar conformational behavior would be expected for **Ia** and **[2c-3]** macrocycles. Considering this evidence and our previous experience with similar macrocyclic disulfides,[23] we concluded that the observed isomer for **[1_2-$2a_2$]** corresponded to the macrobicyclic cage-like structure, where the larger ring size allows a more flexible structure in solution.

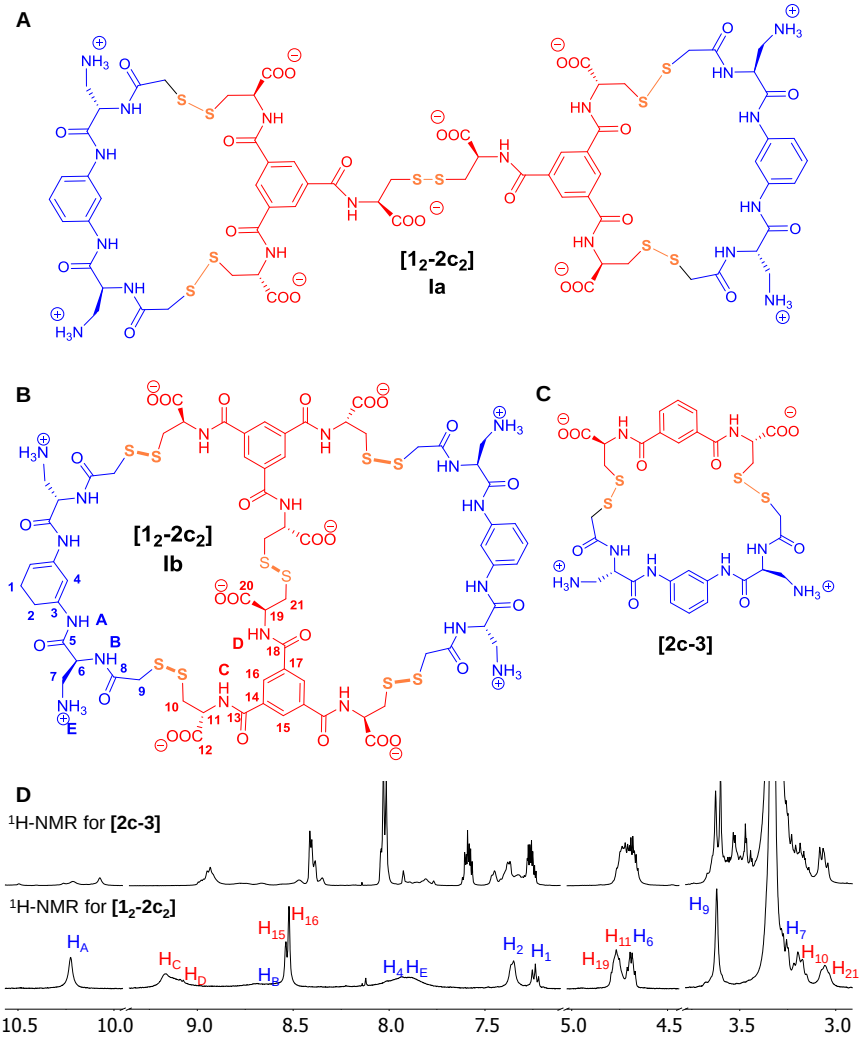


Figure 4. Structure for compound [**1_2-2a$_2$**], **Ia** (A) and **Ib** (B). C) Structure of compound [**2a-2f**], D) Comparative ^{1}NMR spectra (500MHz, 100% DMSO-d_6, 298K) showing the amide region for molecules [**1_2-2a$_2$**] and [**2a-3**].

With these experiments, we became confident about the nature of the **[1_2-2a$_2$]** compound giving a **Ib** structure and ruling out the bis(macrocyclic) topology **Ia**, which necessarily presents the same macrocyclic moiety as **[2a-3]**. In addition, we performed some further control experiments. MALDI-TOF analysis also confirmed the mass of the [2+2] compound demonstrating the formation of a heterotetramer. Moreover, MALDI-TOF MS/MS experiments were attempted to break the structure through collision-induced dissociation (CID) mode using argon as the collision gas. In these experiments we obtained different fragments of the structure (Figure S101) but we were unable to detect the fragment corresponding to **[1-2a]$^+$**, which would be expected by the symmetric rupture of structure **Ia**, or any compound resulting from a one-single bond cleavage. Instead, the loss of either a disulfide unit, CH_2SS or CO_2 were detected. Furthermore, the formation of **[1_2-2a$_2$]** was monitored over time (Figure S25-S26). At

short reactions times the formation of **[1-2a]** was observed but, remarkably, it took longer than other libraries previously reported in our group to reach a thermodynamic equilibrium.[19b] This data suggests that the formation of the main species is fairly complex and, again, this fact supports the formation of **Ib** in detriment of **Ia**, which would be formed rather quickly.

To gain a deeper knowledge about the structure of the hetero tetramer, we also performed molecular modeling studies using the Macromodel software (Molecular mechanics calculations with the OPLS3 force field in water). Therefore, several systematic Monte Carlo conformational searches were undertaken, each one starting from different initial conformations of the two possible symmetric isomers of the [2+2] structure (**Ia** and **Ib**). The lowest energy conformers for both isomers are shown in Figure 5. In both cases, the conformations of the systems are ruled by the polar interactions between the charged residues. Thus, in the two minima, the six carboxylate anions are implicated in salt bridges and the structures show a large number of polar H-bonds (14 in each geometry). These produce quite compact and folded structures. However, the main differences are due to the geometry of the aromatic bis(amide) moieties from the bipodal BBs (highlighted with green C-atoms in Figure 5). In **Ia**, the smaller ring size imposes a more strained geometry that precludes the co-planarity of the aromatic bis(amide) fragments, which is reflected in the syn/anti relative disposition of the amide bonds in both moieties. On the contrary, the same fragments in the **Ib** isomer are co-planar to the aromatic ring and with a symmetric syn/syn conformation, pointing their N-H bonds to the inner part of the macrobicycle ring. This is in agreement with the symmetry observed in the NMR spectra of **[1_2-$2a_2$]**, where a single signal was observed for these four aromatic amide NH protons. Moreover, the **Ib** cage-like isomer has a lower energy than the bis-macrocycle **Ia** (by 13.6 kJ/mol from OPLS3 Force Field). Despite this energy difference could be overestimating the strain effects in a quite flexible structure, the obtained results additionally support the spontaneous formation of the **Ib** isomer over the **Ia** counterpart, since the dynamic mixture of disulfides operates under thermodynamic control leading to the most stable situation.

The molecular model for the **Ib** cage also gives some clues about the reasons for the selectivity toward this macrobicycle. The carboxylate anions from the tripodal BBs are tightly H-bound to the amide and ammonium groups of the bipodal BBs. In most of the simulations, these interactions show a chelate-like disposition of the CH_2NH_3+ side chains of the **2c** moiety in a cyclic H-bond geometry. This binding should be highly favored for the short residues, which decrease the conformational freedom of the cationic side chains. We expect that the increase of the conformational freedom of the cationic residues would disfavor this arrangement, as it will be shown in the next section.

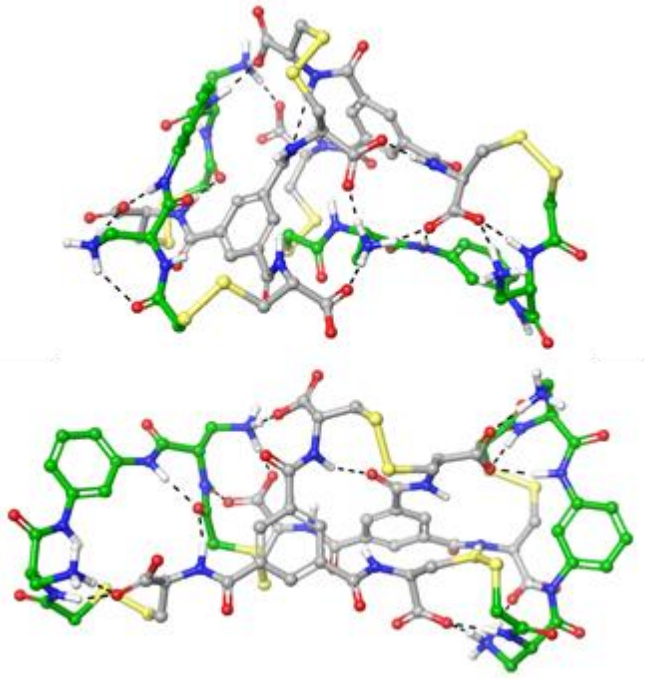


Figure 5. Energy minima for **Ia** (up) and **Ib** (down) heterotetramer, obtained after conformational searches (Macromodel, OPLS3 force field in water). For clarity, non-polar hydrogen atoms are not shown, while the carbon atoms from the bipodal BB moiety (**2a**) are coloured in green. The H-bonds are shown with black dashed lines.

Effect of side-chain length: in the previous section we reasoned that the selection mechanism is based on structurally well-defined polar contacts (salt bridges and H-bonds) between the pseudopeptide side-chain functionalities. We decided to investigate how tolerant is the selection process towards side-chain length changes. To this aim, several homologous amino compounds with different chain length (2 or 3 methylene units, corresponding to amino acids Dab (**2d**) and Orn (**2e**)) were prepared and investigated in dynamic libraries with the trimesic acid-cysteine derivative **1**. Interestingly, when **1** (0.5 mM) was mixed with either compounds compound **2d** or **2e** (0.5 mM) in the usual conditions no predominance towards the [2+2] tetramer was observed for these homologous, positively charged compounds (Figure S20-S21). More strikingly when **1** (0.5 mM) was reacted with a mixture of **2c, 2d**, and **2e** (0.5 mM each) in a competition experiment under the usual conditions (aqueous solution at pH 6.5, 25% DMSO) compound **[1_2-$2c_2$]** was formed in a way that consumed all **1** and **2c** building blocks (Figure 6A). Then, the remaining bipodal components formed the corresponding homo and heterodimers (as revealed by HPLC-MS analysis). Thus, for these structures the formation of **[1_2-$2c_2$]** is exclusive even in the presence of other competing charged species as part of the library. Therefore, the chain length plays a decisive role for the stabilization of a major compound through supramolecular interactions and homologous compounds differing as little

as one methylene unit behave in a divergent manner. The delicate structural complementarity was further shown by an analogous experiment using bipodal compound **3** instead of the tripodal unit **1**. In this case, when a mixture containing an equimolar quantities of the homologous amino components **2c**, **2d** and **2e** and the bipodal carboxylate **3** was tested no selectivity towards any of the amino derivatives was observed (Figure 6B). Accordingly, compound **3** was combined with the different positively charged bipodal units, obtaining as a result a wide range of products in the final distribution of the compounds. Once again, these results indicate the delicate structural complementarity necessary for the emergence of self-recognition processes. In particular, this experiment demonstrates the need of the third arm to selectively form **[1_2-$2c_2$]** and it is a further evidence for the **Ib** isomer as it forms a bigger macrocycle which brings the opposite charges in close proximity. This disposition would explain the dissimilar behavior for the other bipodal units where a single chain extension prevents establishing the adequate stabilizing interactions. These experimental results are in agreement with the computational calculations described in the previous section; the conformational freedom of the larger residues would preclude adopting the closed conformation necessary to establish these polar interactions.

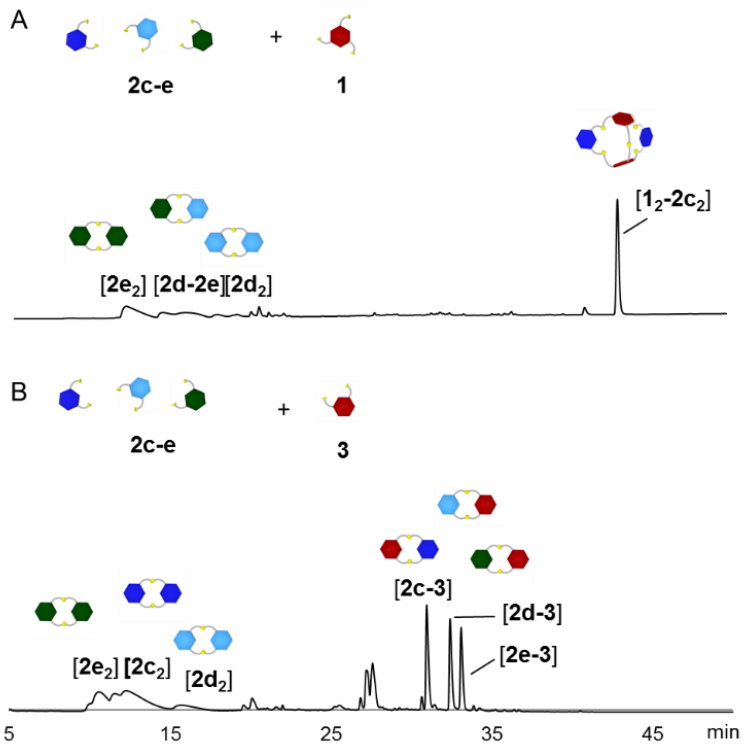


Figure 6. A) DCL formed by **1** and **2a,d,e** (0.5mM each) in aqueous solution (25% DMSO, pH 6.5), B) DCL formed by **2f** and **2a,d,e** (0.5mM each) in aqueous solution (25% DMSO, pH 6.5).

Conclusions

We have shown the intricate relationships that govern the product distributions in dynamic libraries of pseudopeptides. While a normal, statistical, product distribution consisting in homo and heterodimers is found for most of the libraries, in some cases the equilibrium is driven towards the formation of an unexpected heterotetramer. In order to form this discrete cage-like structure, stabilizing interactions must come in place but the effect of these non-covalent interactions is sufficient to change the size of the macrocycles formed from 26 to 42 members. The polar nature of these contacts is demonstrated by the reduction in area of the major product in the HPLC traces when the reactions were performed in different pH or in a higher ionic strength media. The discrimination was shown to be also very sensitive to the size of the building blocks being just an increment of a methylene unit detrimental for efficient formation. In addition, the removal of a covalent link between two of the components also prevents the formation of the tetrameric species. The formation of the major compound is therefore, directed by the nature of all counterparts. The stability of the newly formed compound has been shown by different competition experiments with charged monothiols and by MS-MS spectrometry. The structural importance of intramolecular interactions in aqueous media is shown in many biological processes but is especially relevant in the folding of proteins where the system adopts a tridimentional structure that is not only maintained by covalent bonds but by other non-covalent interactions. Accordingly, the understanding of simpler systems can help to explain the behavior of such complex structures. In this context, Dynamic Covalent Chemistry constitutes a powerful tool to discover non-trivial recognition processes. Knowing the rules that guide self-assembly will allow us to design more complex systems and predict their behavior opening the door to the synthesis of molecular devices with a tailored function.

Experimental Section

Synthesis of compound **1**,[21] **3**,[22] and **2a**, **2b**, **2e**[12] have been previously reported. Synthetic details for the preparation of **2c**, **2d** can be found in the ESI.

Reversed-Phase High-Performance Liquid Chromatography (RP-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. (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelength was set at 254 nm.

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).

High Resolution Mass Spectrometry (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 H2O + 20 mM HCOOH) mixtures at 0.3 mL/min were used as mobile phase.

General procedure for the preparation and HPLC analysis of the DCLs

A 66.7 mM BIS-Tris methane buffer was prepared by dissolving 1.39 g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq). The reaction mixtures were prepared by dilution of individual stocks of the building blocks (BBs) **1**, **2a-e** and **3** in DMSO. The concentration of the stocks of free thiols was adjusted by a calibration using Ellman's reagent (see ESI for more details). Unless otherwise specified, the DCLs were prepared at 0.5 mM of the BBs in a 50 mM BIS-Tris methane buffer (pH 6.5) with 25% DMSO. The mixtures were analysed by means of HPLC or UPLC-MS at different reaction times. Complete oxidation is achieved after 24h h. The reactions were then analysed by preparing HPLC samples adding 45 µL of the corresponding reaction mixture to 75 µL of a solution of 89% H_2O, 10% CH_3CN and 1% TFA. Eluent used: 2 min at 10% CH_3CN in H_2O, then linear gradient from 10% to 50% CH_3CN over 58 min

Acknowledgements

Financial Support from MINECO/FEDER (CTQ2015-70117-R and BES-2013-063128), AGAUR (2014 SGR 231) and European Union (COST CM1304) are gratefully acknowledged

Keywords: systems chemistry • dynamic networks • pseudopeptides • self-assembly • noncovalent interactions

[1] a) Y.-D. Wu and S. Gellman, *Acc. Chem. Res*., **2008**, *41*, 1231-1232; b) I. Alfonso, *Chem. Commun*., **2016**, 52, 239-250.
[2] a) E. Faggi, A. Moure, M. Bolte, C. Vicent, S. V. Luis, I. Alfonso. *J. Org. Chem*., **2014**, *79*, 4590-4601; b) S. V. Luis, I. Alfonso, *Acc Chem. Res*., **2014**, *47*, 112-124; c) E. Faggi, C. Vicent, S. V. Luis, I. Alfonso, *Org. Biomol. Chem*., **2015**, *13*, 11721-11731; d) I. Alfonso, Chem. Commun., 2016, 52, 239-250; e) E. Faggi, Y. Pérez, S. V. Luis, I. Alfonso, *Chem. Commun.* **2016**, *52*, 8142-8145.
[3] a) P. Kokkonen, M. Rahnasto-Rilla, P. H. Kiviranta, T. Huhtiniemi, T. Laitinen, A. Poso, E. Jarho, M. Lahtela-Kakkonen, *ACS Med. Chem. Lett.*, **2012**, 3, 969-974; b) E. Zervoudi, E. Saridakis, J. R. Birtley, S. S. Seregin, E. Reeves, P. Kokkala, Y. A. Aldhamen, A. Amalfitano, I. M. Mavridis, E. James, D. Georgiadis E. Stratikos, *Proc. Natl. Acad. Sci.*, **2013**, *110*,19890-19895; c) N. Micale, K. Scarbaci, V. Troiano, R. Ettari, S. Grasso and M. Zappalà, *Med. Res. Rev*., **2014**, *34*, 1001-1069.
[4] a) M. Reches, E. Gazit, *Science*, **2003**, *300*, 625-627; b) L. Ziserman, H.-Y. Lee, S. R. Raghavan, A. Mor, D. Danino*, J. Am. Chem. Soc*., **2011**,*133*, 2511-2517 c) V. Haridas, M. B. Bijesh, A. Chandra, S. Sharmaa, A. Shandilyaa. *Chem*. *Commun*., **2014**, *50*, 13797-13800.
[5] I. Alfonso. ***Chem. Commun.***, *2016*, ***52***, 239-250.
6 G. C. Barrett and D. T. Elmore, *Amino Acids and Peptides*, Cambridge University Press, 1998
[7] a) J.M. Lehn, *Angew Chem. Int. Ed. Engl.* **1990**, *29*, 1304-1319; b) J.M. Lehn, *Chem. Soc. Rev.* **2007**, *36*, 151-160;
[8] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. W., J. K. M. Sanders, S. Otto, *Chem. Rev*. **2006**, *106*, 3652-3711.
[9] a) N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş, J. K. M. Sanders, *Science*, **2012**, *338*, 785-783. b) J. Solà, M. Lafuente, J. Atcher, I. Alfonso, *Chem. Commun.* **2014**, 50, 4564-4566; b) M.Lafuente, J. Atcher, J. Solà, I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 17002-17009.
[10] a) J. W. Sadownik, D. Philp, Angew. Chem. Int. Ed. 2008, 47, 9965 -9970; Angew.Chem. 2008, 120,10113 -10118; b) R. Nguyen, L. Allouche, E. Buhler, N. Giuseppone, *Angew. Chem. Int. Ed.* **2009**, *48*,1093-1096; *Angew. Chem*. **2009**, *121*,1113 -1116; c) J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans, S. Otto, *Science* **2010**, *327*,1502-506; d) M. Colomb-Delsuc, E. Mattia, J. W. Sadownik, S. Otto, *Nat. Commun*. **2015**, *6*, 7427; e) J. W. Sadownik, E. Mattia, P. Nowak, S. Otto, *Nat. Chem.*

2016, *8*, 264 -269; f) C.Chen, J. Tan, M.-C.Hsieh, T. Pan, J. T. Goodwin, A. K. Mehta, M. A. Grover,D.G.Lynn, *Nat. Chem*. **2017***, 9*,799 -804.
[11] a) T. Gánti, *The Principles of Life*, Oxford University Press, Oxford, **2003**; b) K. Ruiz-Mirazo, C. Briones, A. de la Escosura, *Chem. Rev.* **2014**, *114*, 285-266.
[12] a) J. Atcher, A. Moure, I. Alfonso, *Chem. Commun.* **2013**, *49*, 487-489. b) J. Atcher, A. Moure, J. Bujons, I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 6869-6878. c) J. Atcher, J. Solà, I. Alfonso. *Org. Biomol. Chem*., **2017**,*15*, 213-219. d) A. M. Valdivielso, F. Puig-Castellví, J. Atcher, J. Solà, R. Tauler, I. Alfonso. *Chem. Eur. J.* **2017**, *23*,10789 -10799.
[13] J. Atcher, I. Alfonso. *RSC Advances*, **2013**, *3*, 25605-25608.
[14] a) S. Otto, S. Kubik, *J. Am. Chem. Soc.* **2003**, *125*, 7804–7805 a) M. Rauschenberg, S. Bomke, U. Karst, B.J. Ravoo, *Angew. Chem. Int. Ed*. **2010**, *49*, 7340 –7345; *Angew. Chem*. **2010**, *122*, 7498 –7503;
[15] a) F. B. L. Cougnon, N. Ponnuswamy, N. A. Jenkins, G. D. Pantos, J. K. M. Sanders, *J. Am. Chem.Soc*. **2012**, *134*,19129 –19135; b) R. T. S. Lam, A. Belenguer, S. L. Roberts, C. Naumann, T. Jarrosson, S. Otto, J. K. M. Sanders, *Science* **2005**, *308*, 667–669; c) H. Y. Au-Yeung, G. D. Pantos, J. K. M. Sanders, *Proc. Natl. Acad.Sci. USA*, **2009**, *106*,10466 –10470.
[16] a) S. Otto, R. L. E. Furlan, J. K. M. Sanders, *Science*, **2002**, 297, 590-593 b) J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans, S. Otto. *Science*, **2010**, *327*, 1502-1506
[17] a) A. M. Whitney, S. Ladame, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2004**, *43*, 1143-1146; *Angew. Chem*. **2004**, 116, 1163-1166 b) K. R. West, K. D. Bake, S. Otto, *Org. Lett*. **2005**, *7*, 2615-2618; c) Y. Krishnan-Ghosh, A. M. Whitneya S. Balasubramanian, *Chem. Commun*, **2005**, 3068-3070. d) A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2008**, *47*, 2677-2680; *Angew. Chem*. **2008**, *120*, 2717-2720. e) W. Drożdż, C. Bouillon, C. Kotras, S. Richeter, M. Barboiu, S. Clément, A. R. Stefankiewicz, S. Ulrich, *Chem. Eur. J*., **2017**, *23*, 18010-18018
[18] a) J. Solà, M. Lafuente, J. Atcher, I. Alfonso, *Chem. Commun.* **2014**, 50, 4564-4566. b) M. Lafuente, J. Atcher, J. Solà, I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 17002-17009.
[19] M. Lafuente, J. Solà, I. Alfonso. *Angew. Chem Int. Ed. In press*. DOI/10.1002/anie.201802189
[20] S. Kassem, T. van Leeuwen, A. S. Lubbe, M. R. Wilson, B. L. Feringha, D. A. Leigh. Chem. Soc. Rev**.**, **2017**, *46*, 2592-2621
[21] K. R. West, K. D. Bake, S. Otto, *Org. Lett*. **2005**, *7*, 2615 -2618.
[22] T. M. Postma, W. R. J. D. Galloway, Warren, F. B. Cougnon, G. D. Pantos, J. E. Stokes, D. R.Spring. *Synlett*, **2013**, 24, 765-769.
[23] For a related macrocycle see J. Atcher, J. Bujons, I. Alfonso. *Chem. Commun*. **2017**, *53*, 4274-4277.

2.5.1. Electronic Supplementary Information (ESI) for:

PUBLICATION C

Structurally selective assembly of a specific macrobicycles from a dynamic library of pseudopeptidic disulfides

Maria Lafuente, Jordi Solà and Ignacio Alfonso**

To be submitted

Table of contents

GENERAL METHODS

General: Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka or Merck) and were used without further purification. Flash chromatographic purifications were performed using SiO_2 (Silica gel 60 A, 35-70 µ, Carlo Ebra) and preparative reversed-phase purifications were performed on a Biotage® Isolera Prime™ equipment using KP-C_{18}-12g HS (Biotage® SNAP Cartridge). TLCs were performed using 6x3 cm SiO_2 pre-coated aluminium plates (ALUGRAM® SIL G/UV_{254}).

Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC): Analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using:
For the analysis of building blocks **2c** and **2d**: a reversed-phase X-Terra C_{18} (15 x 0.46 cm, 5 µm) column. (CH_3CN + 0.07% TFA and H_2O + 0.1% TFA) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelengths were set at 220 nm.
For the analysis of the DCLs: a reversed-phase kromaphase C_{18} (25 x 0.46 cm, 5µm) column. (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelength was set at 254 nm.

Nuclear Magnetic Resonance (NMR): Spectroscopic experiments were carried out on a Varian INOVA 500 spectrometer (500 MHz for ^1H and 125 MHz for ^{13}C) and a Varian Mercury 400 instrument (400 MHz for ^1H and 101 MHz for ^{13}C). Chemical shifts (δ_H) are quoted in parts per million (ppm) and referenced to the appropriate NMR solvent peak(s). 2D-NMR experiments COSY, HSQC and HMBC were used where necessary in assigning NMR spectra. Spin-spin coupling constants (*J*) are reported in Hertz (Hz).

pH measurements were performed at room temperature on a Crison GLP21 pH-meter with the electrode Crison 50 14T.

High Resolution Mass Spectrometry (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.

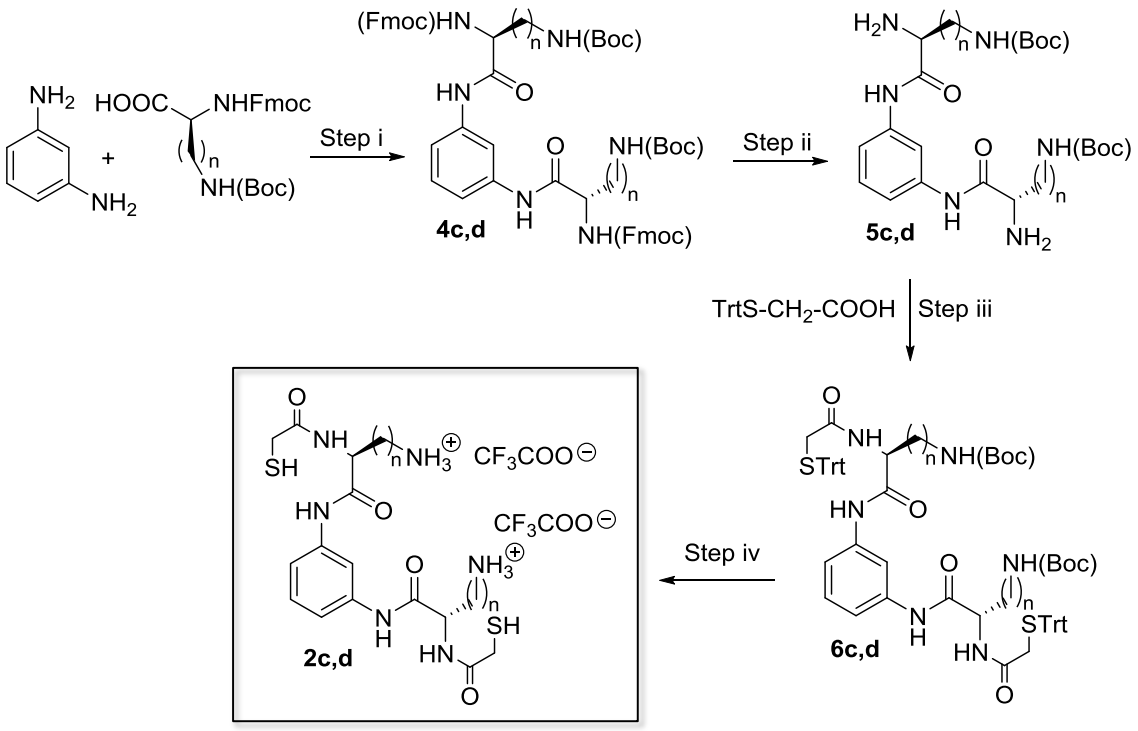
Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-TOF/TOF): A MALDI-TOF/TOF mass spectrometer (AutoFLEX III, Bruker Daltonics, Bremen, Germany) equipped with a pulsed N_2 laser (337 nm) controlled by the Flexcontrol 3.4 software package was used to obtain MS and MS/MS data. The accelerating voltage was 19 kV and the reflectron voltage 21 kV. Spectra are the sum of 200 shots with a frequency of 200 Hz. The MS/MS spectra were obtained in the collision-induced dissociation (CID) mode using argon as the collision gas.

Molecular Modelling: Molecular modeling was performed with the MacroModel package within the Maestro software. The corresponding minima were generated by a Monte Carlo conformational search using the tool available in MacroModel and minimization steps with the OPLS3 Force Field in implicit water. Different starting geometries were assayed for the two isomers and all the generated local minima were compared in energy. The minima in Figure 5 are those with the lowest energy for each isomer.

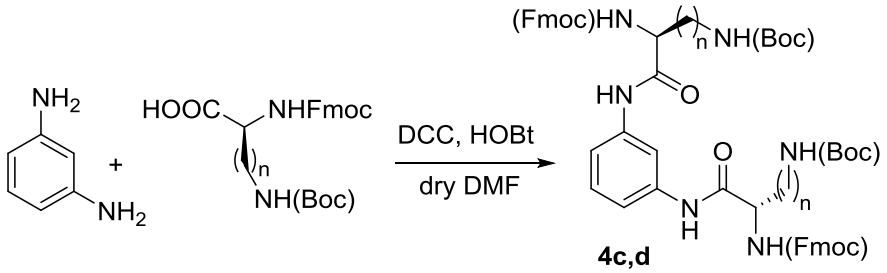
SYNTHESIS OF THE BUILDING BLOCKS

Trithiol **1**[1] and dithiol **3**[2] were synthesized as previously reported. Compounds **2a-b, 2e** were described previously by our group;[3] tritylsulfanylacetic acid was prepared as previously reported.[4]

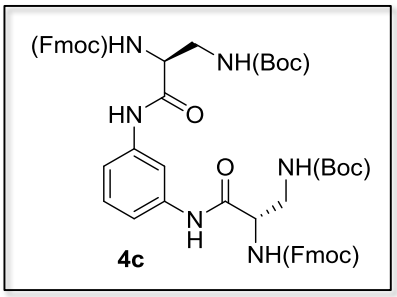
Synthetic scheme for the preparation of 2c, 2d



Step i: Experimental procedure for the synthesis 4c, 4d

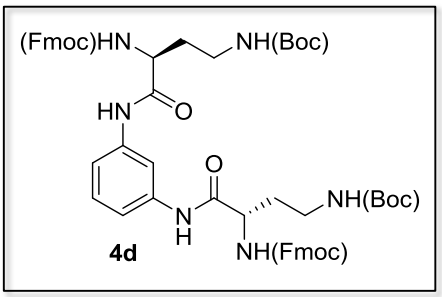


[1] K. R. West, K. D. Bake and S. Otto, *Org. Lett.*, **2005**, *7*, 2615-2618.
[2] T. M. Postma, W. Galloway, F. B. L. Cougnon, G. D. Pantos, J. E. Stokes and D. R. Spring, *Synlett*, **2013**, *24*, 765−769.
[3] J. Atcher, J. Solà and I. Alfonso, *Org. Biomol. Chem.*, **2017**, *15*, 213-219.
[4] A. P. Kozikowski, Y. Chen, A. Gaysin, B. Chen, M. A. D'Annibale, C. M. Suto and B. C. Langley, *J. Med. Chem.*, **2007**, *50*, 3054-3061.



4c: Fmoc-DAP(Boc)-OH (1.75 g, 4.07 mmol) was dissolved in dry DMF (10 mL); dicyclohexylcarbodiimide (DCC, 915 mg, 4.41 mmol) and 1-hydroxybenzotriazole (HOBt, 660 mg, 4.88 mmol) were added over the solution. The reaction mixture was stirred during 2 min and then *m*-phenylenediamine (200 mg, 1.84 mmol) was added over the mixture. The reaction mixture was stirred at room temperature under an inert atmosphere of N_2 for 48 hours, after which complete conversion of the starting material was observed by TLC. The mixture was filtered, and the filtrate was diluted with dichloromethane (40 mL), washed with saturated aqueous $NaHCO_3$ (3 x 10 mL) and brine (10 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography (SiO_2) using $CHCl_2$: EtOH as eluent (from 1% to 2.5% EtOH) to give 800 mg of **4c** (52% yield) as a white solid.

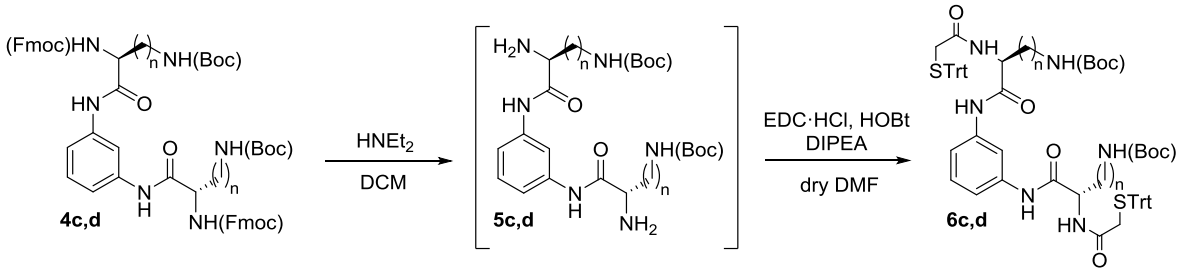
Rf (SiO_2, $CHCl_2$/ EtOH (2% EtOH))= 0.32.
HRMS (ESI$^+$) calcd. for $C_{52}H_{56}N_6O_{10}$ $[M-H]^+$ (m/z): 925.0520, found: 925.2185.
^{1}H NMR (400 MHz, $CDCl_3$-*d*: CD_3OD, 9:1): δ = 9.51 (s, 2H, CON<u>H</u>), 7.74 (s, 1H, CH_{Ar}), 7.63 (d, *J* = 7.6 Hz, 3H, CH_{Ar}), 7.50 (t, *J* = 6.2 Hz, 4H, CH_{Ar}), 7.34 – 7.07 (m, 12H, CH_{Ar}), 4.28 (m, 6H, 4H $NHCO_2C\underline{H}_2$ + 2H $NHCO_2CH_2C\underline{H}$), 4.09 (t, *J* = 7.0 Hz, 2H, C*<u>H</u>CH_2), 3.16-3.46 (m, 4H, C*HC<u>H</u>$_2$), 1.29 (s, 18H, CH_3) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ= 168.8 (2 x CO), 143.6 (4 x CO), 141.2 (4 x C_{Ar}), 138.1 (4 x C_{Ar}), 129.2 (2 x C_{Ar}), 127.6 (6 x C_{Ar}), 127.0 (4 x C_{Ar}), 124.9 (2 x C_{Ar}), 119.8 (6 x C_{Ar}), 116.1 (2 x C_{Ar}), 80.1 (2 x <u>C</u>$(CH_3)_3$), 67.1 (2 x COO<u>C</u>H_2), 55.9 (2 x $COOCH_2$<u>C</u>H), 47.0 (2 x <u>C</u>*HCH_2), 42.5 (2 x C*H<u>C</u>H_2), 28.1 (6 x <u>C</u>H_3) ppm.

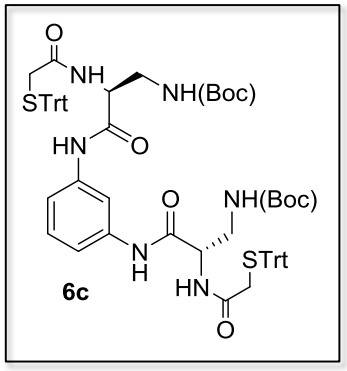


4d: Fmoc-DAB(Boc)-OH (515 mg, 1.17 mmol) was dissolved in dry DMF (3 mL), dicyclohexylcarbodiimide (DCC, 311 mg, 1.57 mmol) and 1-hydroxybenzotriazole (HOBt, 232 mg, 1.57 mmol) were added over the solution. The reaction mixture was stirred during 2 min and then *m*-phenylenediamine (57 mg, 0.53 mmol) was added over the mixture. The reaction mixture was stirred at room temperature under an inert atmosphere of N_2 for 48 hours, after which complete conversion of the starting material was observed by TLC. The mixture was filtered, and the filtrate was diluted with dichloromethane (15mL), washed with saturated aqueous $NaHCO_3$ (3 x 10mL) and brine (10 mL), dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The residue was purified by flash column chromatography using Hexane: EtOAc as eluent (from 40% to 60% EtOAc) to give 0.435 g of **4d** (86% yield) as a white solid.

Rf (SiO_2, EtOAc/hexane, 3:7)= 0.29.
HRMS (ESI$^-$) calcd. for $C_{54}H_{60}N_6O_{10}$ $[M+H]^+$ (m/z): 953.1060, found: 953.4489.
^{1}H NMR (400 MHz, $CDCl_3$-*d*): δ = 9.71 (s, 2H, CONH), 7.76 (d, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.61 (d, *J* = 7.5 Hz, 4H, CH_{Ar}), 7.39 (t, *J* = 7.4 Hz, 4H, CH_{Ar}), 7.30 (t, *J* = 7.5 Hz, 8H, CH_{Ar}), 6.07 (s, 2H, C*HNH), 5.09 (s, 2H, NH(COOtBu)), 4.38 (m, 6H, 4H $NHCO_2CH_2$ + 2H $NHCO_2CH_2CH$), 4.22 (t, *J* = 7.2 Hz, 2H, C*HCH$_2$), 2.99-3.59 (m, 4H, C*HCH$_2$CH$_2$), 1.96 (m, 4H, C*HCH$_2$), (1.43 (s, 18H, CH_3) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ= 169.5 (2 x CO), 157.8 (4 x CO), 143.9 (2 x C_{Ar}), 141.4 (4 x C_{Ar}), 138.6 (4 x C_{Ar}), 129.5 (1 x C_{Ar}), 127.9 (4 x C_{Ar}), 127.2 (4 x C_{Ar}), 125.3 (4 x C_{Ar}), 120.1 (4 x C_{Ar}), 115.9 (2 x C_{Ar}), 111.7 (1 x C_{Ar}), 80.6 (2 x $C(CH_3)_3$), 67.2 (2 x $COOCH_2$), 52.7 (2 x $COOCH_2CH$), 47.3 (2 x C*H), 37.2 (2 x C*HCH$_2$CH$_2$), 35.3 (2 x C*HCH$_2$), 28.5 (6 x CH_3) ppm.

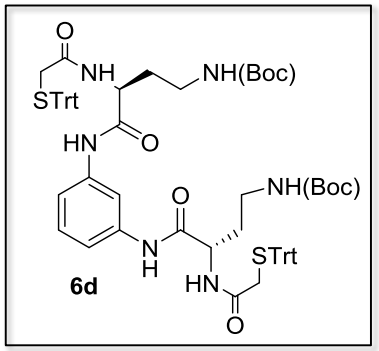
Step ii + Step iii: Experimental procedure for the synthesis 6c, 6d





6c: **4c** (500 mg, 0.54 mmol) was dissolved in a mixture of dichloromethane (5 mL) and diethylamine (2.5 mL). The solution mixture was stirred at room temperature for 5 hours until no remaining starting material was observed by TLC. The mixture was then diluted with $CHCl_3$ (10 mL) and the solvent removed under reduced pressure; this procedure was repeated three times to ensure complete removal of diethylamine excess. The remaining crude of **5c** was used without further purification. The crude product was dissolved in DMF (6mL), trityl-protected tritylsulfanylacetic acid (400 mg, 1.20 mmol), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC·HCl, 345 mg, 1.8 mmol), HOBt (210 mg, 1.56 mmol) and *N*,*N*diisopropylethylamine (DIPEA, 0.42 mL, 2.4 mmol) were then added over the solution. The reaction was stirred overnight at room temperature under a nitrogen atmosphere. After 18 h the mixture was diluted with AcOEt (40 mL), washed with saturated aqueous $NaHCO_3$ (3x10 mL) and dried over Na_2SO_4, filtered and concentrated under reduced pressure. The residue was purified by flash column chromatography using hexane: AcOEt as eluent (from 50% to 80% AcOEt) to give 430 mg of **6c** (73% yield over two steps) as a white solid.

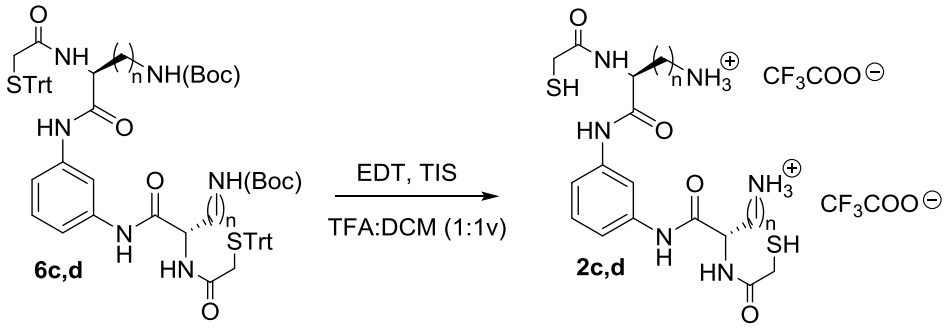
Rf (SiO_2**,** EtOAc/hexane, 1:1)= 0.39.
HRMS (ESI$^-$) calcd. for $C_{64}H_{68}N_6O_8S_2$ $[M+H]^+$ (m/z): 1113.4020, found: 1113.4558.
^{1}H NMR (400 MHz, $CDCl_3$-*d*): δ = 9.11 (s, 2H, CONHC*H), 8.00 (s, 2H, CONHCH_2S(Trt)), 7.41 (dd, *J* = 7.5, 1.7 Hz, 12H, CH_{Ar}), 7.25 (m, 14H, CH_{Ar}), 7.18 – 7.07 (m, 8H, CH_{Ar}), 5.33 (s, 2H, CH_2NH), 4.23 (d, *J* = 7.6, 4.1 Hz, 2H, C*HCH_2), 3.56-3.30 (m, 4H, C*HCH_2), 3.34 (m, 4H, CH_2S(Trt)), 1.35 (s, 18H, CH_3) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ= 170.2 (2 x CO), 168.0 (2 x CO), 157.1 (2 x CO), 143.9 (6 x C_{Ar}), 138.5 (2 x C_{Ar}), 129.6 (12 x C_{Ar}), 129.4 (1 x C_{Ar}), 128.3 (12 x C_{Ar}), 127.2 (6 x C_{Ar}), 115.8 (2 x C_{Ar}), 111.3 (1 x C_{Ar}), 80.3 (2 x C$(CH_3)_3$), 68.2 (2 x C$(C_6H_5)_3$), 55.3 (2 x C*HCH_2), 41.6 (2 x C*HCH_2), 35.9 (2 x CH_2S(Trt)), 28.4 (6 x CH_3) ppm.

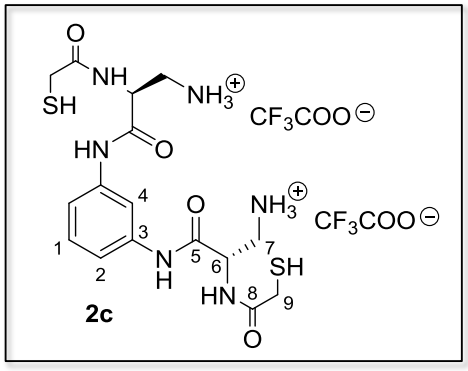


6d: **4d** (435 mg, 0.47 mmol) was dissolved in dichloromethane (2.0 mL) and diethylamine (2.0 mL). The solution mixture was stirred at room temperature for 4 hours until no starting material was observed by TLC. The mixture was diluted with $CHCl_3$ (10 mL) and the solvent removed under reduced pressure; this procedure was repeated three times to ensure complete removal of diethylamine excess. The remaining crude containing **5d** was used without further purification. DMF (4 mL) was added followed by trityl-protected tritylsulfanylacetic acid (160 mg, 0.47 mmol), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC·HCl, 120 mg, 0.61 mmol), HOBt (110 mg, 0.81 mmol) and *N*,*N*diisopropylethylamine (DIPEA, 0.18 mL, 1.82 mmol). The solution was stirred at room temperature under an inert atmosphere of N_2 for 18 hours. The mixture was diluted with AcOEt (10 mL), washed with saturated aqueous $NaHCO_3$ (3x5 mL) and dried over Na_2SO_4, filtered and concentrared under reduced pressure. The residue was purified by flash chromatography using hexane: AcOEt as eluent (from 30% to 40% AcOEt) to give 0.103 g of **6d** (73% yield over two steps) as a white solid.

Rf (SiO_2; EtOAc/hexane, 1:1)= 0.48.
HRMS (ESI$^-$) calcd. for $C_{66}H_{72}N_6O_8S_2$ $[M+H]^+$ (m/z): 1141.4560, found: 1141.4823.
^{1}H NMR (400 MHz, $CDCl_3$-*d*): δ = 9.75 (s, 2H, CONHC*H), 8.10 (s, 1H, CH_{Ar}), 7.42 (d, *J* = 7.8 Hz, 12H, CH_{Ar}), 7.26 (m, 15H, CH_{Ar}), 7.18 (m, 6H, CH_{Ar}), 7.02 (s, 2H, C*HNH), 5.13 (s, 2H, CH_2CH_2NH), 4.22 (m, 2H, C*HCH_2), 3.09 (m, 4H, CH_2S), 2.80-3.50 (m, 4H, C*HCH_2CH_2), 1.71 (m, 4H, C*HCH_2), 1.44 (s, 18H, CH_3) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ = 168.8 (2 x CO), 168.1 (2 x CO), 157.8 (2 x CO), 144.1 (6 x C_{Ar}), 138.7 (2 x C_{Ar}), 129.7 (12 x C_{Ar}), 129.5 (1 x C_{Ar}), 128.4 (12 x C_{Ar}), 128.3 (1 x C_{Ar}), 127.1 (6 x C_{Ar}), 115.7 (1 x C_{Ar}), 111.2 (1 x C_{Ar}), 80.4 (2 x C$(CH_3)_3$), 68.0 (2 x C$(C_6H_5)_3$), 51.3 (2 x C*H), 37.1 (2 x C*HCH_2CH_2), 36.2 (2 x CH_2S), 34.9 (2 x C*HCH_2), 28.5 (6 x CH_3) ppm.

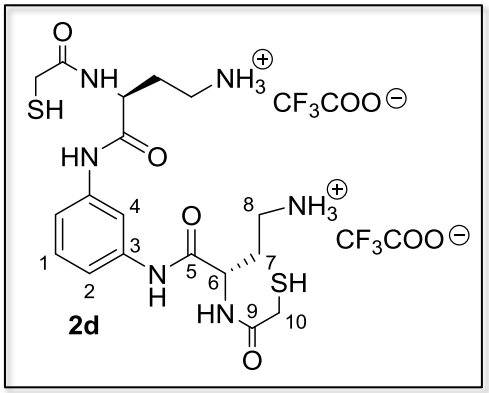
Step iv: Experimental procedure for the synthesis 2c, 2d





2c: **6c** (200 mg, 0.18 mmol) was dissolved in dichloromethane (1 mL) and trifluoroacetic acid (1 mL), triisobutylsilane (560 µL, 2.16 mmol) and of 1,2-ethanedithiol (271 µL, 3.24 mmol) were added rapidly and unde stirring. The reaction mixture was stirred at room temperature for 2 hours after which the reaction solvents were partially removed using a N_2 flow. The product was precipitated by the addition of diethyl ether (15 mL), filtered and washed with diethyl ether. The product was purified using reversed-phase flash chromatography (C_{18}, gradient: from 1% to 12% CH_3CN in H_2O). 64.8 mg of **2c** (55% yield) were obtained as a white solid.

HRMS (ESI$^-$) calcd. for $C_{16}H_{24}N_6O_4S_2$ $[M+H]^+$ (m/z): 429.1378, found: 429.1368.
^{1}H NMR (400 MHz, CD_3OD-d_4): δ = 8.01 (t, J = 2.2 Hz, 1H, H_4), 7.36 (t, J = 2.1 Hz, 1H, H_1), 7.31 (m, 2H, H_2), 4.83 (dd, J = 7.3, 5.4 Hz, 2H, H_6), 3.31 (s, 4H, H_9, overlapped with solvent residual peak), 3.21-3.49 (dd, J = 13.2, 7.3 Hz, 4H, H_7).
^{13}C NMR (101 MHz, CD_3OD-d_4): δ = 174.1 (2 x C_8), 168.5 (2 x C_5), 139.6 (2 x C_3), 130.2 (1 x C_1), 118.0 (2 x C_2), 114.0 (1 x C_4), 52.8 (2 x C_6), 41.7 (2 x C_7), 28.3 (2 x C_9).



2d: **6d** (87 mg, 0.076 mmol) was dissolved in dichloromethane (1 mL) and trifluoroacetic acid (1 mL) was added followed by triisobutylsilane (187 µL, 0.91 mmoL) and 1,2-ethanedithiol (115 µL, 1.37 mmol) 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 precipitate was filtered off and washed with diethyl ether. The product was purified using reversed-phase flash chromatography (C18 column, gradient: from 1% to 12% CH_3CN in H_2O) and 6 mg of **2d** (11% yield) were obtained as a white solid.

HRMS (ESI$^-$) calcd. for $C_{18}H_{28}N_6O_4S_2$ $[M+H]^+$ (m/z): 457.1692, found: 457.1682.

Product **2d** was found to aggregate and rapidly oxidize in solution so it could not be characterized by NMR methods. Instead, its oxidized form [**2d$_2$**] was prepared and fully characteritzed:
^{1}H NMR (400 MHz, CD_3OD-d_4): δ = 8.04 (m, 2H, H_4), 7.19 (m, 6H, H_1, H_2), 4.67 (m, 4H, H_6), 3.65 (m, 8H, H_{10}), 3.14 (m, 8H, H_8), 2.19(dd, 8H, H_7) ppm.
^{13}C NMR (101 MHz, CD_3OD-d_4): δ = 172.1 (4 x C_5), 170.7 (4 x C_9),139.38 (4 x C_3), 130.1 (2 x C_1), 117.7 (4x C_2), 114.0 (2 x C_4), 53.3 (4 x C_6), 43.5 (4 x C_{10}), 37.7 (4 x C_8), 30.9 (4 x C_7) ppm.

NMR spectra, HRMS (ESI+) spectrum and HPLC trace of 2c, 2d

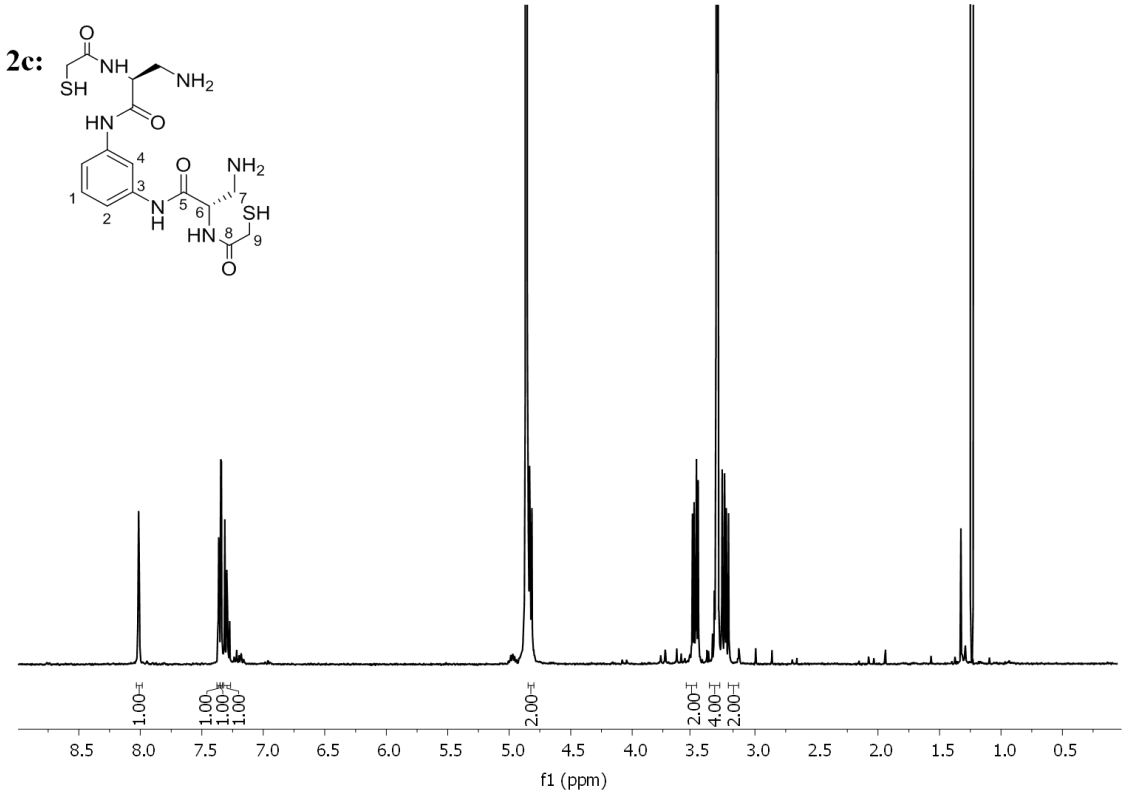


Figure S1. ^{1}H (400 MHz, 298 K in CD_3OD-d_4) spectrum of **2c**.

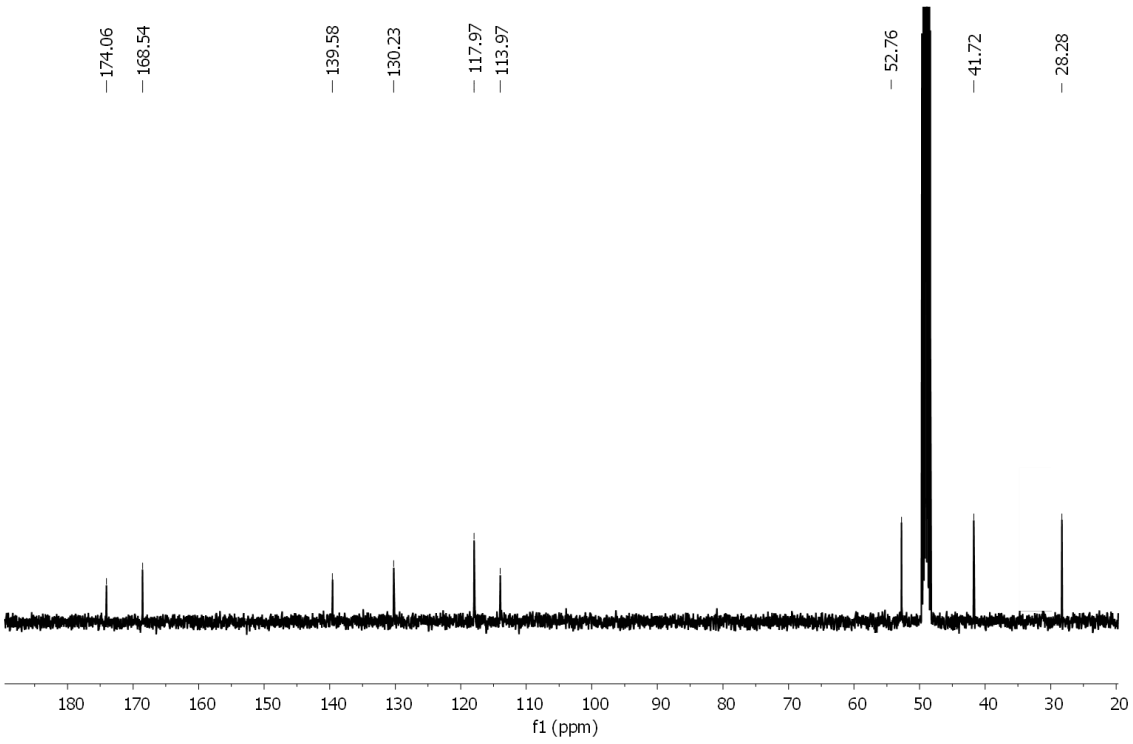


Figure S2. ^{13}C (101 MHz, 298 K in CD_3OD-d_4) spectrum of **2c**.

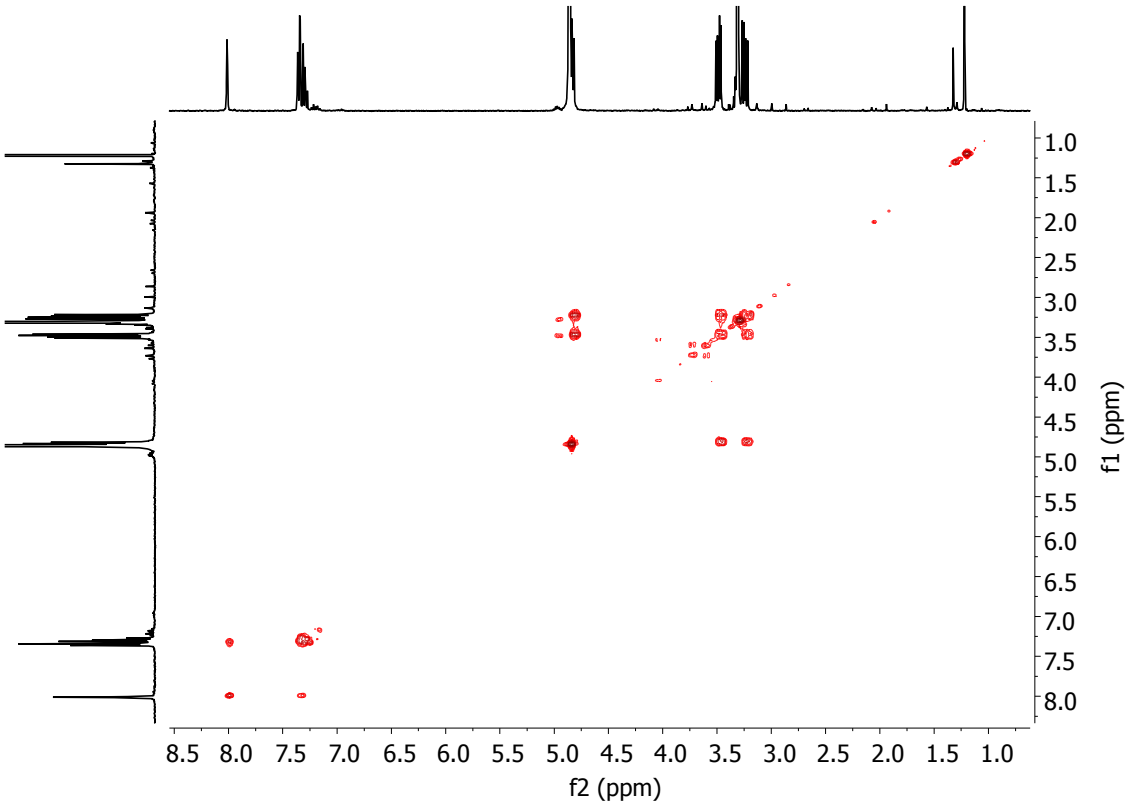


Figure S3. ^{1}H-^{1}H gCOSY (400 MHz, 298 K in CD_3OD-d_4) spectrum of **2c**.

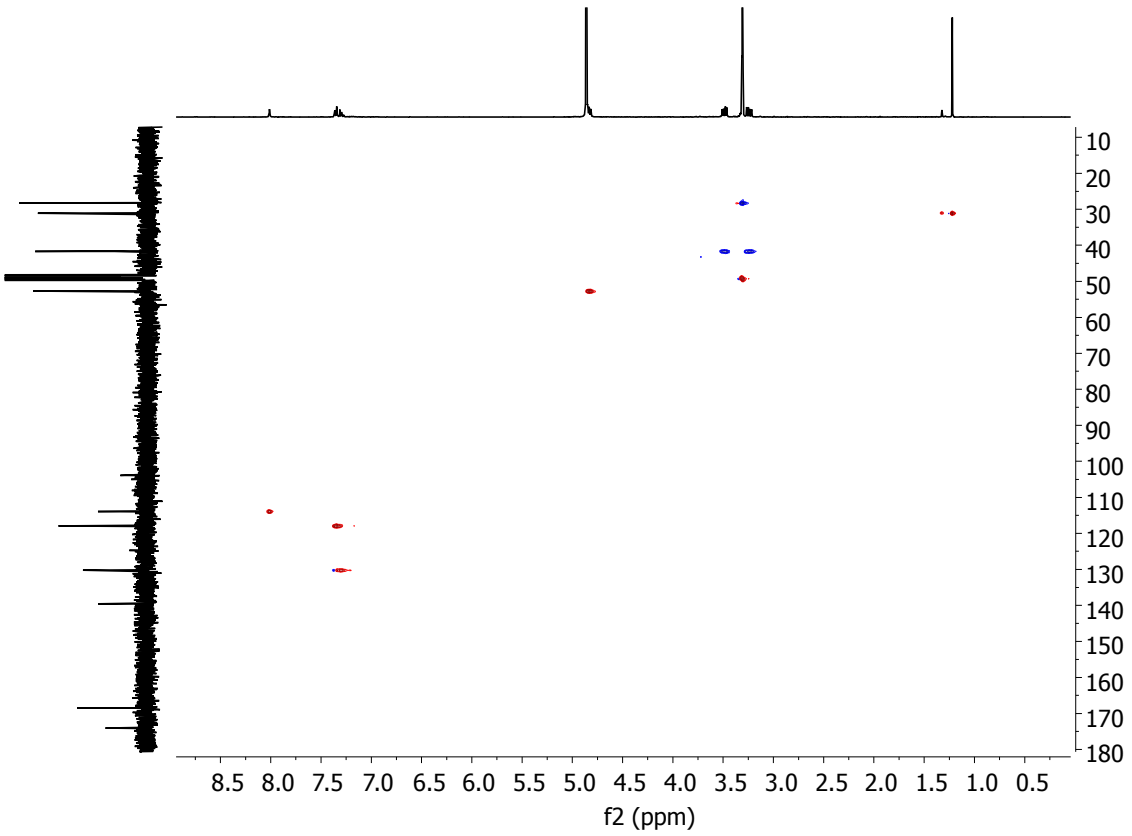


Figure S4. ^{1}H-^{13}C gHSQC (400 MHz, 298 K in CD_3OD-d_4) spectrum of **2c**.

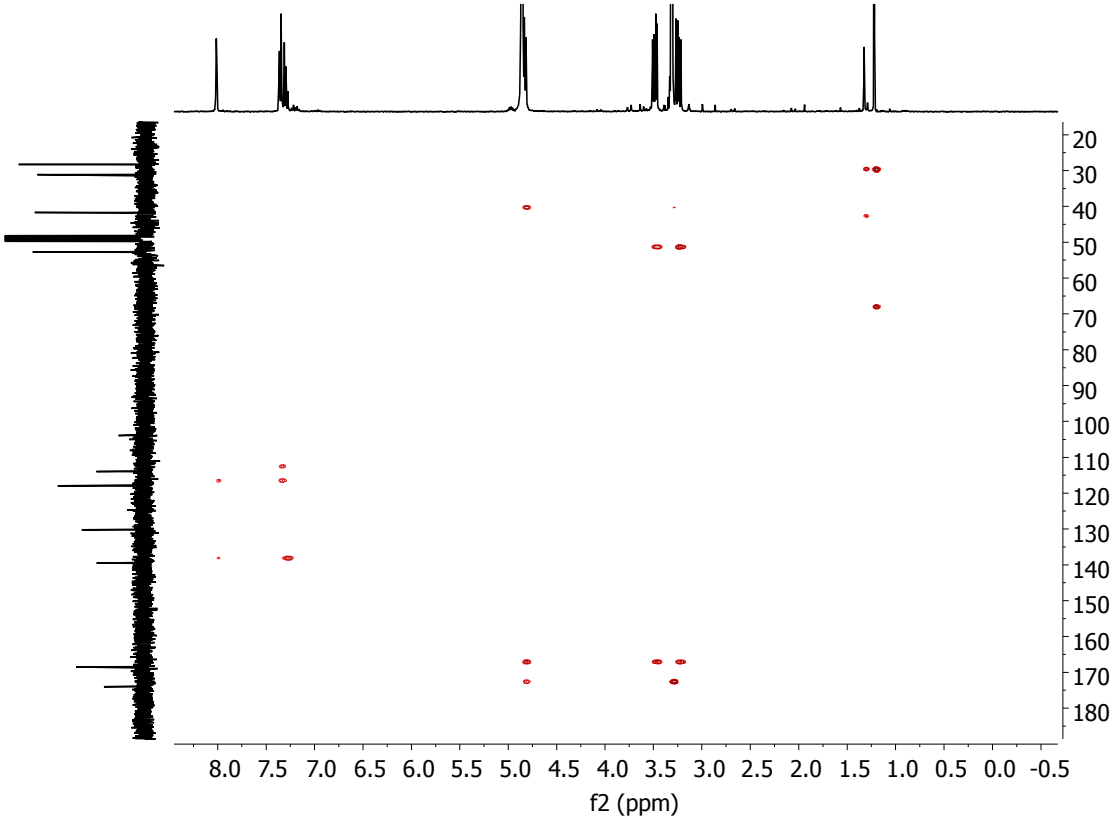


Figure S5. ^{1}H-^{13}C gHMBC (400 MHz, 298 K in CD_3OD-d_4) spectrum of **2c**.

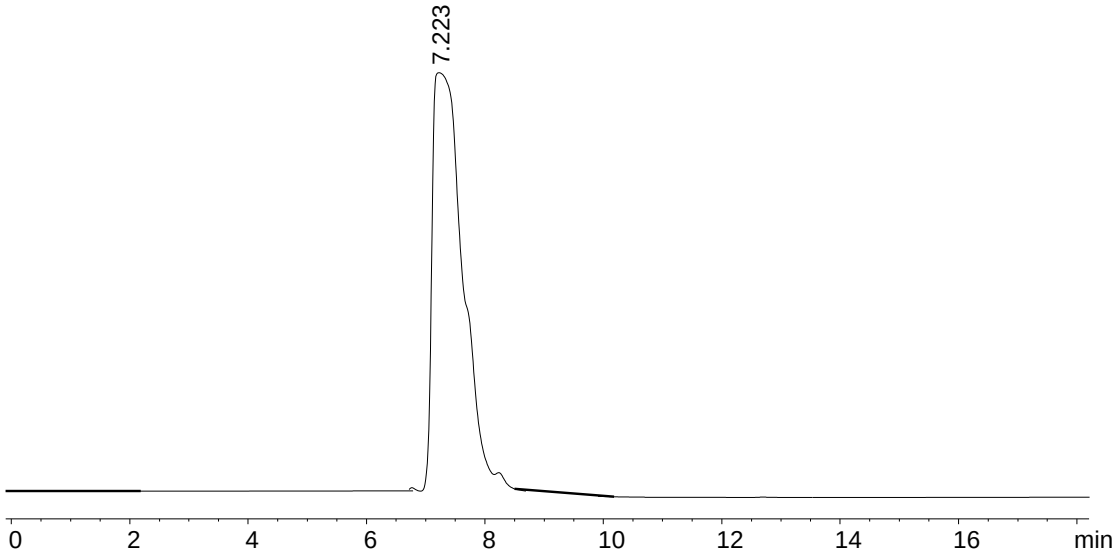


Figure S6: HPLC traces of **2c** (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

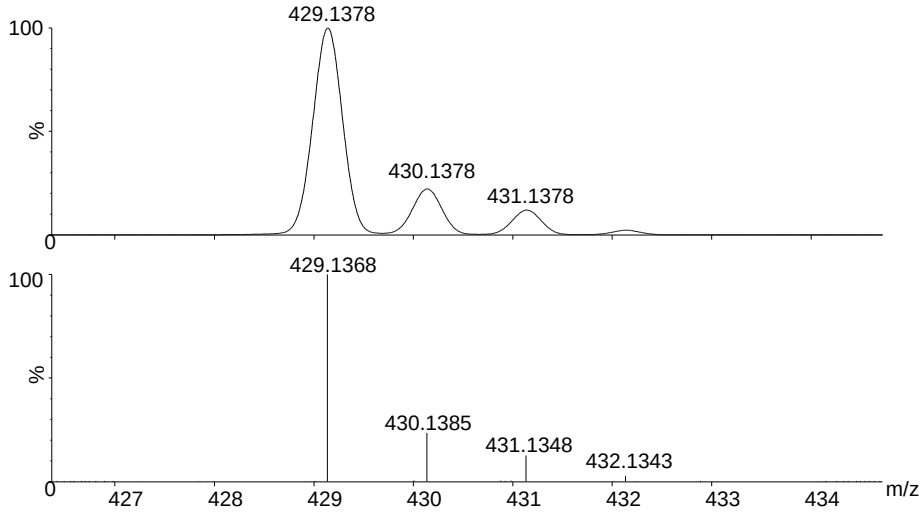


Figure S7: Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **2c**.

2d:

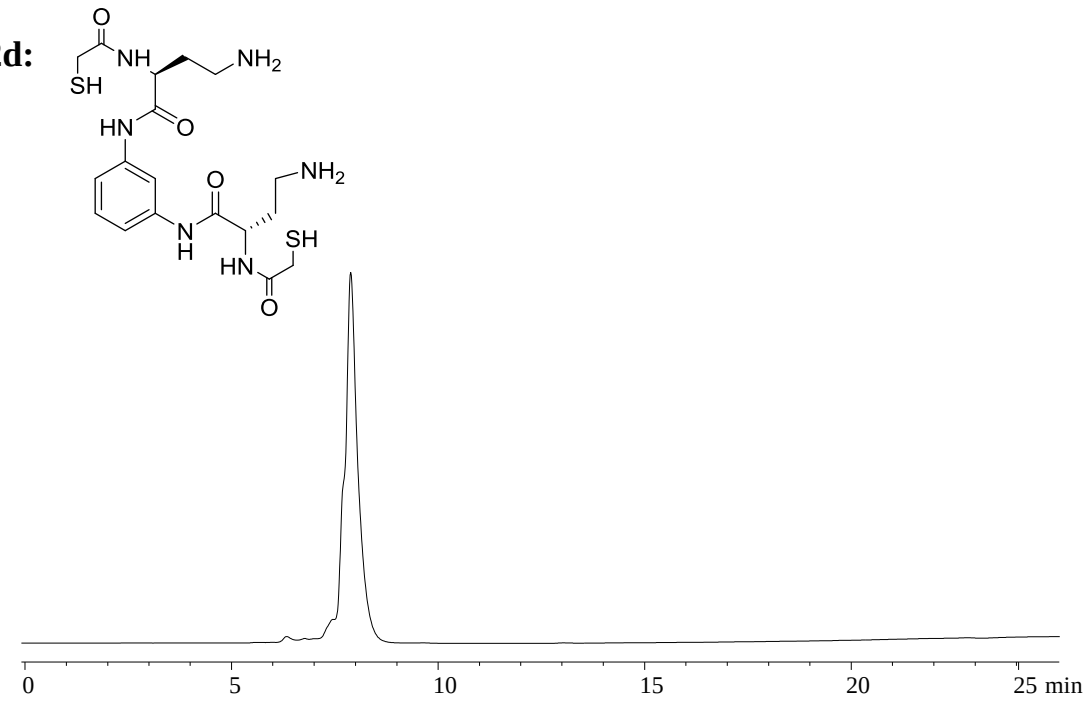


Figure S8: HPLC traces of **2d** (2 min at 1% CH_3CN in H_2O, then linear gradient from 1% to 100% CH_3CN over 25 min).

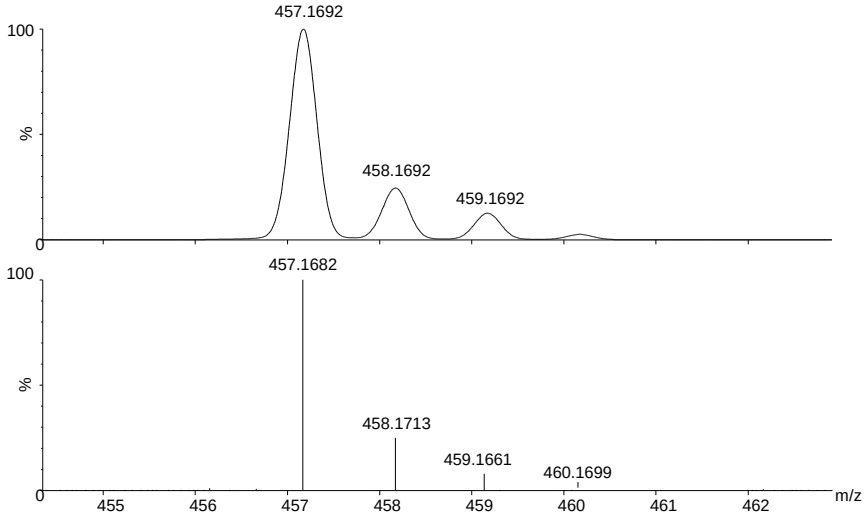


Figure S9: Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **2d**.

Product **2d** was found to aggregate and rapidly oxidize in solution so it could not be characterized by NMR methods. Instead, its oxidized form [**2d$_2$**] was prepared and fully characterized:

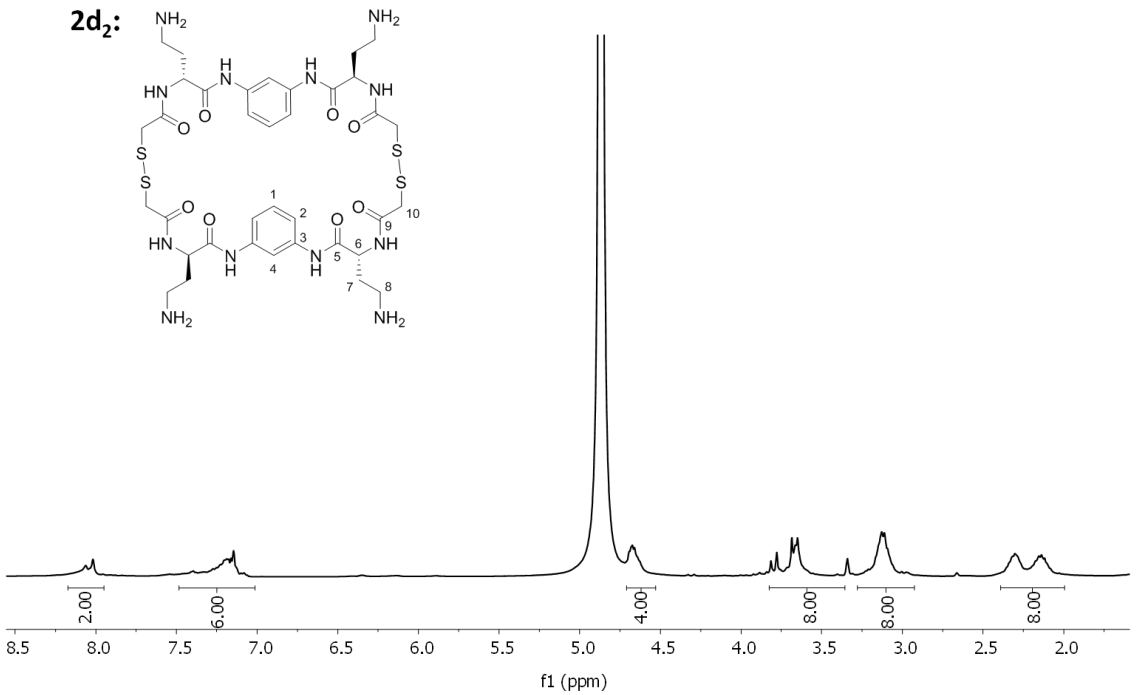


Figure S10. ^{1}H (400 MHz, 298 K in CD_3OD-d_4) spectrum of [**2d$_2$**].

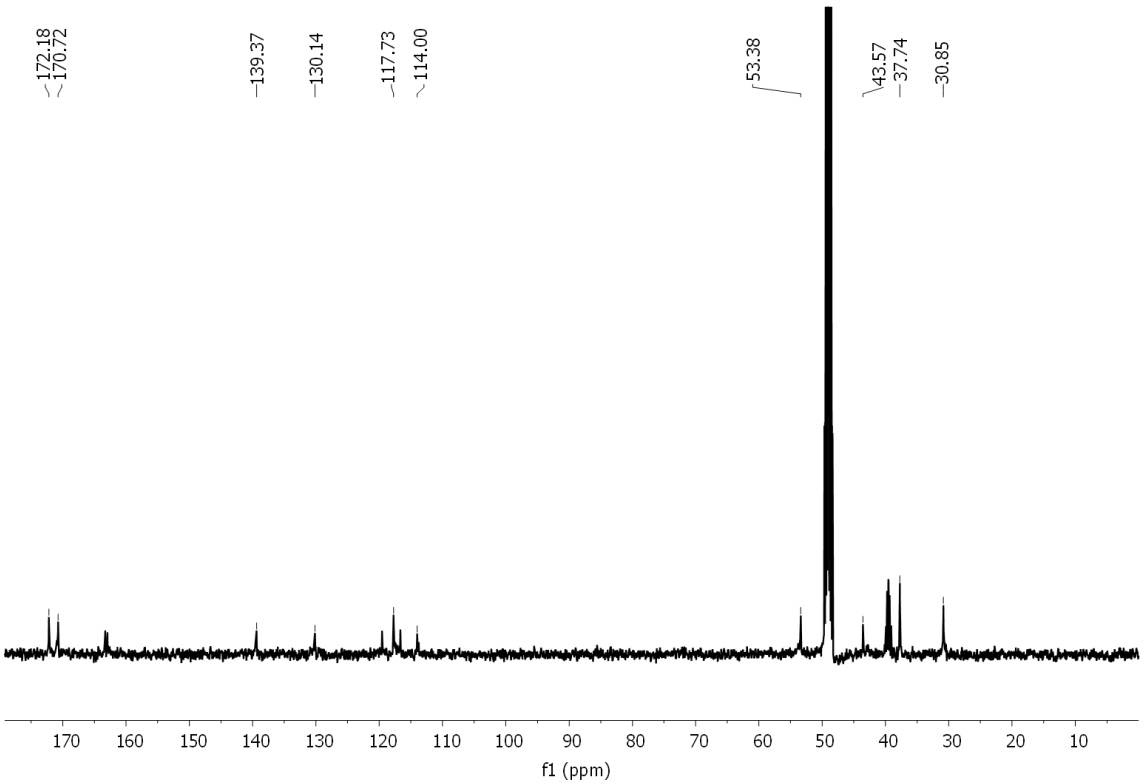


Figure S11. ^{13}C (101 MHz, 298 K in CD_3OD-d_4) spectrum of [**2d$_2$**].

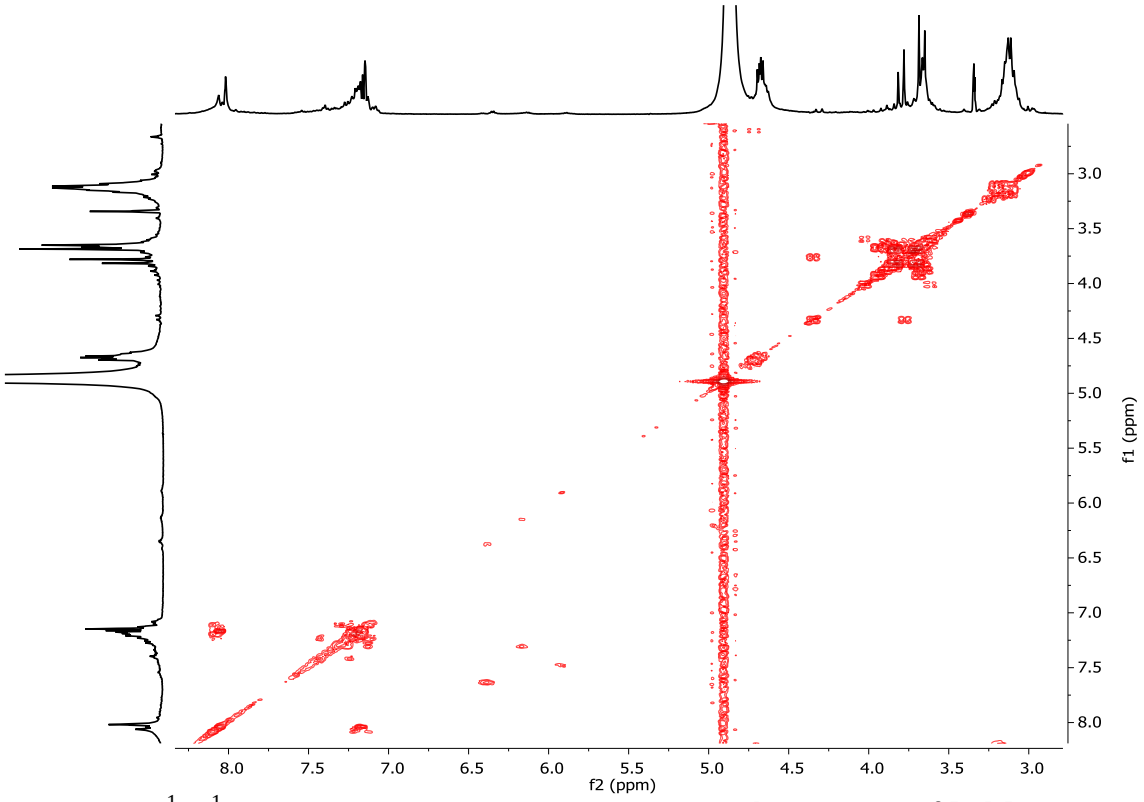
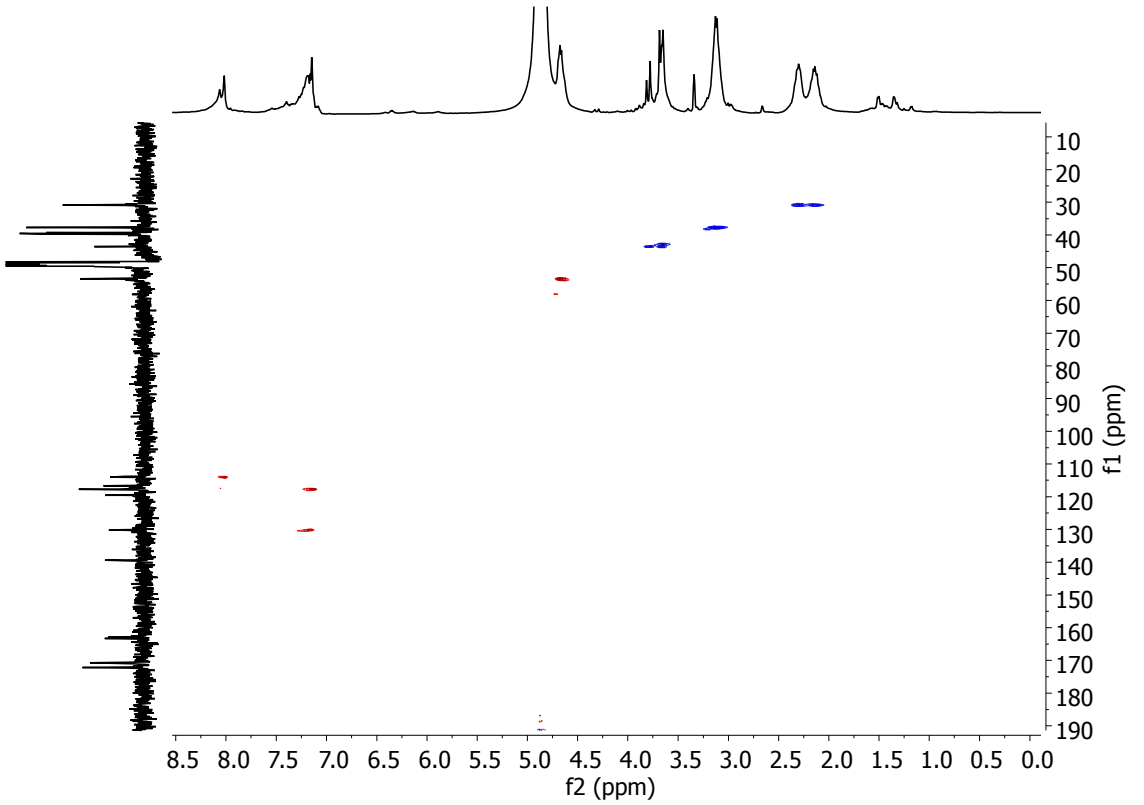


Figure S12. ^{1}H-^{1}H gCOSY (400 MHz, 298 K in CD_3OD-d_4) spectrum of [**2d$_2$**].



. **Figure S13**. ^{1}H-^{13}C gHSQC (400 MHz, 298 K in CD_3OD-d_4) spectrum of [**2d$_2$**].

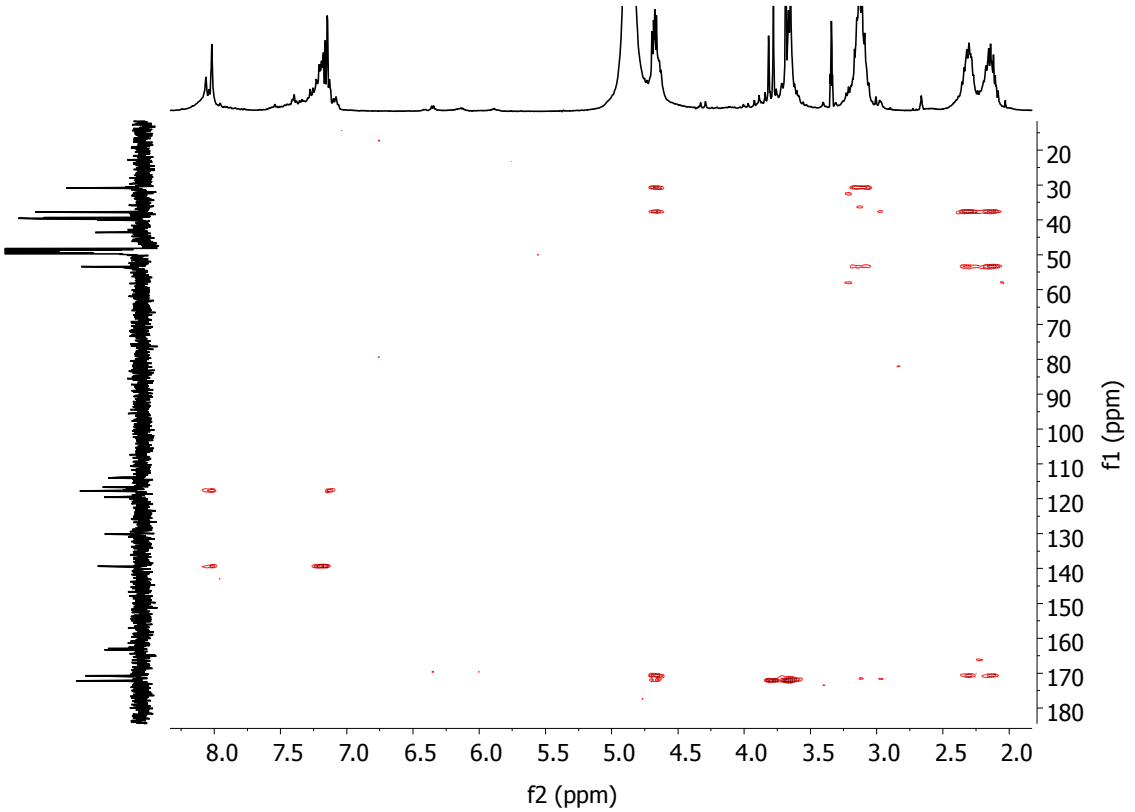


Figure S14. ^{1}H-^{13}C gHMBC (400 MHz, 298 K in CD_3OD-d_4) spectrum of [**2d$_2$**].

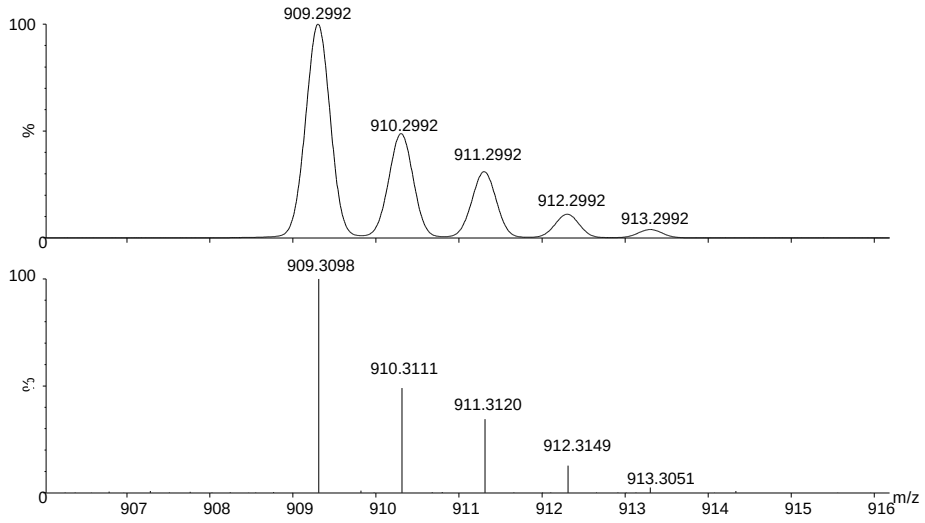


Figure S15: Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]_+$ of [**2d$_2$**].

DYNAMIC COMBINATORIAL LIBRARIES

General procedure for the quantification of the free thiols

A 4 mM solution of 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB or Ellman's reagent) was prepared by dissolving 7.93 mg in 5 mL of a 0.1 M phosphate buffer (pH 8.0). Then, 200 μL of this freshly prepared solution were placed in a well of a 96-well plate. Finally, 2.5 μL of the sample were also added in the well and, after 2 min of incubation at room temperature, the absorbance of the well was measured at 412 nm. The microplate reader was set to shake the samples for 5 seconds before each measurement. For all the batches and reaction times, the absorbance of a blank sample was also measured. The blanks were prepared by adding 2.5 μL of DMSO to 200 μL of the Ellman's reagent solution. The net absorbance was calculated by subtracting the absorbance of the corresponding blank. All the studied oxidations were carried out at room temperature, in capped vials and without any stirring.

Calibration curve

A 40 mM stock solution of cysteine was prepared by dissolving 242.32 mg in 50 mL of milli-Q water. From this, the rest of the stocks (36, 32, 28, 24, 20, 16, 12, 8, 4, 2 and 1 mM) were 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 the regression line was obtained by using the linear least square method (see Figure S16). The data showed good linear behavior within the range of working concentrations demonstrating the reliability of the quantification method.

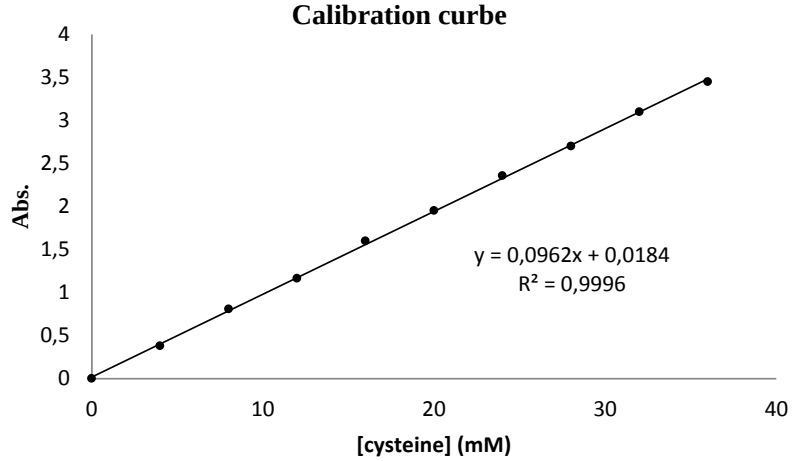


Figure S16: linear least square calibration curve.

General procedure for the preparation and HPLC analysis of the DCLs

General procedure for the preparation of the DCLs

A 66.7 mM BIS-TRIS methane buffer was prepared by dissolving 1.39 g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq). For the experiments performed at different pH values, a 100 mM BIS-Tris propane buffer was used for pH 6.5-9.5, and a 100 mM $HCOOH/HCOO^-$ buffer was used for pH 4.5. The BIS-Tris propane buffer was prepared by dissolving 2.82g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution

to 6.5, 7.5, 8.5 and 9.5 by the addition of HCl (aq). The HCOOH/HCOO$^-$ buffer was prepared by dissolving 0.45 mL of the formic acid in 100 mL of milli-Q water and adjusting the pH of the solution to 4.5 by the addition of NaOH (aq).
Individual concentrated stock solutions for the different building blocks (BBs) were prepared in DMSO. Mixture stock solutions containing the necessary BBs for the generation of the libraries (**1**, **2a-e** and **3**) were prepared from these individual solutions. The reaction mixtures were then prepared by dilution of the stock solutions ensuring no differences in concentration between the reaction mixtures of the same batch. The reaction mixtures samples were prepared by adding 50 μL of the corresponding BBs stock mixture in DMSO to 150 μL of a solution of 66.7 mM BIS-TRIS buffer or HCOOH/HCOO$^-$ buffer. Unless otherwise specified, the DCLs were prepared at final concentrations of 0.5 mM for the di- and tripodal BBs (**1, 2** and **3**) in a 50 mM Buffer with 25% DMSO. The concentrations for L-cysteine (Cys) and cysteamine (CyA) are specified for each of the experiments. All the studied oxidations were carried out at room temperature, in capped vials and without any stirring.

General procedure for the analysis of the DCLs
The mixtures were analyzed by means of HPLC or UPLC-MS at different reaction times. Complete oxidation is achieved after 24h for reactions containing 25% DMSO. The HPLC samples were prepared by adding 45 μL of the corresponding reaction mixture to 75 μL of a solution of 89% H_2O, 10% CH_3CN and 1% TFA. The different experiments were analyzed with two different HPLC programs: in libraries containing the bipodals units **2d** and **2e** 2 min at 2% CH_3CN in H_2O, then linear gradient from 2% to 40% CH_3CN over 48 min was used. For the other experiments the eluent used consisted in 2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 40% CH_3CN over 48 min.

Mixture of 1 +2a in the presence or absence L-Cys (25% DMSO, pH=6.5)

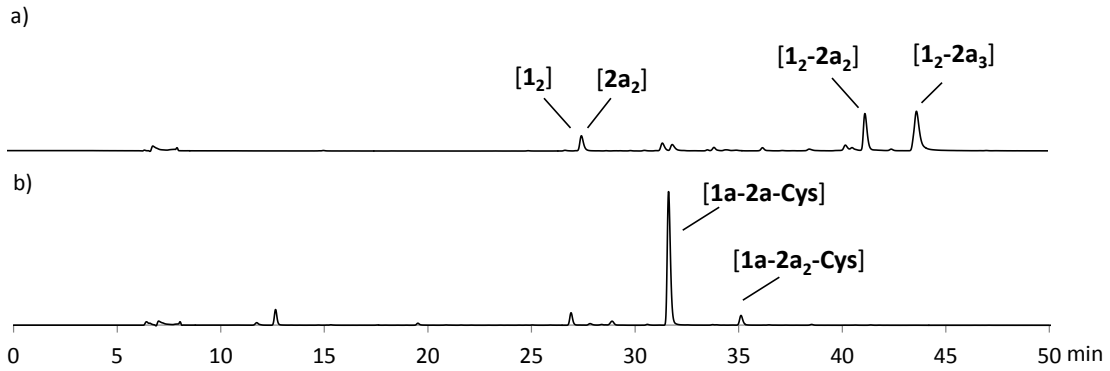


Figure S17. HPLC traces of the mixture **1**+**2a** (0.5mM each) in the absence (a), and in the presence of 2.5mM of L- **Cys** (b).

Mixture of 1 +2b in the presence or absence L-Cys (25% DMSO, pH=6.5)

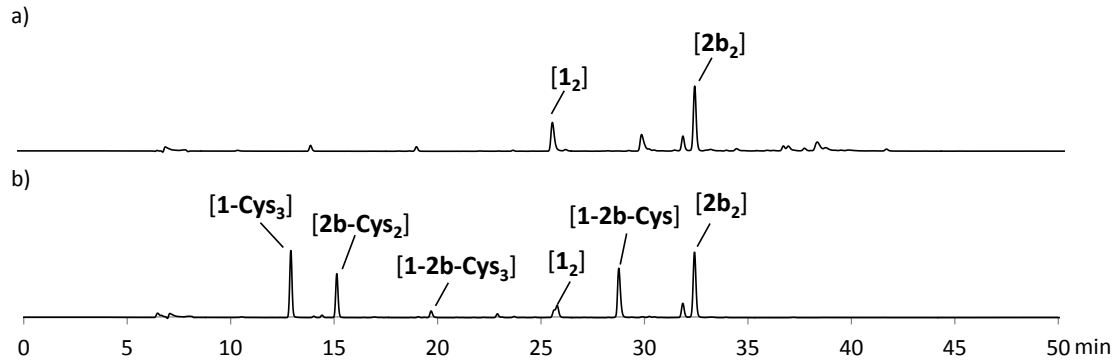


Figure S18. HPLC traces of the mixture **1**+**2b** (0.5mM each) in the absence (a), and in the presence of 2.5mM of L-Cys (b).

Mixture of 1 +2c in the presence or absence L-Cys (25% DMSO, pH=6.5)

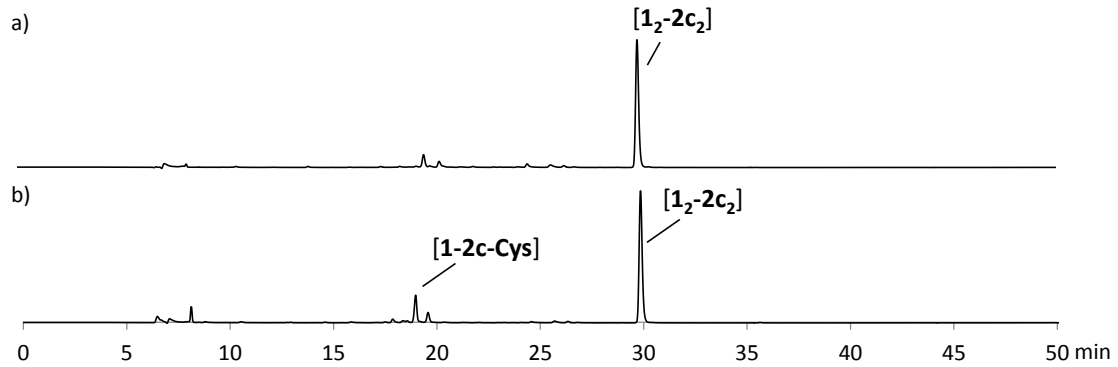


Figure S19. HPLC traces of the mixture **1**+**2c** (0.5mM each) in the absence (a), and in the presence of 2.5mM of L-Cys (b).

Mixture of 1 +2d in the presence or absence L-Cys (25% DMSO, pH=6.5)

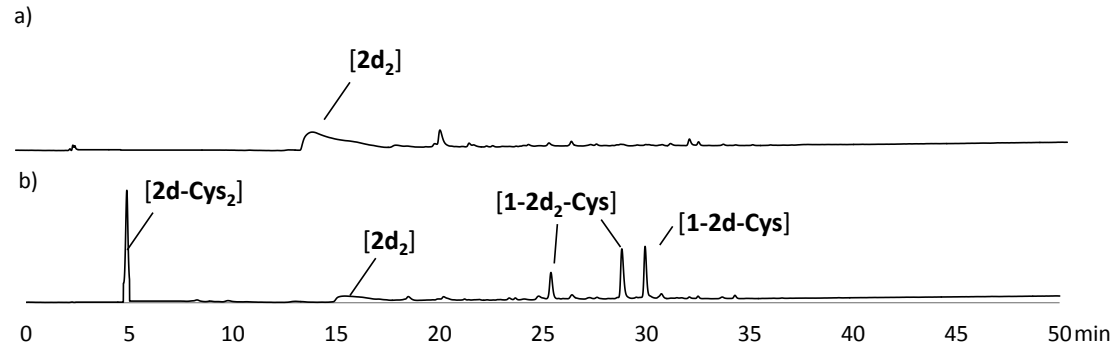


Figure S20. HPLC traces of the mixture **1**+**2d** (0.5mM each) in the absence (a), and in the presence of 2.5mM of L-Cys (b).

Mixture of 1 +2e in the presence or absence L-Cys (25% DMSO, pH=6.5)

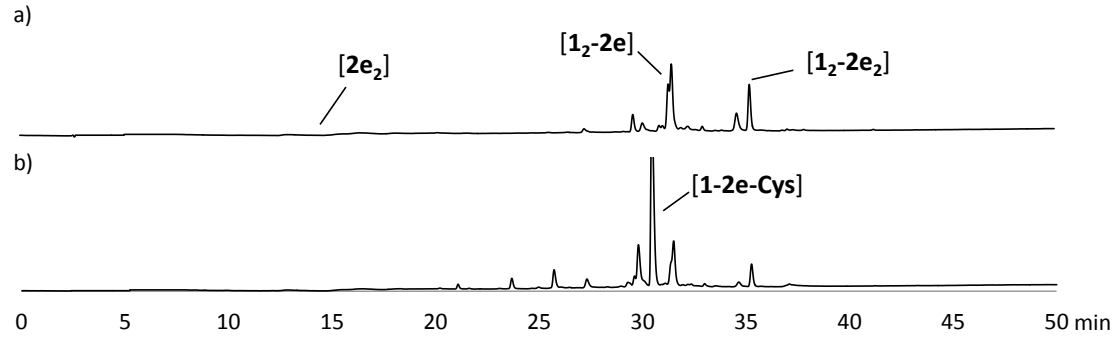


Figure S21. HPLC traces of the mixture **1**+**2e** (0.5mM each) in the absence (a), and in the presence of 2.5mM of L-Cys (b).

Mixture of 1 with different bipodals (2c, 2d or 2e) in the presence or absence L-Cys (25% DMSO, pH=6.5)

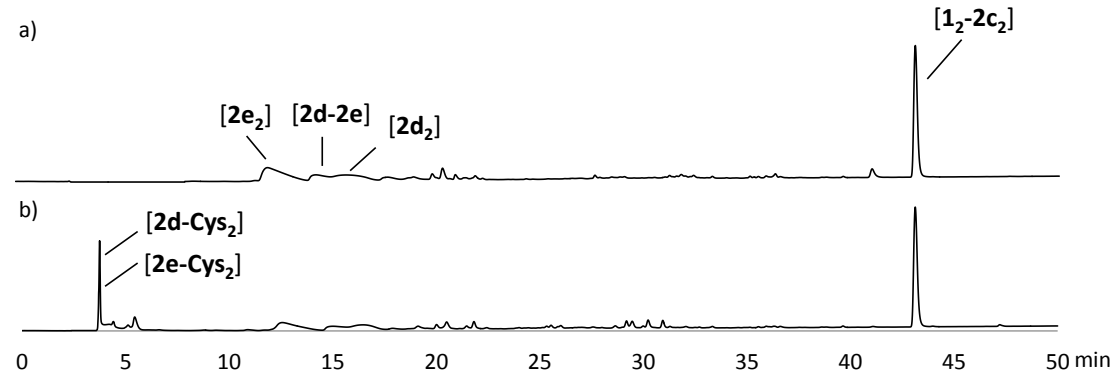


Figure S22. HPLC traces of the mixture of **1**+**2c**+**2d**+**2e** (0.5 mM each) in the absence (a), and in the presence of 2.5mM of L-Cys (b).

Mixture of 3 and different bipodals (2c, 2d or 2e) in the presence or absence of L-Cys (25% DMSO, pH=6.5)

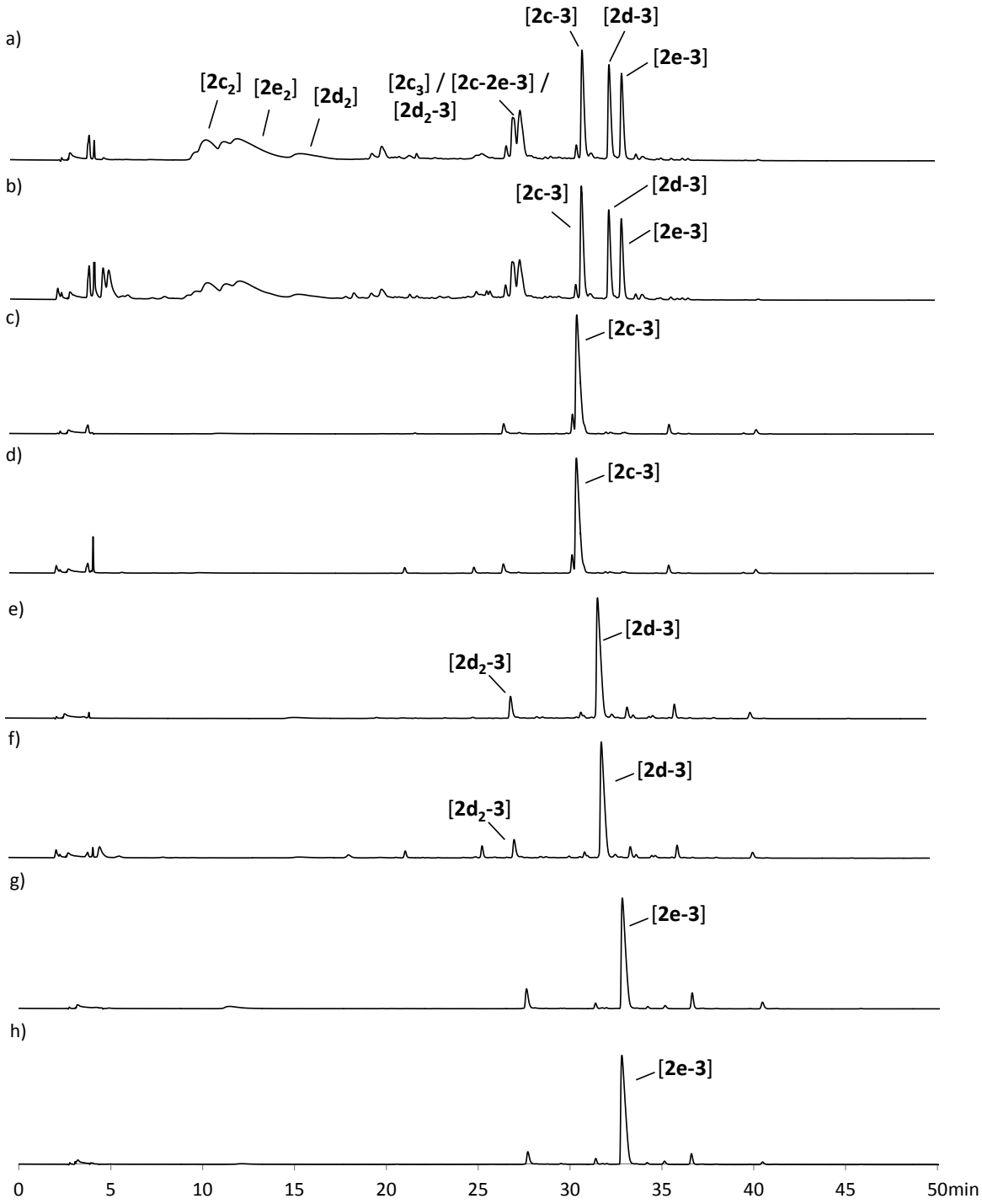


Figure S23. HPLC traces of the mixture of **3**+**2c**/**2d**/**2e** (0.5 mM each) in the presence or absence of L-Cys (2.5mM): (a) **3**+**2c**+**2d**+**2e**, (b) **3**+**2c**+**2d**+**2e** with **Cys**, (c) **3**+**2c**, (d) **3**+**2c** with **Cys**, (e) **3**+**2d**, (f) **3**+**2d** with **Cys**, (g) **3**+**2e**, (h) **3**+**2e** with **Cys**.

Mixture of 1+2c in the presence or absence of Cysteamine, (CyA) (25%DMSO, pH=6.5)

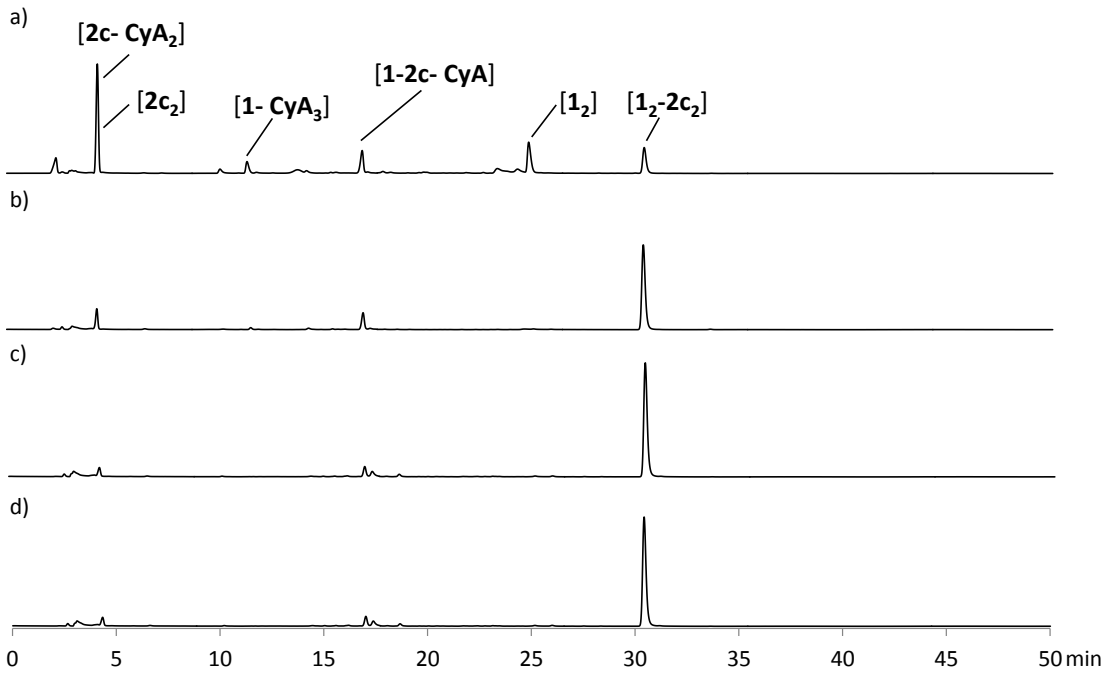


Figure S24. HPLC traces of the mixture **1**+**2c** (0.5mM each) with 25mM of **CyA** (a), 2.5mM of **CyA** (b), 0.25mM of **CyA** (c), without **CyA** (d).

Mixture of 1+2c at different reactions times (25%DMSO, pH=6.5)

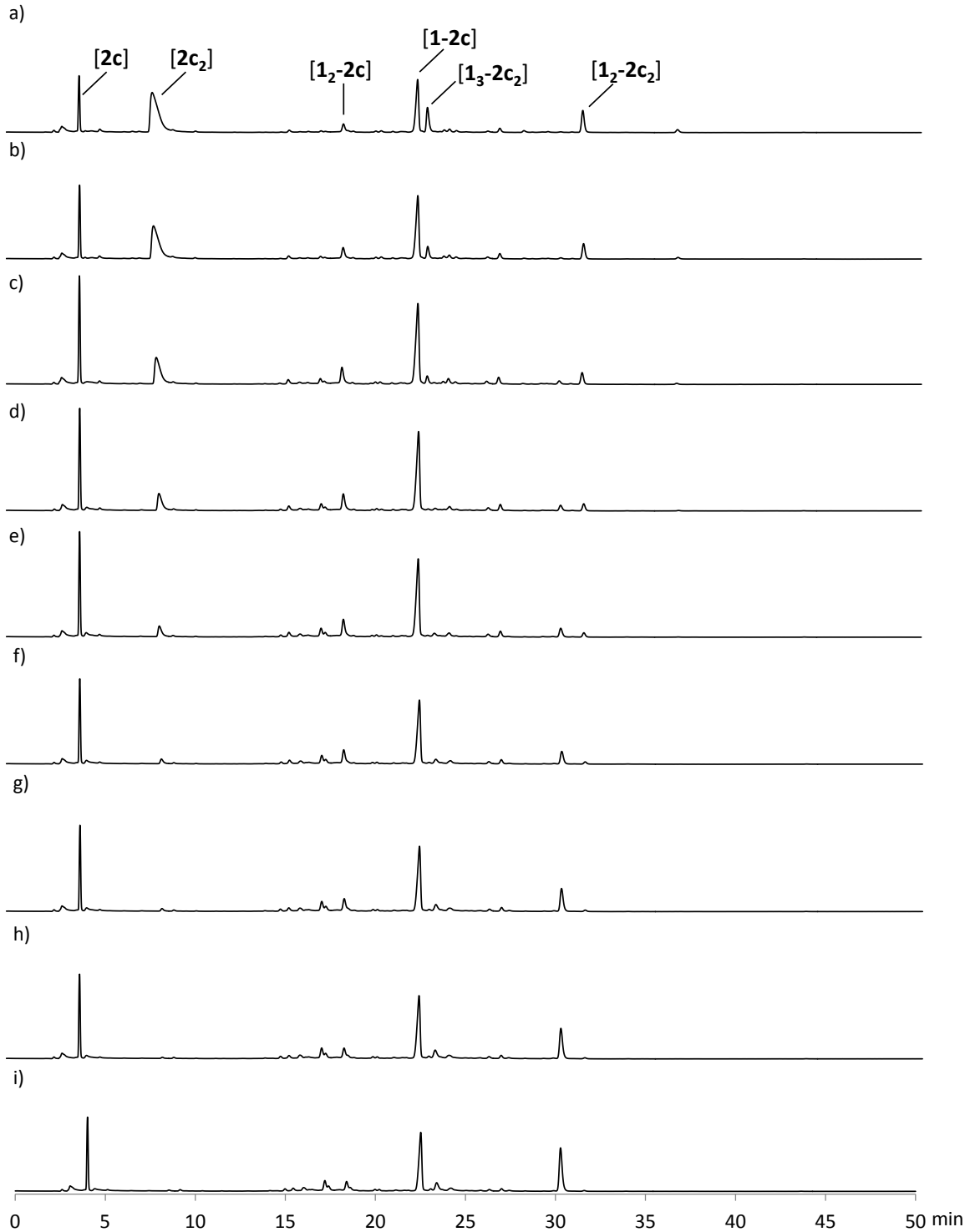


Figure S25. HPLC traces of the mixture **1**+**2c** (0.5mM each), at pH 6.5 with 25% of DMSO at different reaction times: 0 min (a), 2 min (b), 5 min (c), 10 min (d), 15 min (e), 30 min (f), 45 min (g), 1 hour (h), 1.50 hours (i).

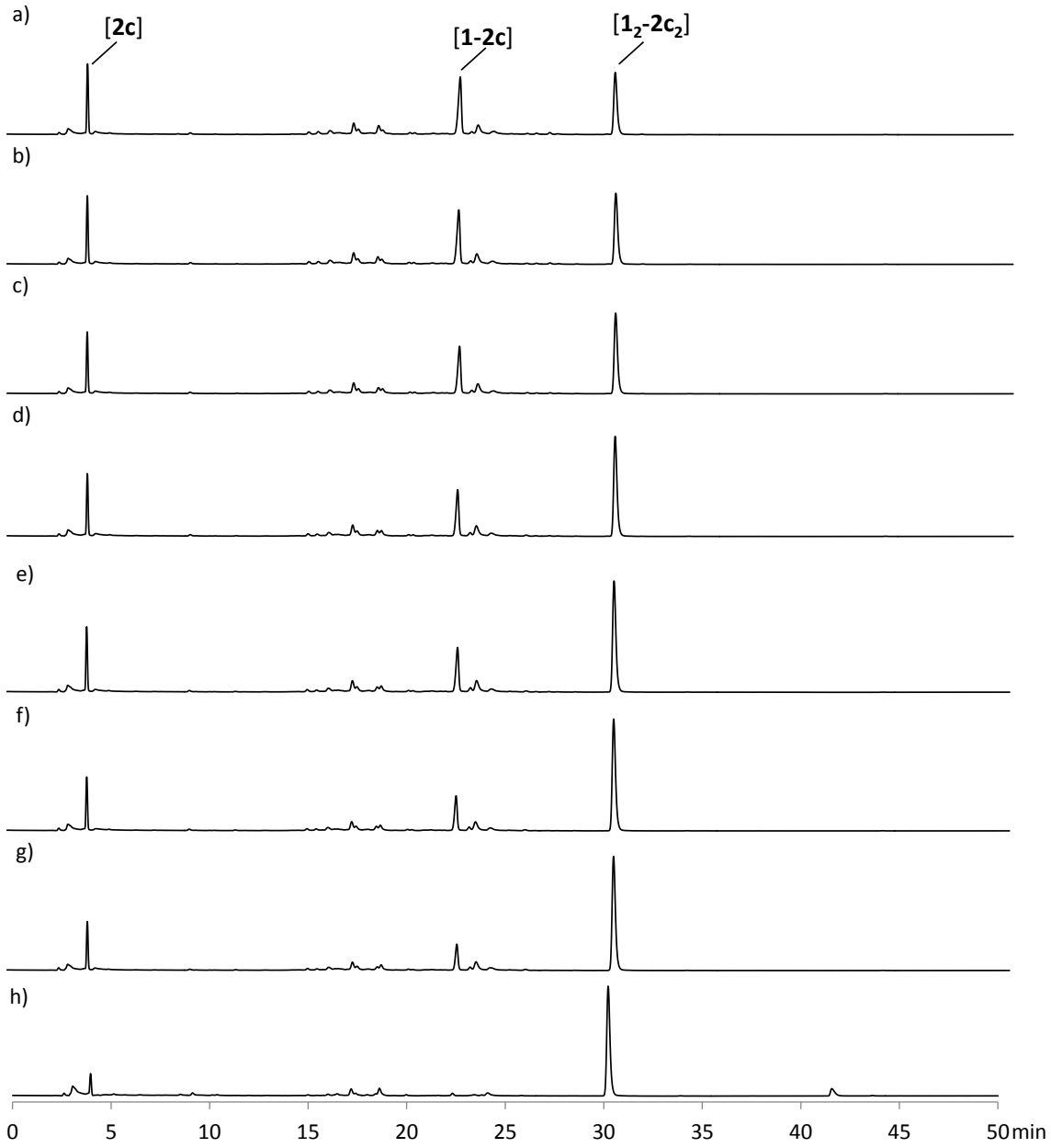


Figure S26. HPLC traces of the mixture **1**+**2c** (0.5mM each), at pH 6.5 with 25% of DMSO at different reaction times: 2 hours (a), 2.50 hours (b), 3 hours (c), 3.50 hours (d), 4 hours (e), 5 hours (f), 6 hours (g), 22 hours (h).

Mixture of 1+2c at different reactions pH values (25% DMSO)

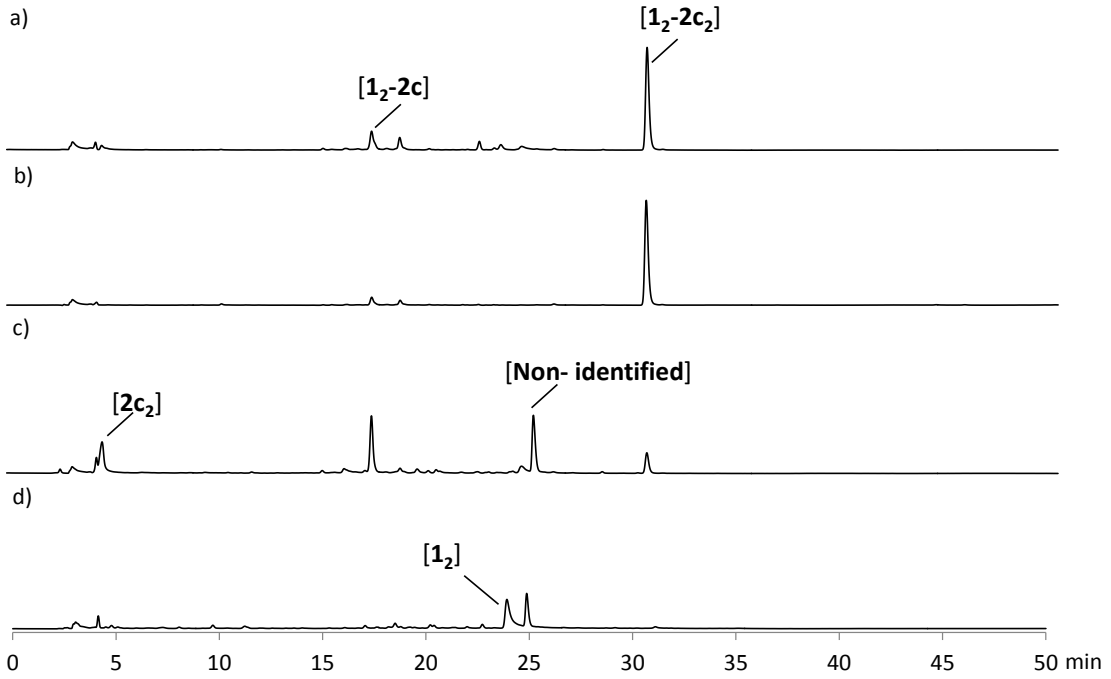


Figure S27. HPLC traces of the mixture **1**+**2c** (0.5mM each) at different pH values: 4.5 (a), 6.5 (b), 8.5 (c), 9.5 (d).

Mixture of 1+2c at different NaCl concentrations (25% DMSO, pH=6.5)

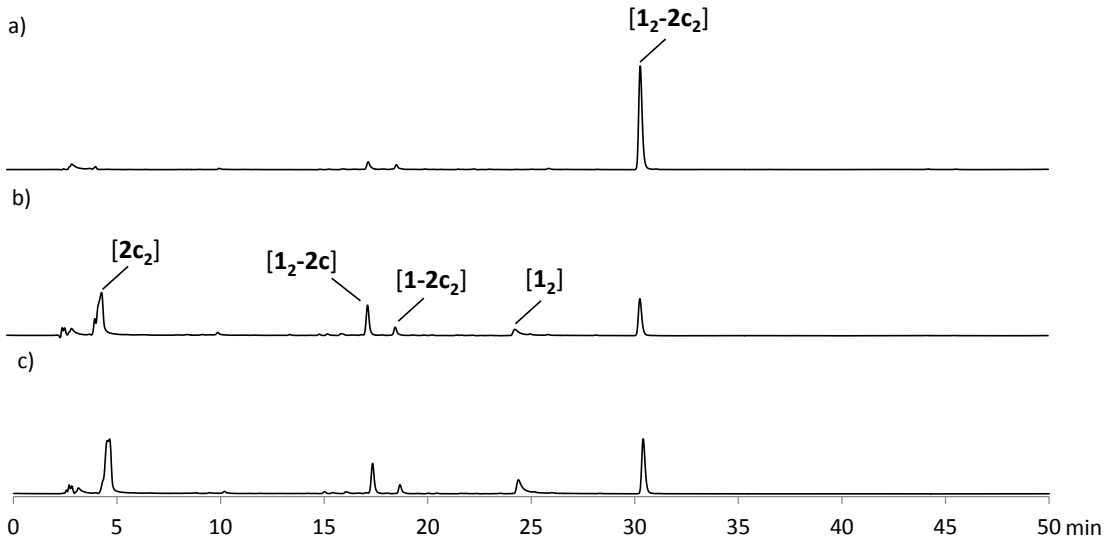


Figure S28. HPLC traces of the mixture **1**+**2c** (0.5mM each) at different sodium chloride concentrations: 0.0 M (a), 1.0 M (b), 2.0 M (c).

Mixture of 1+2c with different buffered solutions (25% DMSO, pH=6.5)

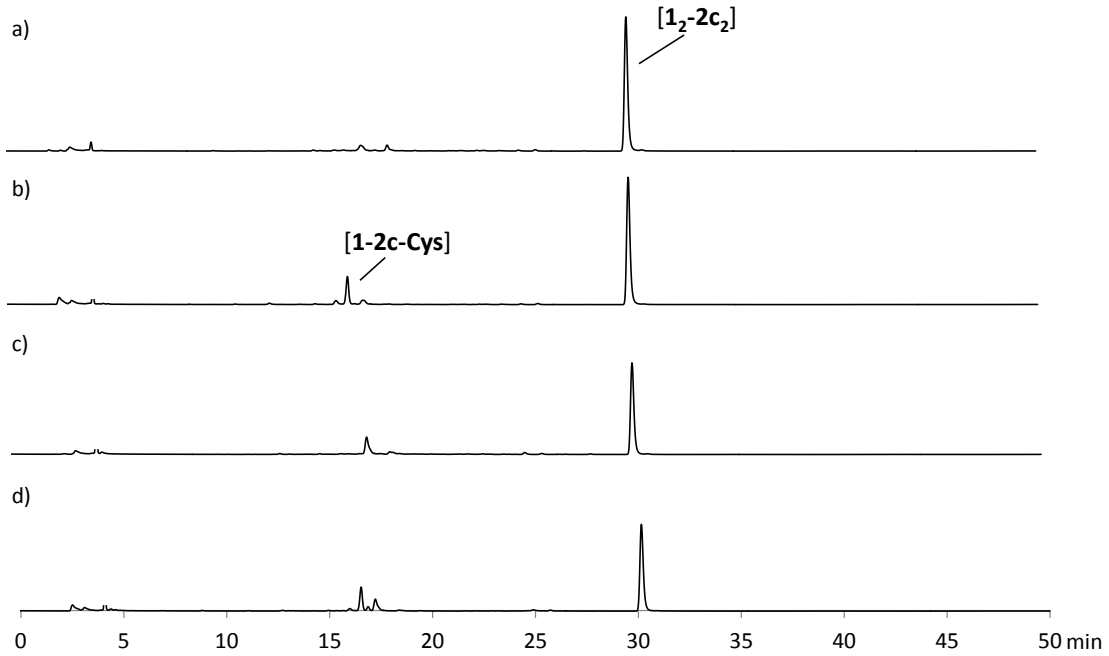


Figure S29. HPLC traces of the mixture **1**+**2c** (0.5mM each) with and without L-Cys using different buffered solutions: **1+2c** in Bis-Tris methane (a), **1+2c+Cys** in Bis-Tris methane (b), **1+2c** in Phosphate Buffer (c), **1+2c+Cys** in Phosphate Buffer (d).

Mixture of 1+2c with different percentage of DMSO (pH=6.5)

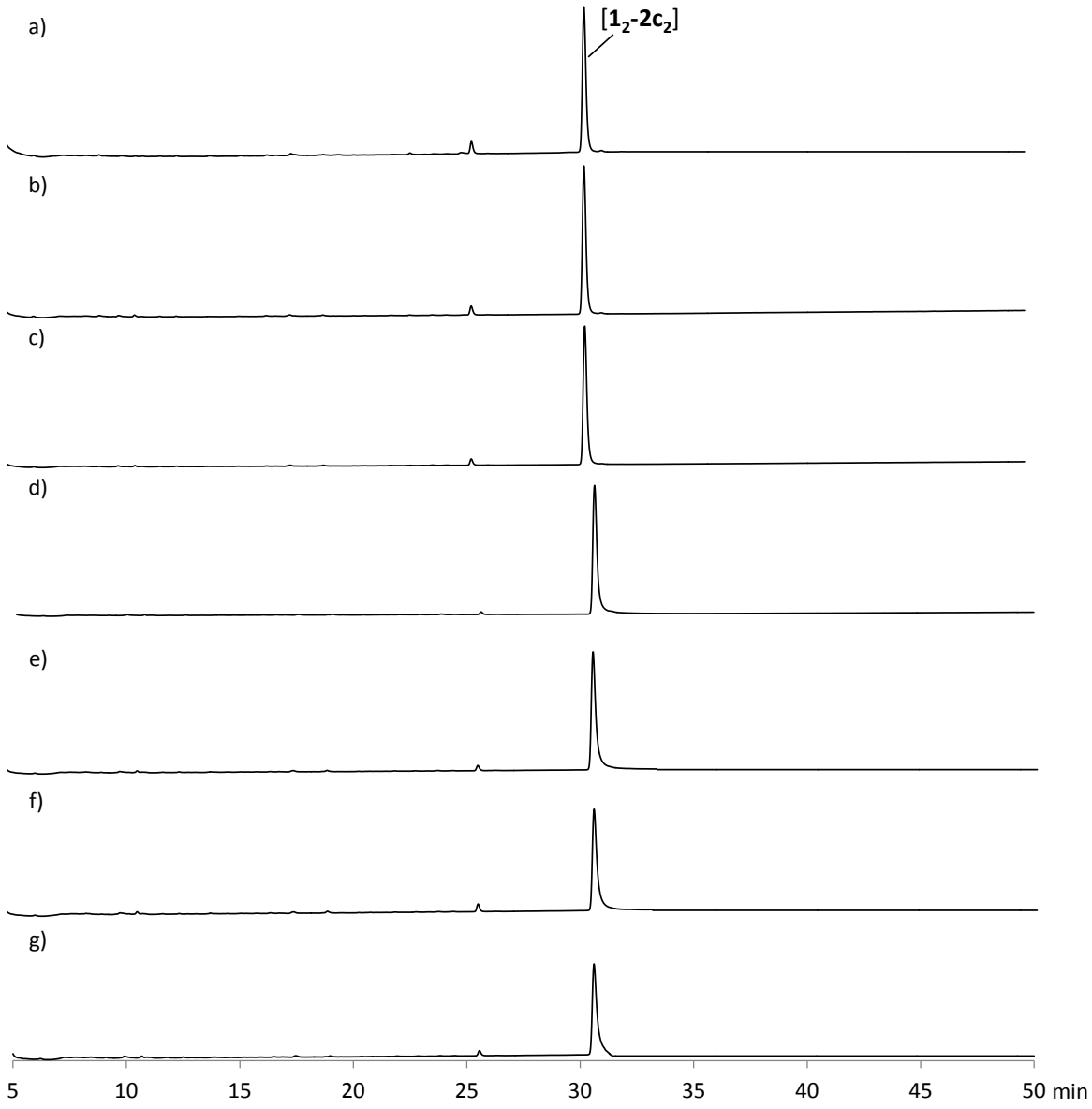


Figure S30. HPLC traces of the mixture **1**+**2c** (0.5mM each) at pH 6.5, after complete oxidation at different percentage of DMSO: 25% (a), 10% (b), 5% (c), 2% (d), 1% (e), 0.5% (f), 0% (g).

MASS SPECTROMETRY OF THE DCL's

General procedure for the analysis of the DCLs by HRMS

The HRMS (UPLC-ESI-TOF) samples were prepared by adding 20 μL of the corresponding reaction mixture to 40 μL of a solution of 89% H_2O, 10% MeCN and 1% TFA. The different experiments were analysed with two different programs, for the libraries containing the bipodal units **2e** and **2d** 2.5 min at 2% CH_3CN (+ 20mM HCOOH) in H_2O (+ 20mM HCOOH), then linear gradient from 2% to 40% CH_3CN over 27.5 min was used. For the other experiments the eluent mixture used was: 2.5 min at 5% CH_3CN (+ 20mM HCOOH) in H_2O (+ 20mM HCOOH), then linear gradient from 5% to 50% CH_3CN over 27.5 min.

Mixture of 1 +2a (0.5 mM each, 25% DMSO, pH=6.5)

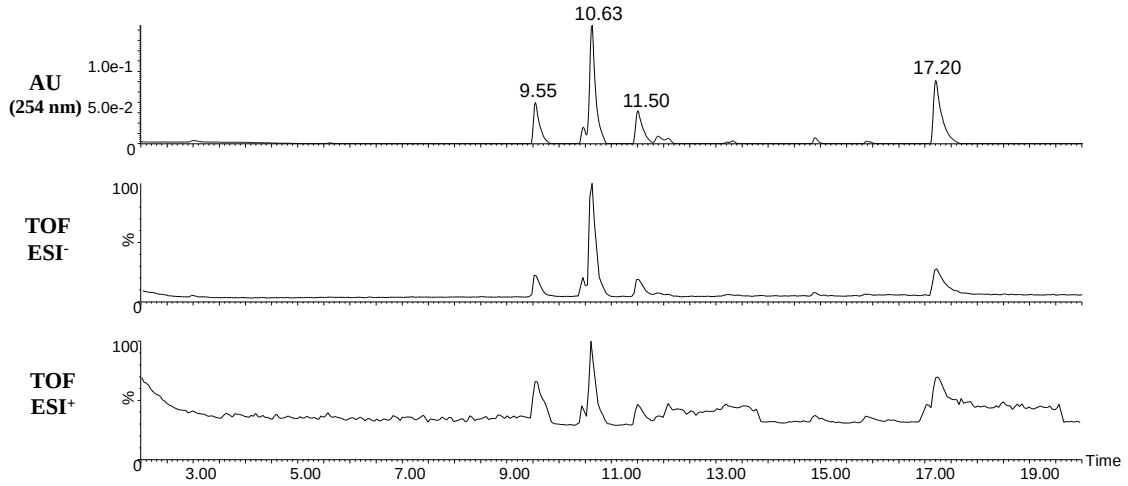


Figure S31. UPLC_ESI_TOF for **1+2a** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[$2a_2$]
Retention time: 9.55 min.
Chemical Formula: $C_{32}H_{40}N_8O_{12}S_4$
Exact Mass: 856.1649

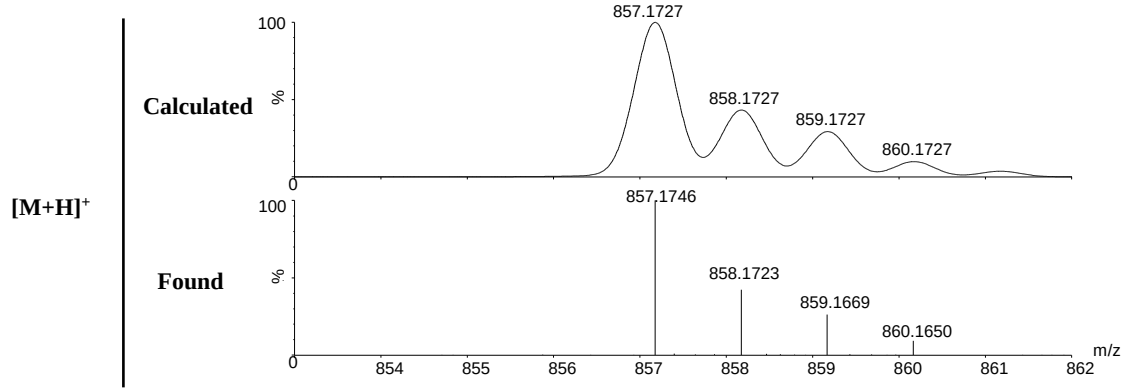


Figure S32. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**$2a_2$**].

[1_2-2a]

Retention time: 10.63 min.
Chemical Formula: $C_{52}H_{56}N_{10}O_{24}S_8$
Exact Mass: 1460.1235

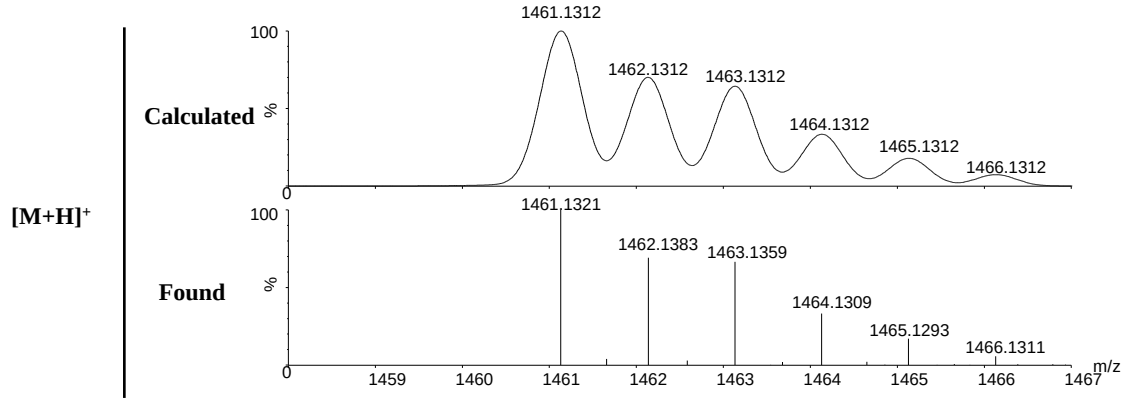


Figure S33. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1_2-2a**].

[1_2-2a$_2$]
Retention time: 11.50 min.
Chemical Formula: $C_{68}H_{76}N_{14}O_{30}S_{10}$
Exact Mass: 1888.2059

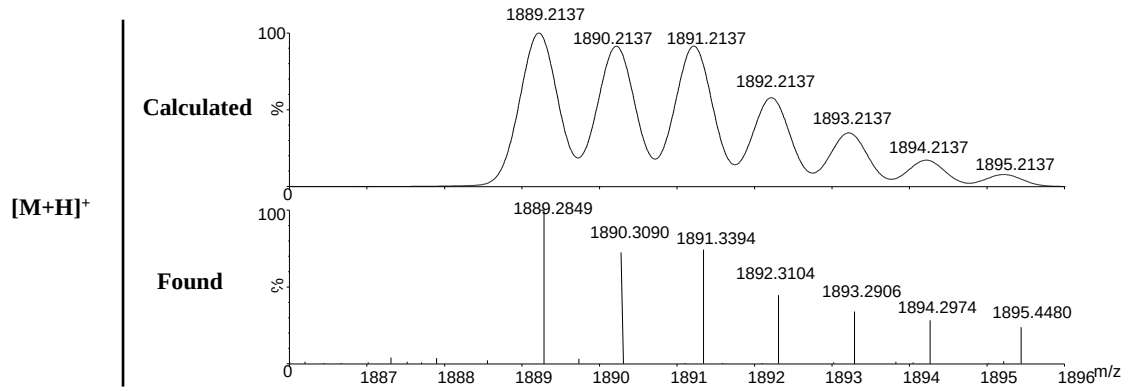


Figure S34. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1_2-2a$_2$**].

[1_2-2a$_3$]
Retention time: 17.20 min.
Chemical Formula: $C_{84}H_{96}N_{18}O_{36}S_{12}$
Exact Mass: 2316.2883

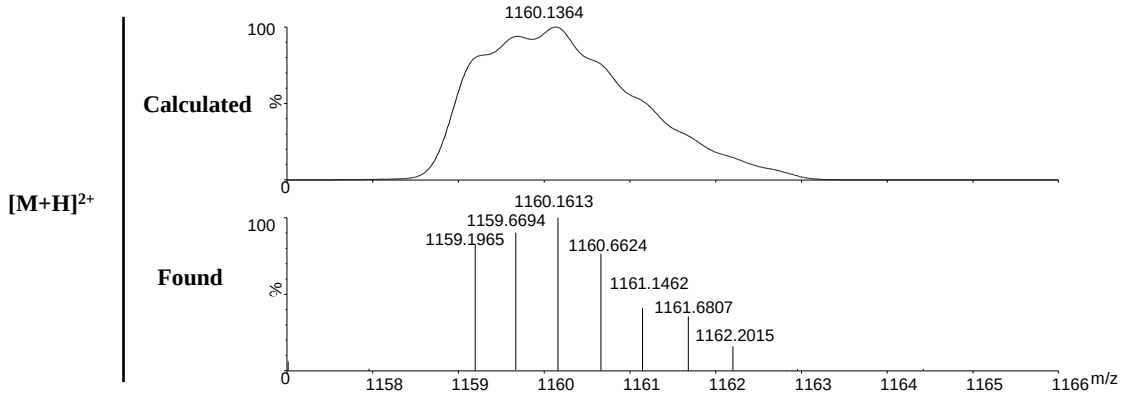


Figure S35. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of [$\mathbf{1_2}$-$\mathbf{2a_2}$].

Mixture of 1 +2a+L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)

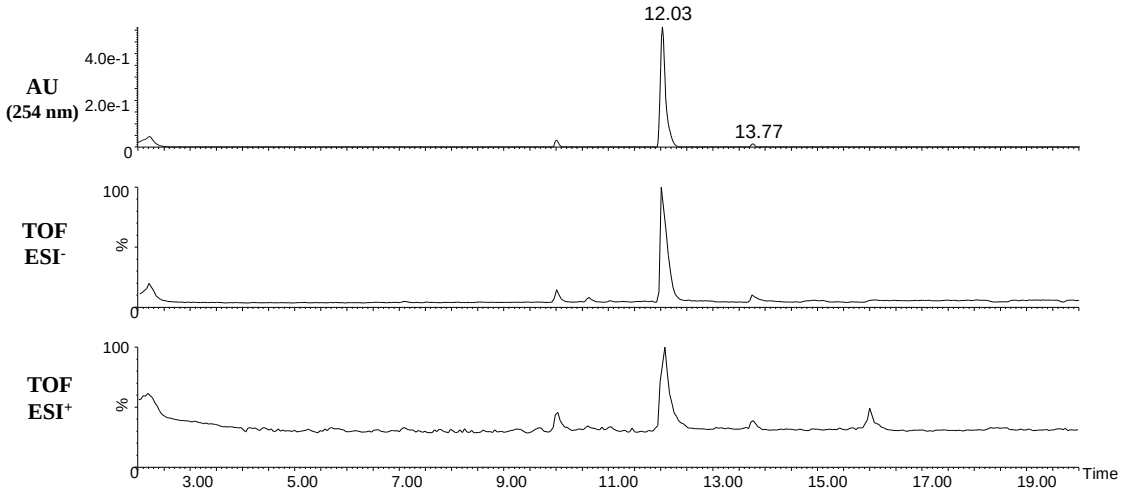


Figure S36. UPLC_ESI_TOF for **1+2a+L-Cys** (0.5 + 0.5 + 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[1-2a-Cys]
Retention time: 12.03 min.
Chemical Formula: $C_{37}H_{44}N_8O_{17}S_6$
Exact Mass: 1064.1149

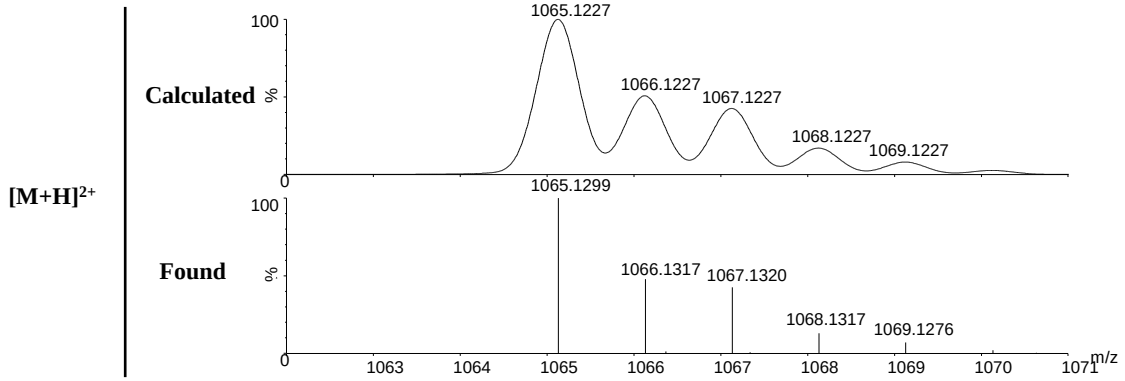


Figure S37. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2a-Cys**].

[1-2a$_2$-Cys]
Retention time: 13.77 min.
Chemical Formula: $C_{53}H_{64}N_{12}O_{23}S_8$
Exact Mass: 1492.1973

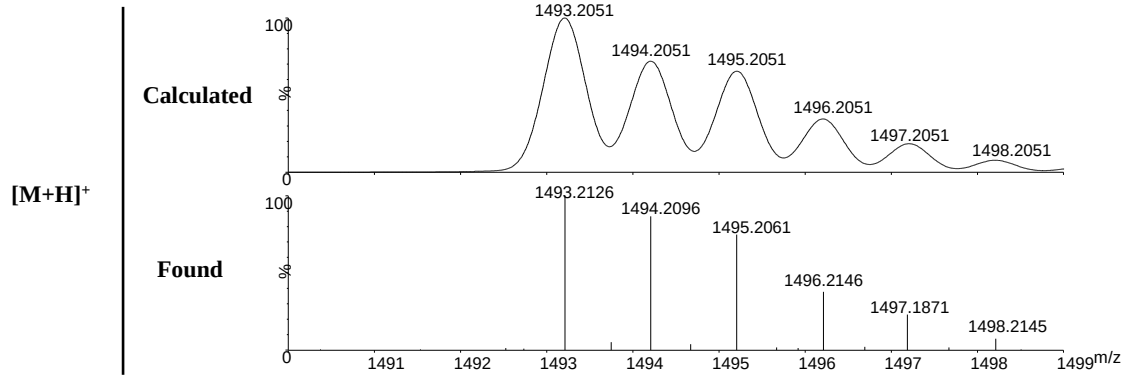


Figure S38. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2a$_2$-Cys]**.

<u>Mixture of 1 +2b (0.5 mM each, 25% DMSO, pH=6.5)</u>

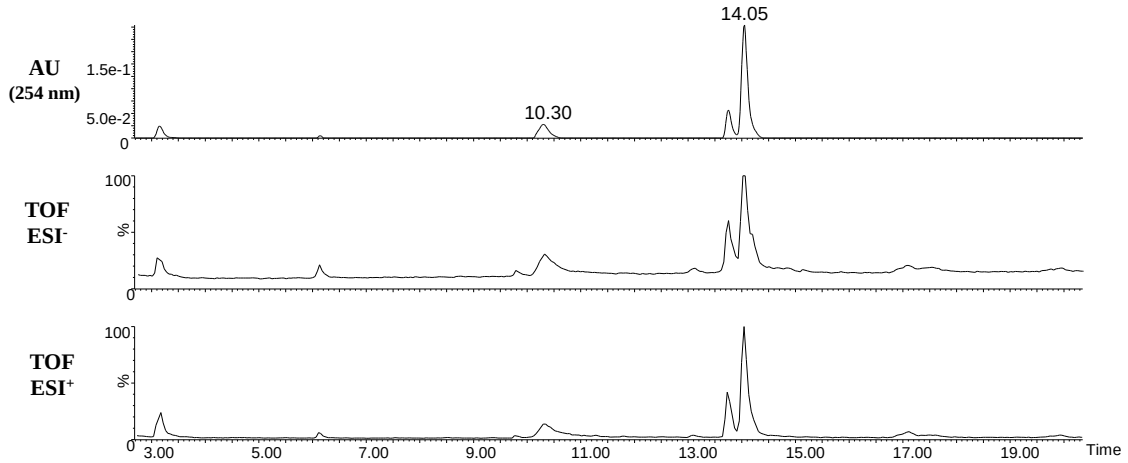


Figure S39. UPLC_ESI_TOF for **1+2b** (0.5 mM each, pH.6.5, 25% DMSO).

<u>Identification of the products:</u>

[1$_2$]
Retention time: 10.30 min.
Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$
Exact Mass: 1032.0410

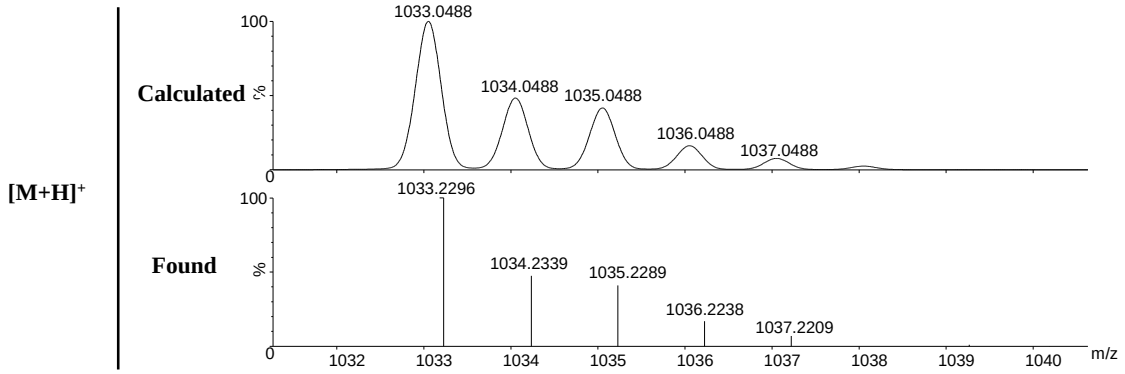


Figure S40. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [$\mathbf{1_2}$].

[$\mathbf{2b_2}$]
Retention time: 10.30 min.
Chemical Formula: $C_{36}H_{40}N_8O_{16}S_4$
Exact Mass: 968.1445

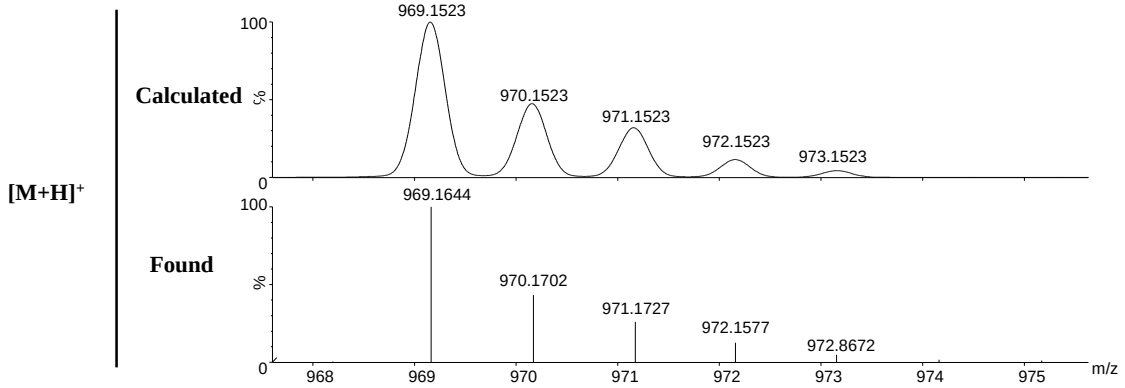


Figure S41. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [$\mathbf{2b_2}$].

<u>Mixture of 1 +2b+L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)</u>

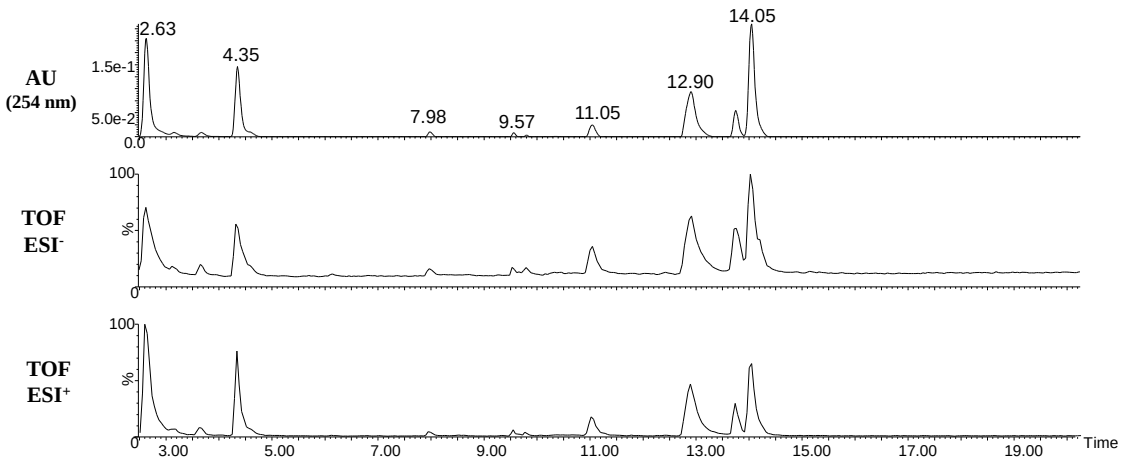


Figure S42. UPLC_ESI_TOF for **1+2b+L-Cys** (0.5 + 0.5 + 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[1-Cys$_3$]
Retention time: 2.63 min.
Chemical Formula: $C_{27}H_{36}N_6O_{15}S_6$
Exact Mass: 876.0563

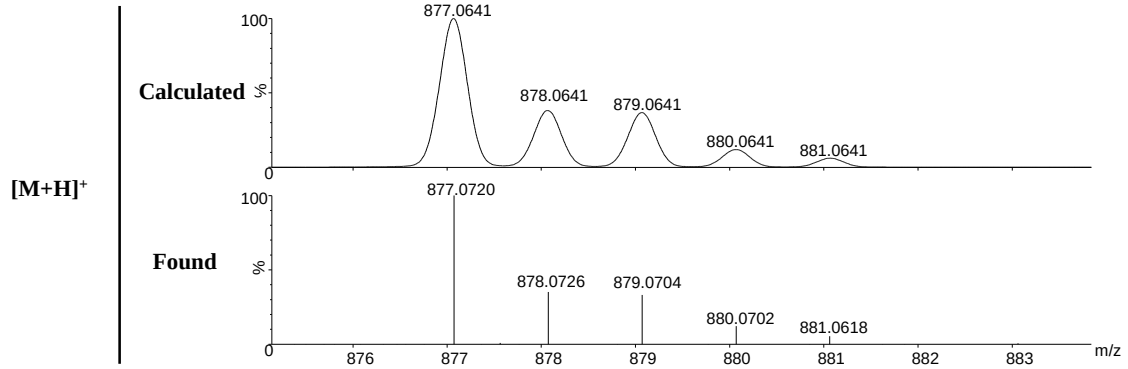


Figure S43. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-Cys$_3$]**.

[2b-Cys$_2$]
Retention time: 4.35 min.
Chemical Formula: $C_{24}H_{32}N_6O_{12}S_4$
Exact Mass: 724.0961

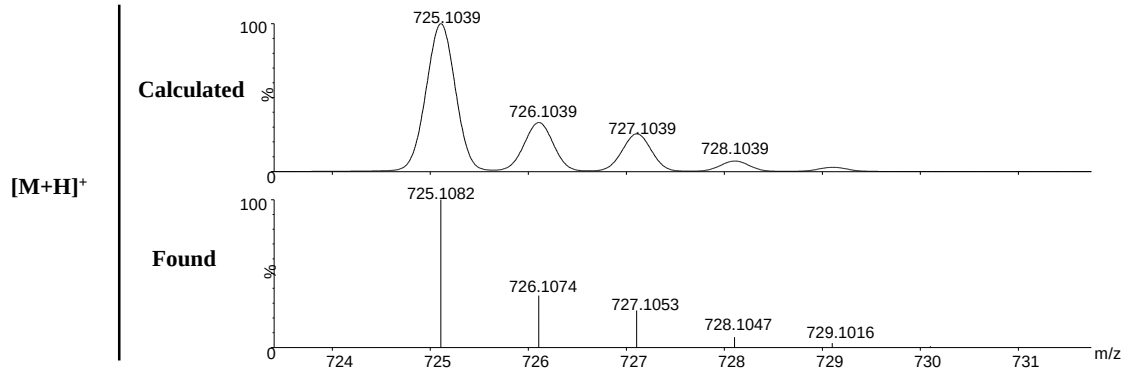


Figure S44. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2b-Cys$_2$]**.

[1-2b-Cys$_3$]
Retention time: 7.98 min.
Chemical Formula: $C_{45}H_{56}N_{10}O_{23}S_8$
Exact Mass: 1360.1285

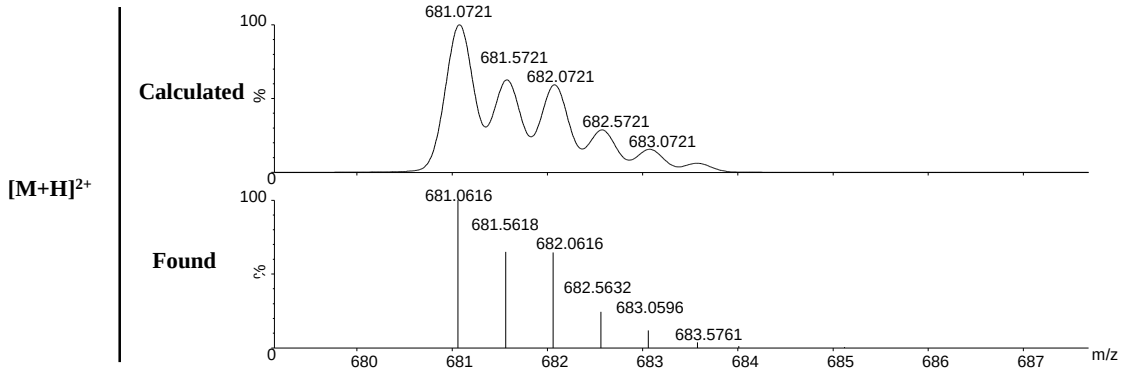


Figure S45. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of **[1-2b-Cys$_3$]**.

[2b$_2$-Cys$_2$]
Retention time: 9.57 min.
Chemical Formula: $C_{42}H_{52}N_{10}O_{20}S_6$
Exact Mass: 1208.1684

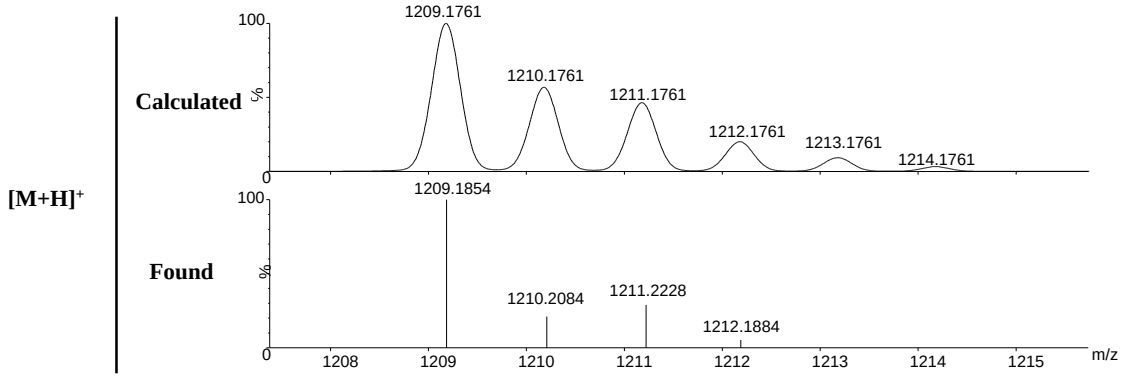


Figure S46. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{+}$ of **[2b$_2$-Cys$_2$]**.

[1$_2$] *(Previously described)*
Retention time: 11.05 min.
Chemical Formula: $C_{36}H_{36}N_6O_{18}S_6$
Exact Mass: 1032.0410

[1-2b-Cys]
Retention time: 12.90 min.
Chemical Formula: $C_{39}H_{44}N_8O_{19}S_6$
Exact Mass: 1120.1047

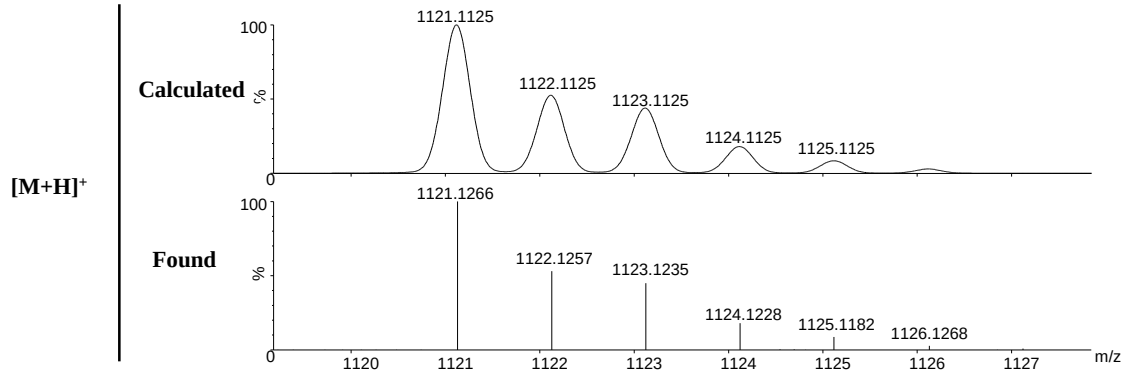


Figure S47. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2b-Cys**].

[2b$_2$] *(Previously described)*
Retention time: 14.05 min.
Chemical Formula: $C_{36}H_{40}N_8O_{16}S_4$
Exact Mass: 968.1445

<u>Mixture of 1 +2c (0.5 mM each, 25% DMSO, pH=6.5)</u>

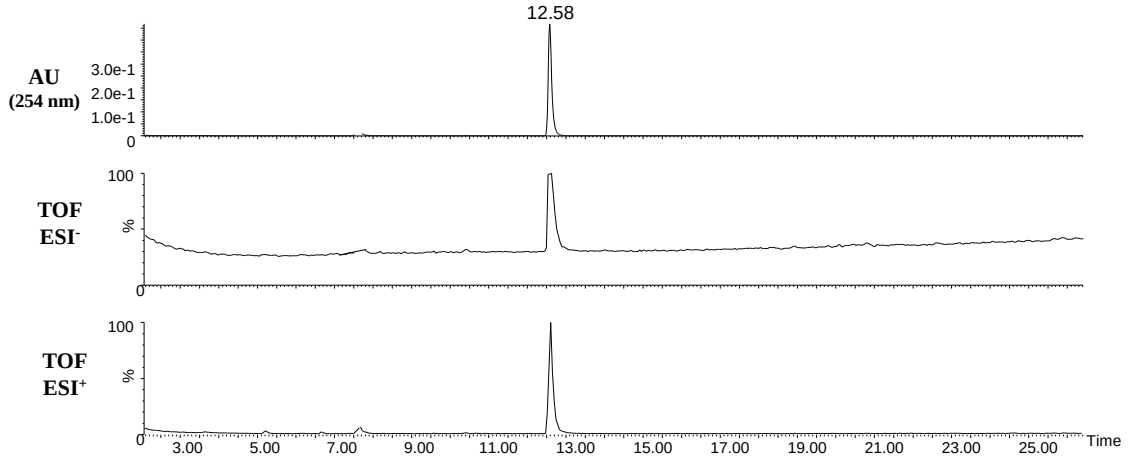


Figure S48. UPLC_ESI_TOF for **1+2c** (0.5 mM each, pH.6.5, 25% DMSO).

<u>Identification of the products:</u>

[1$_2$-2c$_2$]
Retention time: 12.58 min.
Chemical Formula: $C_{68}H_{80}N_{18}O_{26}S_{10}$
Exact Mass: 1884.2698

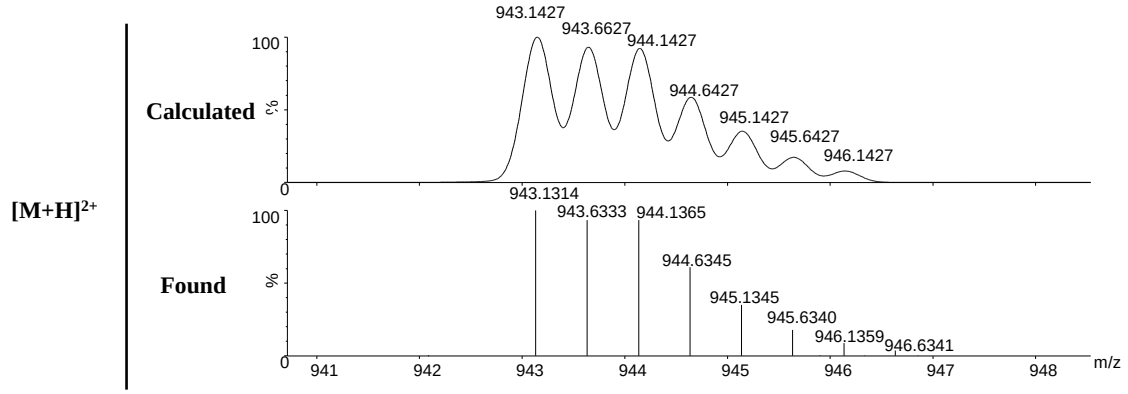


Figure S49. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of **[1_2-2c_2]**.

Mixture of 1 +2c+L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)

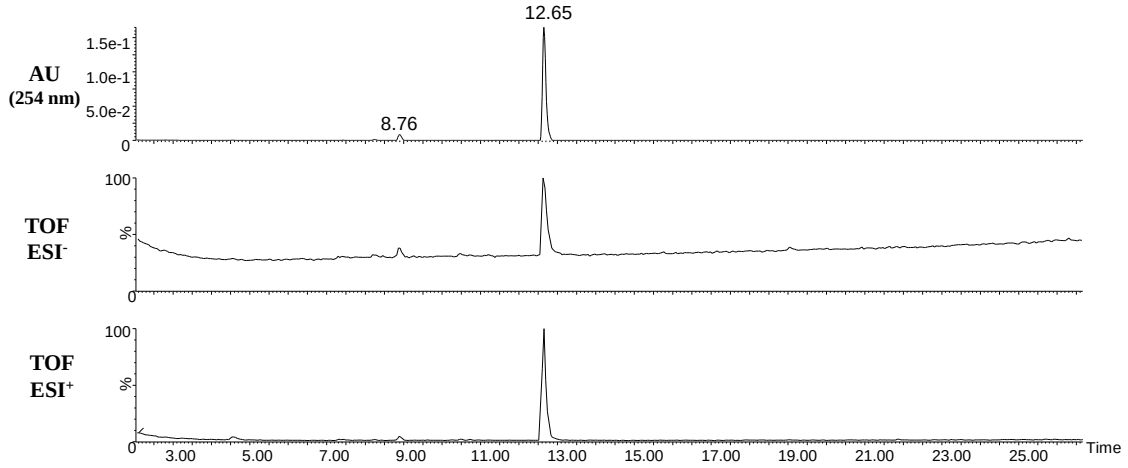


Figure S50. UPLC_ESI_TOF for **1+2c+L-Cys** (0.5 + 0.5 + 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[1-2c-Cys]
Retention time: 8.76 min.
Chemical Formula: $C_{37}H_{46}N_{10}O_{15}S_6$
Exact Mass: 1062.1468

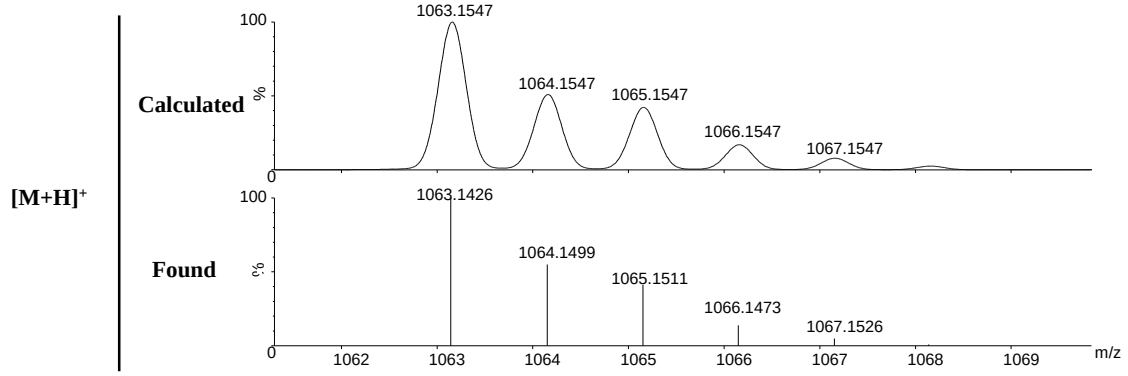


Figure S51. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{+}$ of **[1-2c-Cys**].

[1_2-$2c_2$] *(Previously described)*
Retention time: 12.65 min.
Chemical Formula: $C_{68}H_{80}N_{18}O_{26}S_{10}$
Exact Mass: 1884.2698

Mixture of 1 + 2d (0.5 mM each, 25% DMSO, pH=6.5)

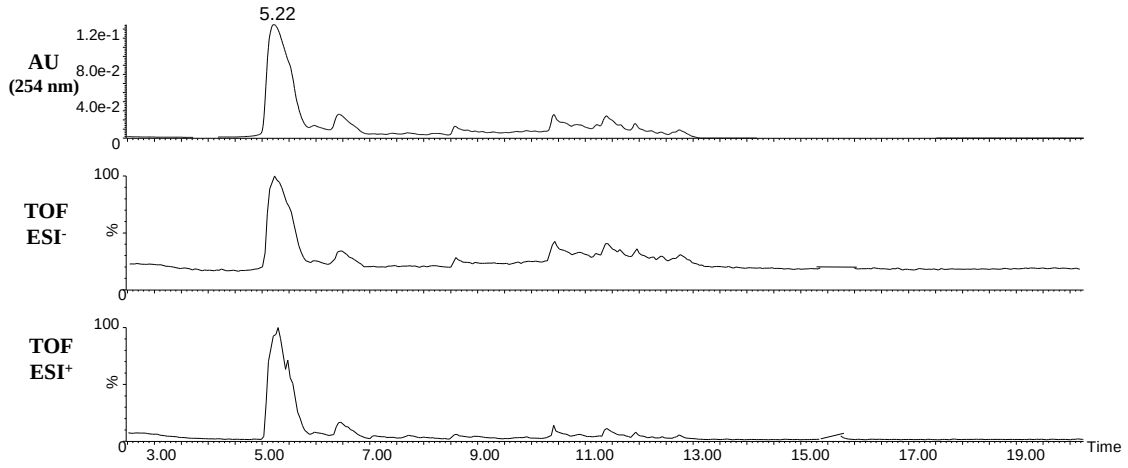


Figure S52. UPLC_ESI_TOF for **1+2d** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[$2d_2$]
Retention time: 5.22 min.
Chemical Formula: $C_{36}H_{52}N_{12}O_8S_4$
Exact Mass: 908.2914

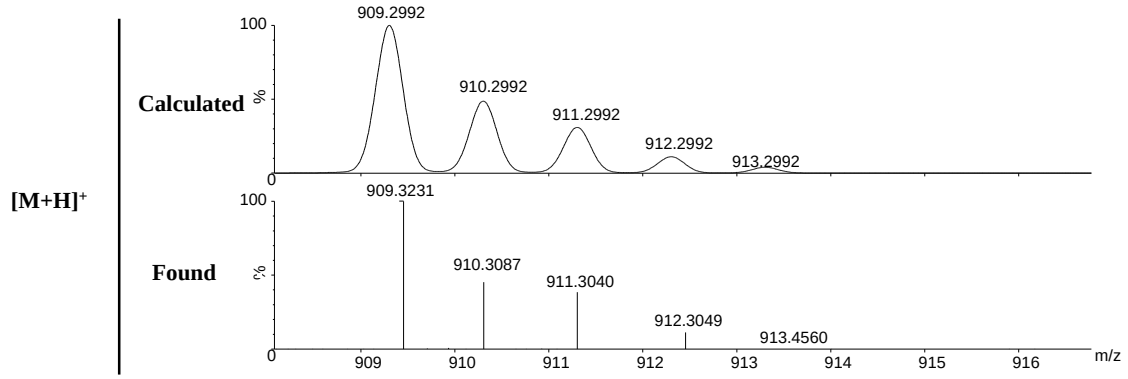


Figure S53. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[$2d_2$]**.

Mixture of 1 +2d+L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)

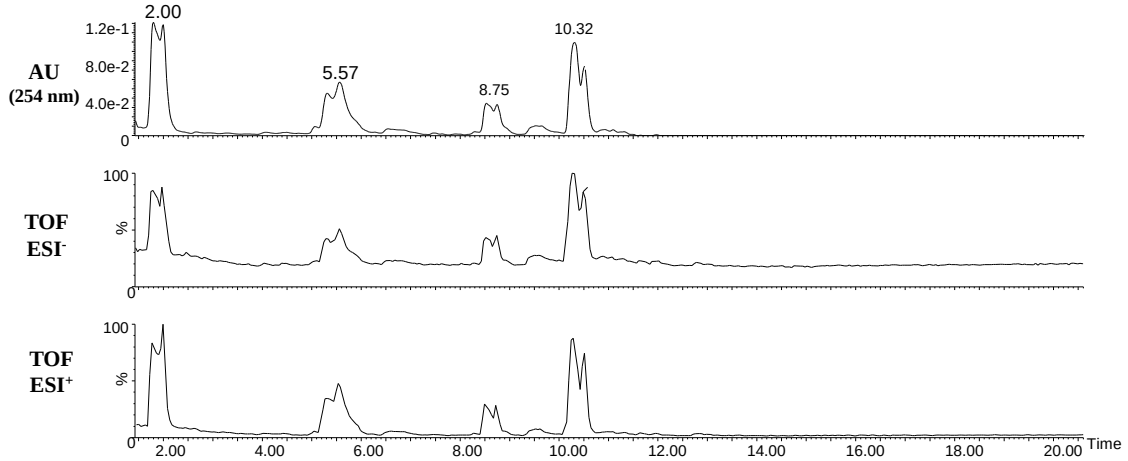


Figure S54. UPLC_ESI_TOF for **1+2d+L-Cys** (0.5 + 0.5 + 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2d-Cys$_2$]
Retention time: 2.00 min.
Chemical Formula: $C_{24}H_{38}N_8O_8S_4$
Exact Mass: 694.1695

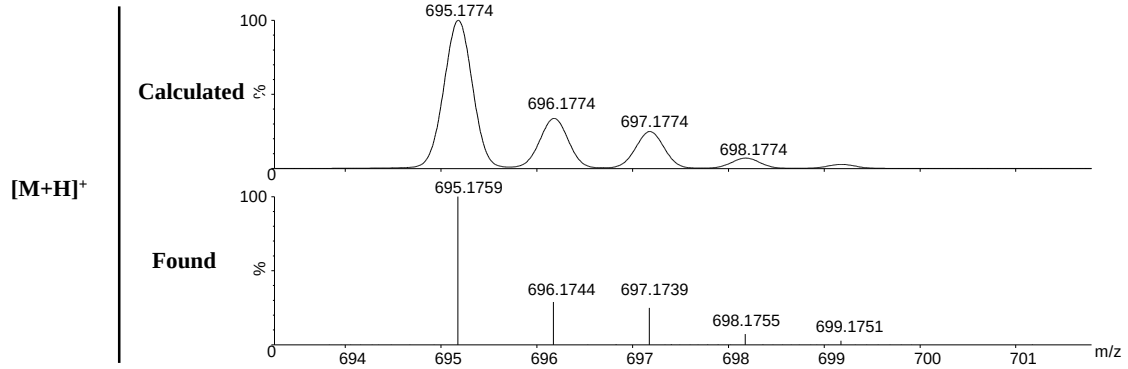


Figure S55. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2d-Cys$_2$]**.

[2d$_2$] *(Previously described)*
Retention time: 5.57 min.
Chemical Formula: $C_{36}H_{52}N_{12}O_8S_4$
Exact Mass: 908.2914

[1-2d$_2$-Cys]
Retention time: 8.75 min.
Chemical Formula: $C_{57}H_{76}N_{16}O_{19}S_8$
Exact Mass: 1544.3238

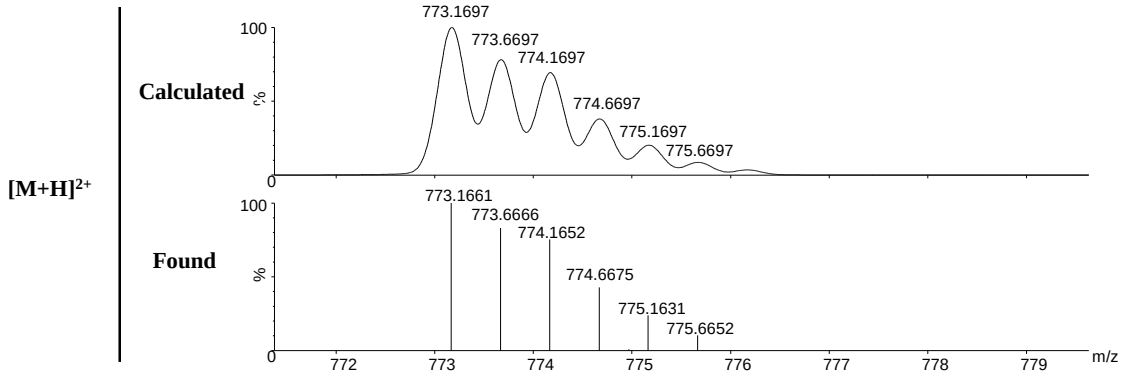


Figure S56. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of **[1-2d$_2$-Cys**].

[1-2d-Cys]
Retention time: 10.32 min.
Chemical Formula: $C_{39}H_{50}N_{10}O_{15}S_6$
Exact Mass: 1090.1781

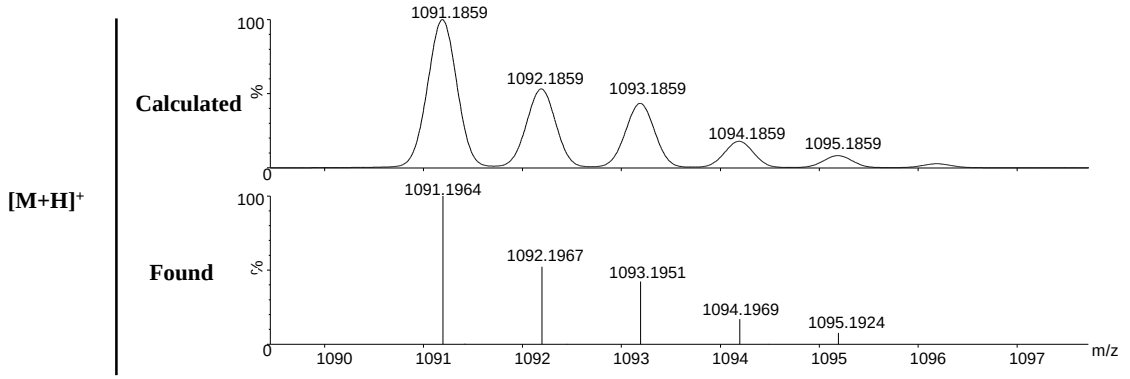


Figure S57. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{+}$ of **[1-2d-Cys**].

<u>Mixture of 1 + 2e (0.5 mM each, 25% DMSO, pH=6.5)</u>

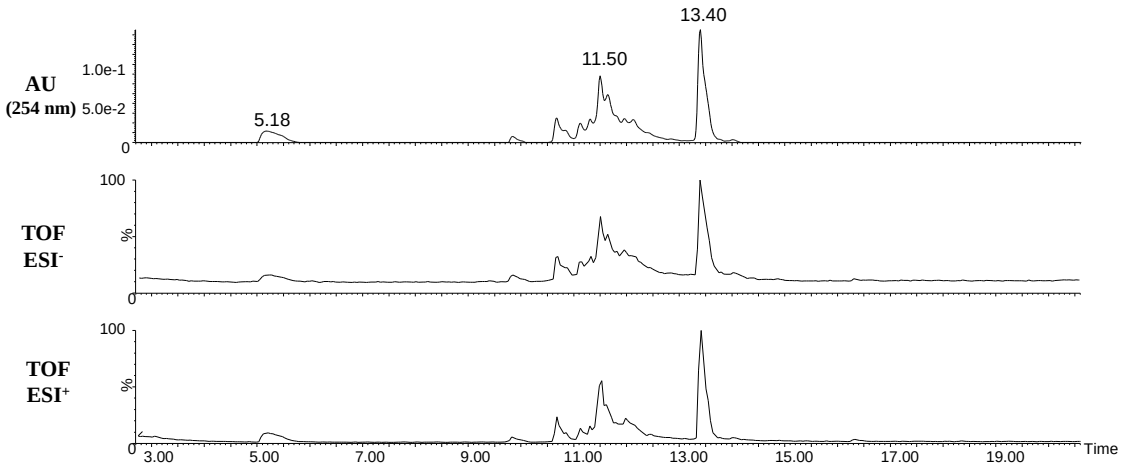


Figure S58. UPLC_ESI_TOF for **1+2e** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2e$_2$]
Retention time: 5.18 min.
Chemical Formula: $C_{40}H_{60}N_{12}O_8S_4$
Exact Mass: 964.3540

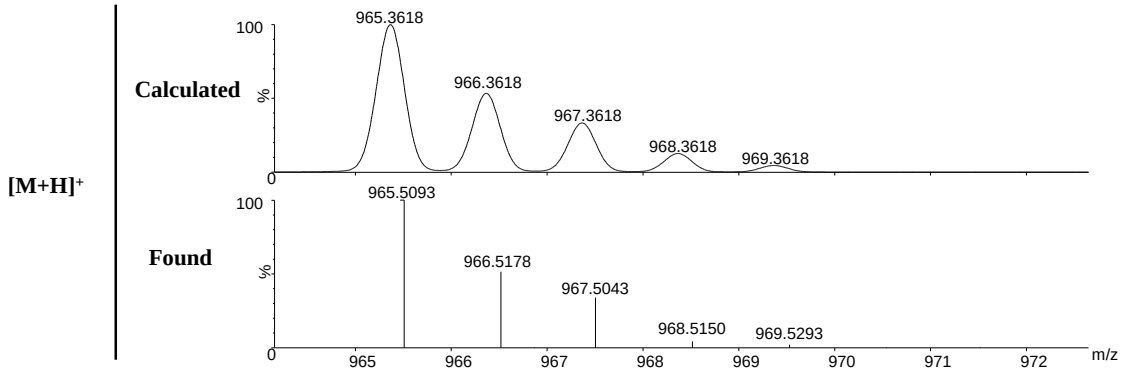


Figure S59. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2e$_2$]**.

[1$_2$-2e]
Retention time: 11.50 min.
Chemical Formula: $C_{56}H_{66}N_{12}O_{22}S_8$
Exact Mass: 1514.2180

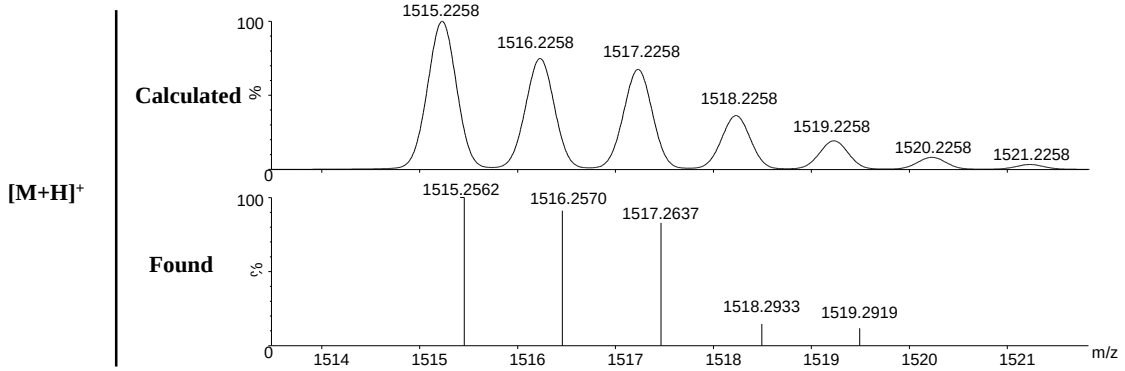


Figure S60. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1$_2$-2e]**.

[1$_2$-2e$_2$]
Retention time: 13.40 min.
Chemical Formula: $C_{76}H_{96}N_{18}O_{26}S_{10}$
Exact Mass: 1996.3950

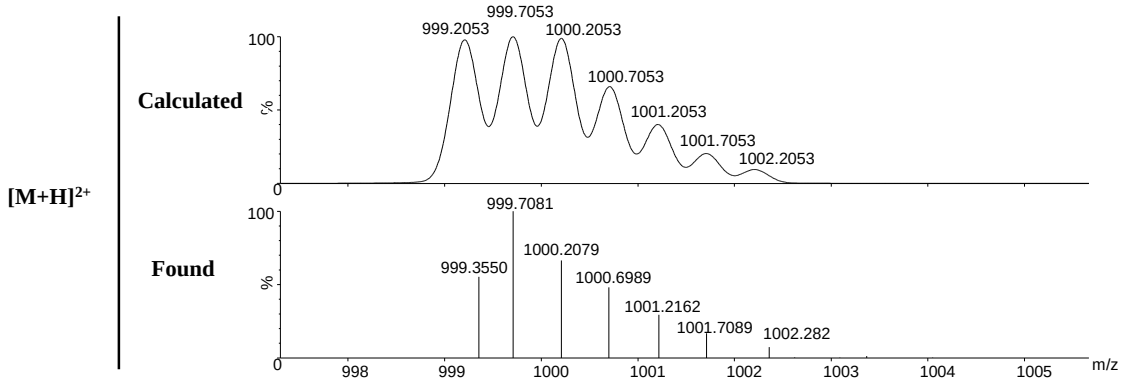


Figure S61. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of **[1_2-2e_2]**.

Mixture of 1 +2e+L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)

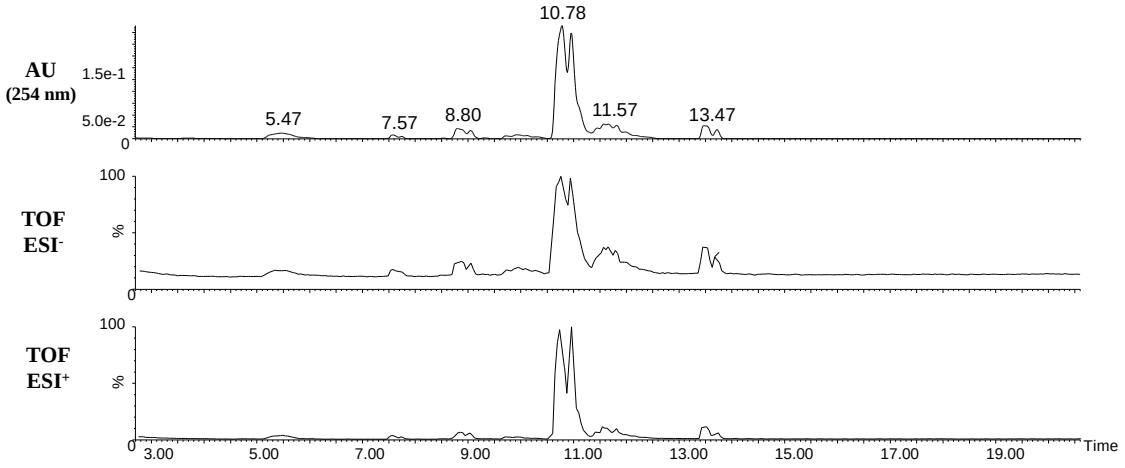


Figure S62. UPLC_ESI_TOF for **1+2e+L-Cys** (0.5 + 0.5 + 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2e_2] *(Previously described)*
Retention time: 5.47 min.
Chemical Formula: $C_{40}H_{60}N_{12}O_8S_4$
Exact Mass: 964.3540

[1-2e-Cys$_3$]
Retention time: 7.57 min.
Chemical Formula: $C_{47}H_{66}N_{12}O_{19}S_8$
Exact Mass: 1358.2333

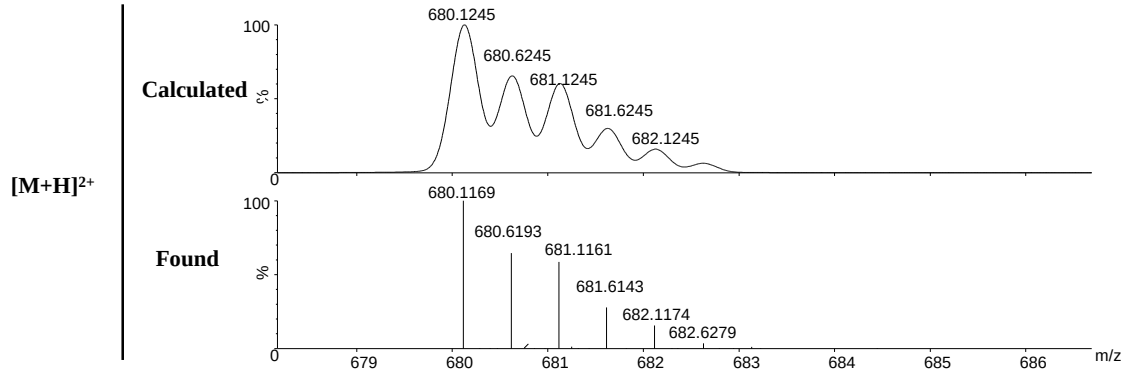


Figure S63. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of **[1-2e-Cys$_3$]**.

[1-2e$_2$-Cys]
Retention time: 8.80 min.
Chemical Formula: $C_{61}H_{84}N_{16}O_{19}S_8$
Exact Mass: 1600.3864

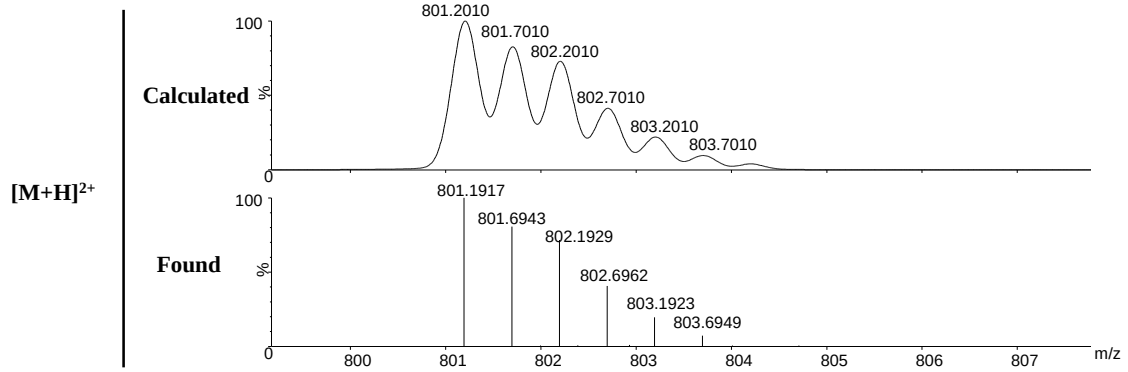


Figure S64. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of **[1-2e$_2$-Cys**].

[1-2e-Cys]
Retention time: 10.78 min.
Chemical Formula: $C_{41}H_{54}N_{10}O_{15}S_6$
Exact Mass: 1118.2094

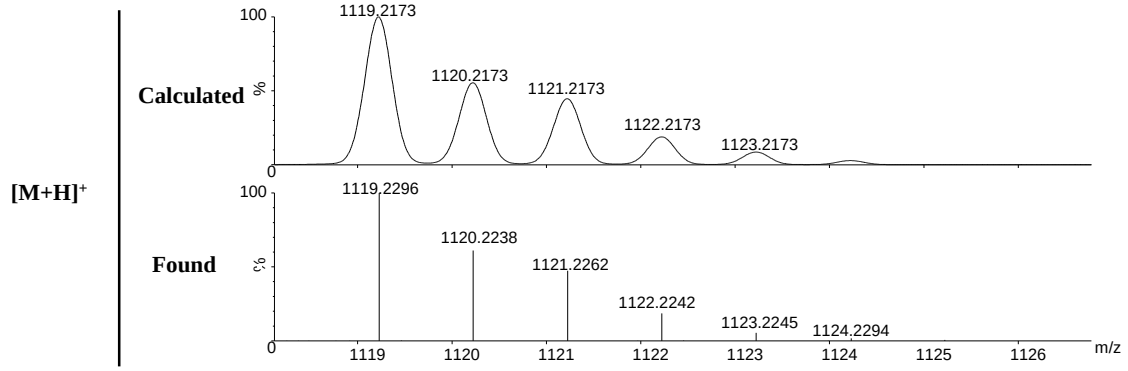


Figure S65. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{+}$ of **[1-2e-Cys**].

[1$_2$-2e] *(Previously described)*
Retention time: 11.57 min.

Chemical Formula: $C_{56}H_{66}N_{12}O_{22}S_8$
Exact Mass: 1514.2180

[1_2-2e_2] *(Previously described)*
Retention time: 13.47 min.
Chemical Formula: $C_{76}H_{96}N_{18}O_{26}S_{10}$
Exact Mass: 1996.3950

Mixture of 1+2c+2d+2e (0.5 mM each, 25% DMSO, pH=6.5)

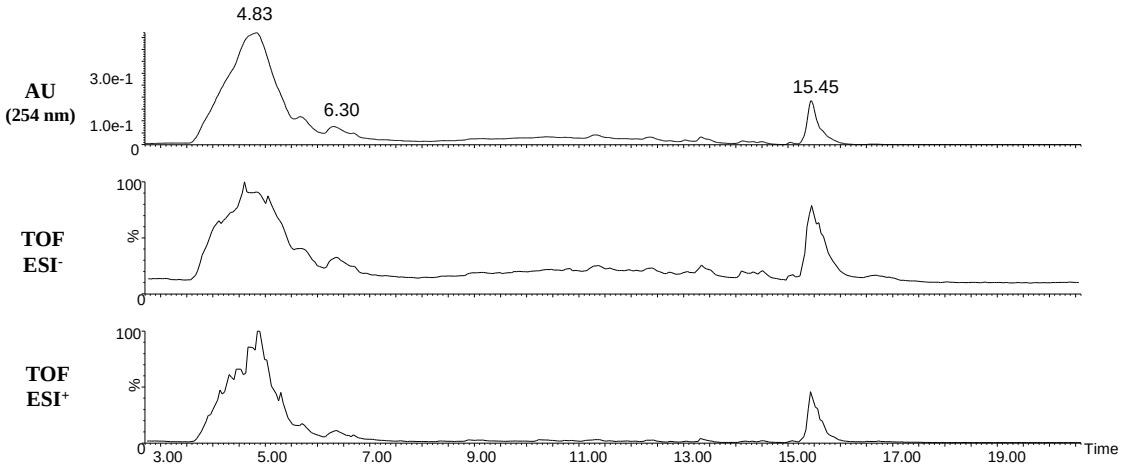


Figure S66. UPLC_ESI_TOF for **1+2c+2d+2e** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2e_2] *(Previously described)*
Retention time: 4.83 min.
Chemical Formula: $C_{40}H_{60}N_{12}O_8S_4$
Exact Mass: 964.3540

[2d_2] *(Previously described)*
Retention time: 4.83 min.
Chemical Formula: $C_{36}H_{52}N_{12}O_8S_4$
Exact Mass: 908.2914

[2d-2e]
Retention time: 6.30 min.
Chemical Formula: $C_{38}H_{56}N_{12}O_8S_4$
Exact Mass: 936.3227

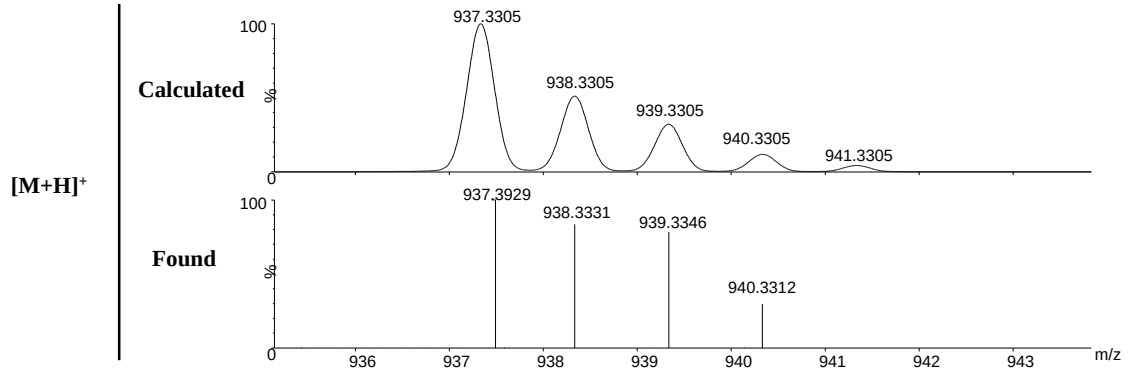


Figure S67. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2d-2e]**.

[1_2-$2c_2$] *(Previously described)*
Retention time: 15.45 min.
Chemical Formula: $C_{68}H_{80}N_{18}O_{26}S_{10}$
Exact Mass: 1884.2698

Mixture of 1+2c+2d+2e+ L-Cys(0.5 + 0.5+ 0.5+ 0.5+ 2.5 mM, 25% DMSO, pH=6.5)

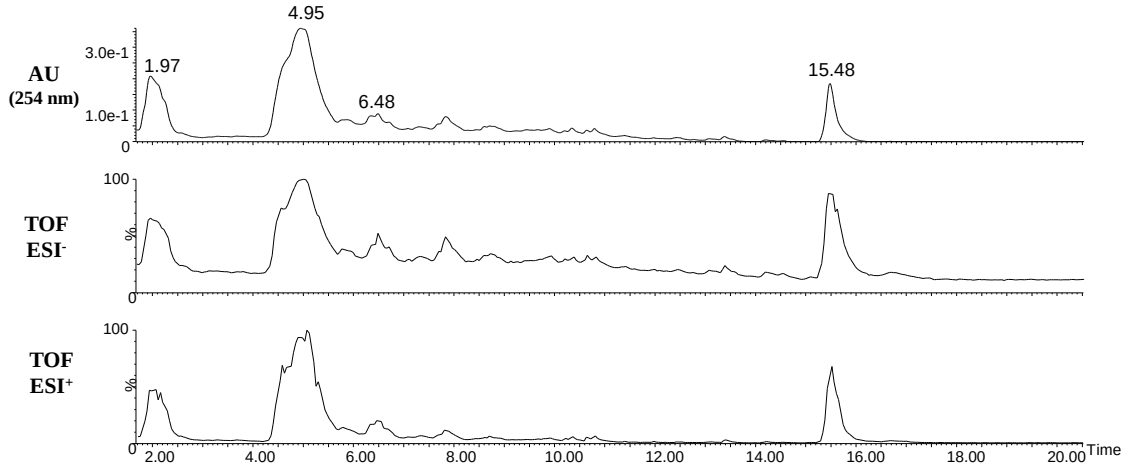


Figure S68. UPLC_ESI_TOF for **1+2c+2d+2e+L-Cys** (0.5 + 0.5+ 0.5+ 0.5+ 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2d-Cys_2] *(Previously described)*
Retention time: 1.97 min.
Chemical Formula: $C_{24}H_{38}N_8O_8S_4$
Exact Mass: 694.1695

[2e-Cys_2]
Retention time: 1.97 min.
Chemical Formula: $C_{26}H_{42}N_8O_8S_4$
Exact Mass: 722.2008

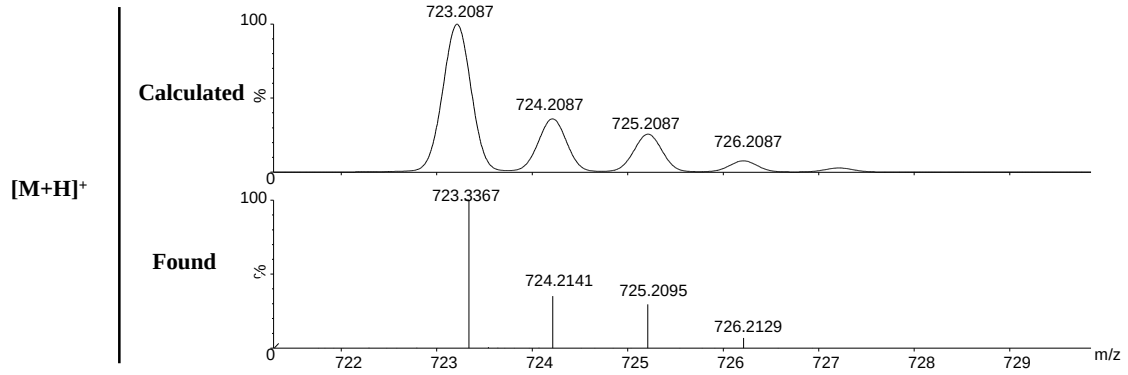


Figure S69. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2e-Cys$_2$]**.

[2e$_2$] *(Previously described)*
Retention time: 4.95 min.
Chemical Formula: $C_{40}H_{60}N_{12}O_8S_4$
Exact Mass: 964.3540

[2d$_2$] *(Previously described)*
Retention time: 4.95 min.
Chemical Formula: $C_{36}H_{52}N_{12}O_8S_4$
Exact Mass: 908.2914

[2d-2e] *(Previously described)*
Retention time: 6.48 min.
Chemical Formula: $C_{38}H_{56}N_{12}O_8S_4$
Exact Mass: 936.3227

[1$_2$-2c$_2$] *(Previously described)*
Retention time: 15.48 min.
Chemical Formula: $C_{68}H_{80}N_{18}O_{26}S_{10}$
Exact Mass: 1884.2698

<u>Mixture of 2c+3 (0.5 mM each, 25% DMSO, pH=6.5)</u>

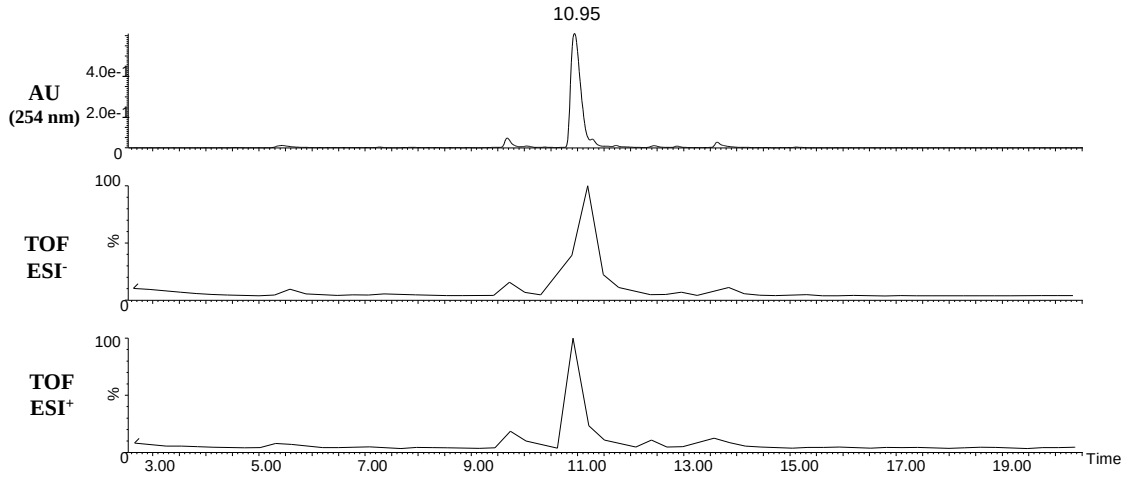


Figure S70. UPLC_ESI_TOF for **2c+3** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2c-3]
Retention time: 10.95 min.
Chemical Formula: $C_{30}H_{36}N_8O_{10}S_4$
Exact Mass: 796.1437

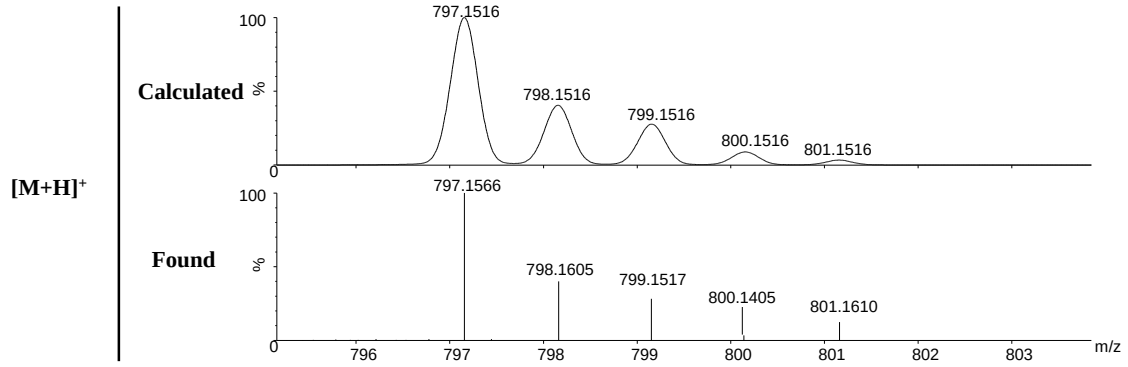


Figure S71. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2c-3]**.

Mixture of 2c+ 3 +L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)

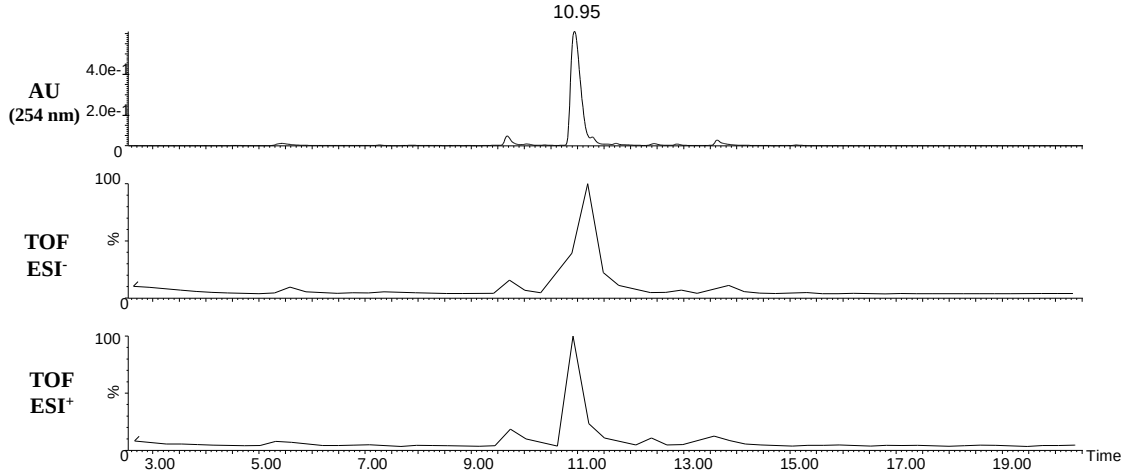


Figure S72. UPLC_ESI_TOF for **2c+ 3 +L-Cys** (0.5 + 0.5 + 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2c-3] *(Previously described)*
Retention time: 10.95 min.
Chemical Formula: $C_{30}H_{36}N_8O_{10}S_4$
Exact Mass: 796.1437

Mixture of 2d+3 (0.5 mM each, 25% DMSO, pH=6.5)

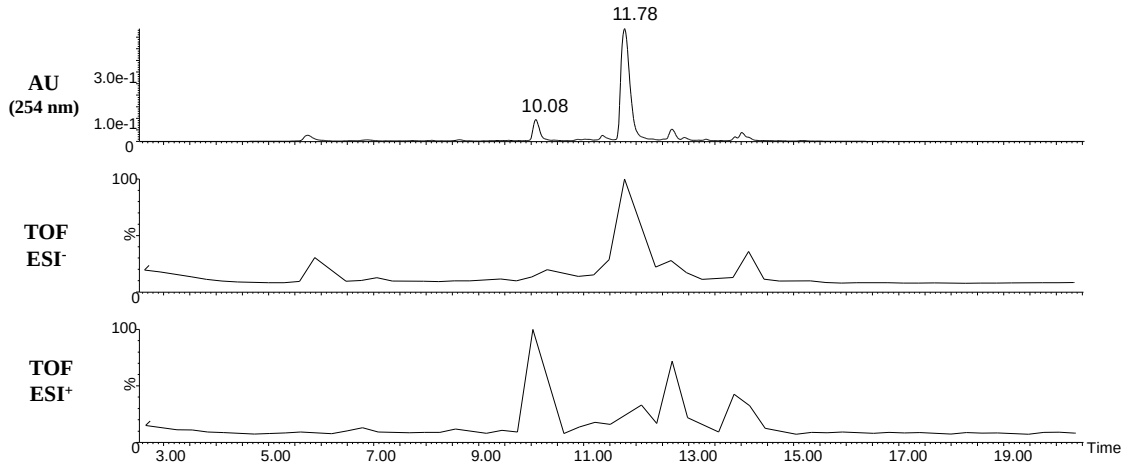


Figure S73. UPLC_ESI_TOF for **2d+3** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2d$_2$-3]
Retention time: 10.08 min.
Chemical Formula: $C_{50}H_{66}N_{14}O_{14}S_6$
Exact Mass: 1278.3207

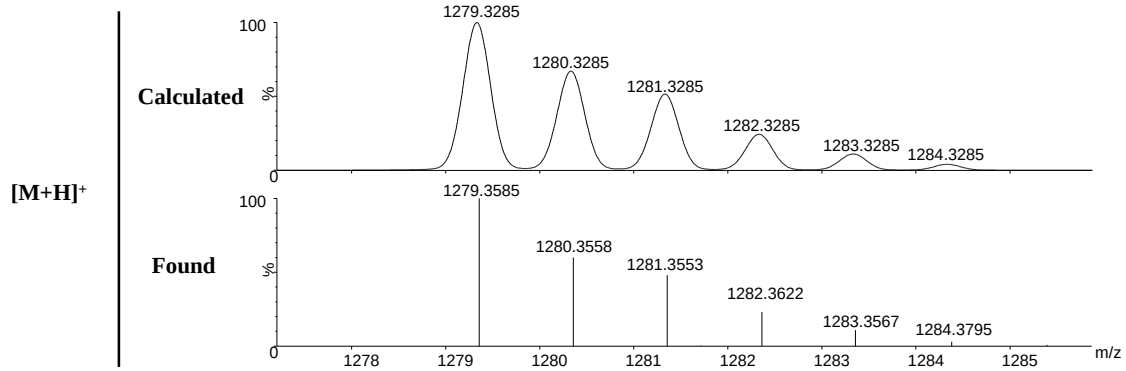


Figure S74. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2d$_2$-3**].

[2d-3]
Retention time: 11.78 min.
Chemical Formula: $C_{32}H_{40}N_8O_{10}S_4$
Exact Mass: 824.1750

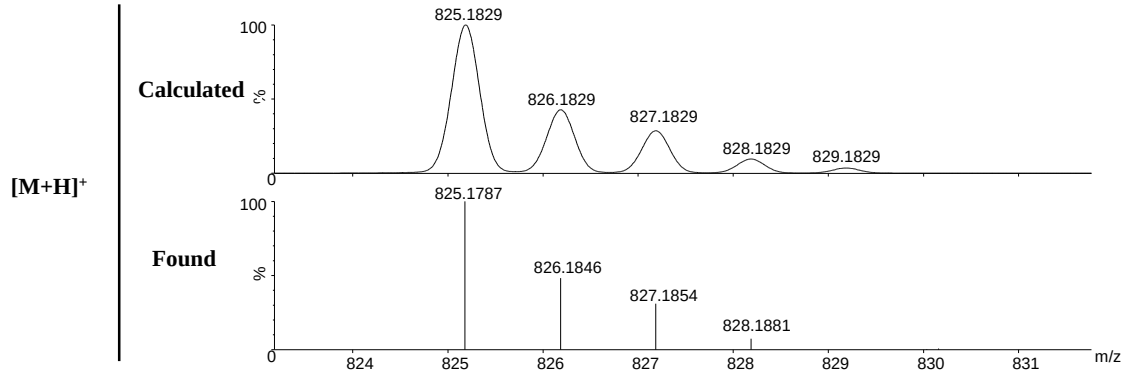


Figure S75. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2d-3**].

Mixture of 2d+ 3 +L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)

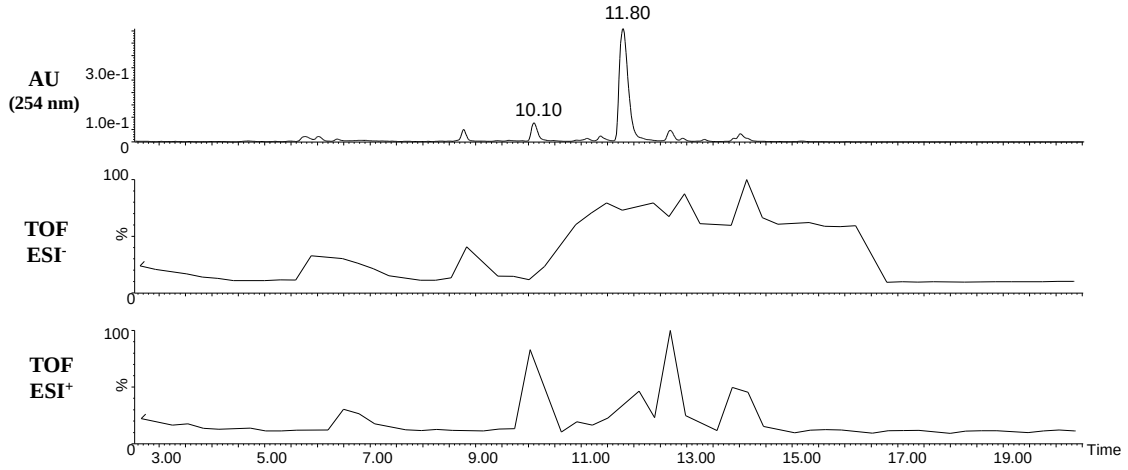


Figure S76. UPLC_ESI_TOF for **2d+ 3 +L-Cys** (0.5 + 0.5 + 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2d$_2$-3] *(Previously described)*
Retention time: 10.10 min.
Chemical Formula: $C_{50}H_{66}N_{14}O_{14}S_6$
Exact Mass: 1278.3207

[2d-3] *(Previously described)*
Retention time: 11.80 min.
Chemical Formula: $C_{32}H_{40}N_8O_{10}S_4$
Exact Mass: 824.1750

Mixture of 2e+3 (0.5 mM each, 25% DMSO, pH=6.5)

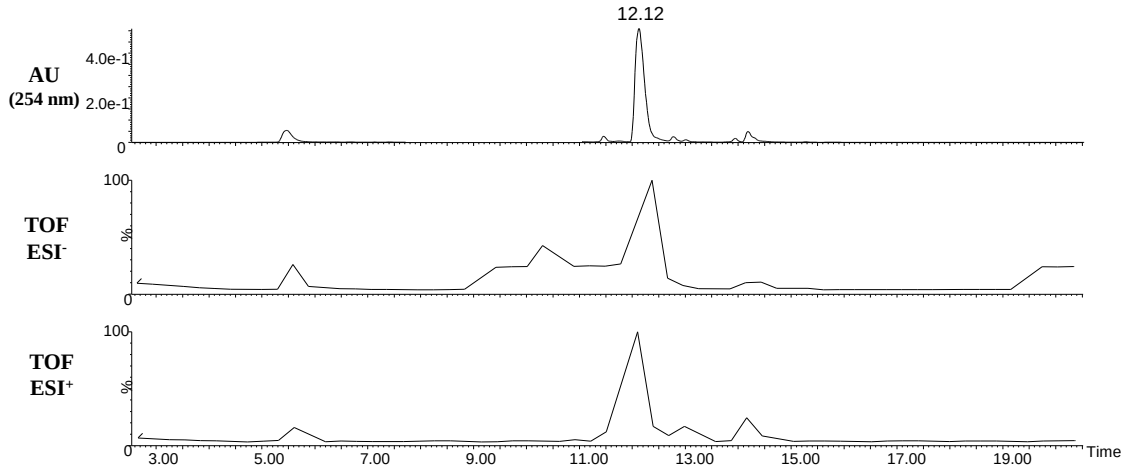


Figure S77. UPLC_ESI_TOF for **2e+3** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2e-3]
Retention time: 12.12 min.
Chemical Formula: $C_{34}H_{44}N_8O_{10}S_4$
Exact Mass: 852.2063

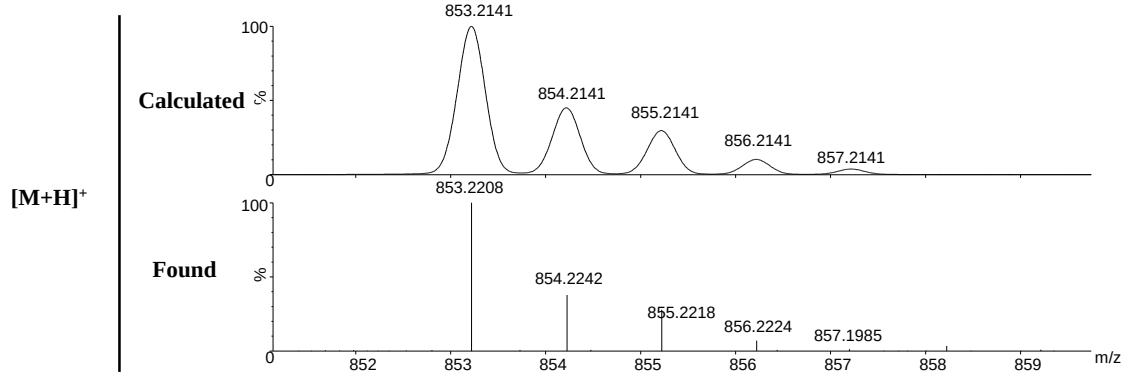


Figure S78. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2e-3**].

Mixture of 2e+ 3 +L-Cys (0.5 + 0.5 + 2.5 mM, 25% DMSO, pH=6.5)

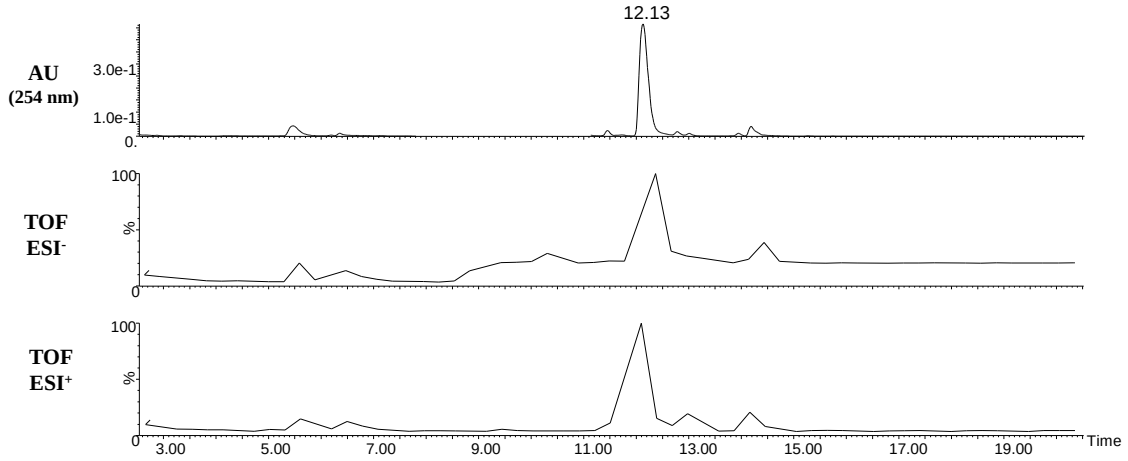


Figure S79. UPLC_ESI_TOF for **2e+3+Cys** (0.5 + 0.5+ 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2e-3] *(Previously described)*
Retention time: 12.13 min.
Chemical Formula: $C_{34}H_{44}N_8O_{10}S_4$
Exact Mass: 852.2063

Mixture of 2c+2d+2e+3 (0.5 mM each, 25% DMSO, pH=6.5)

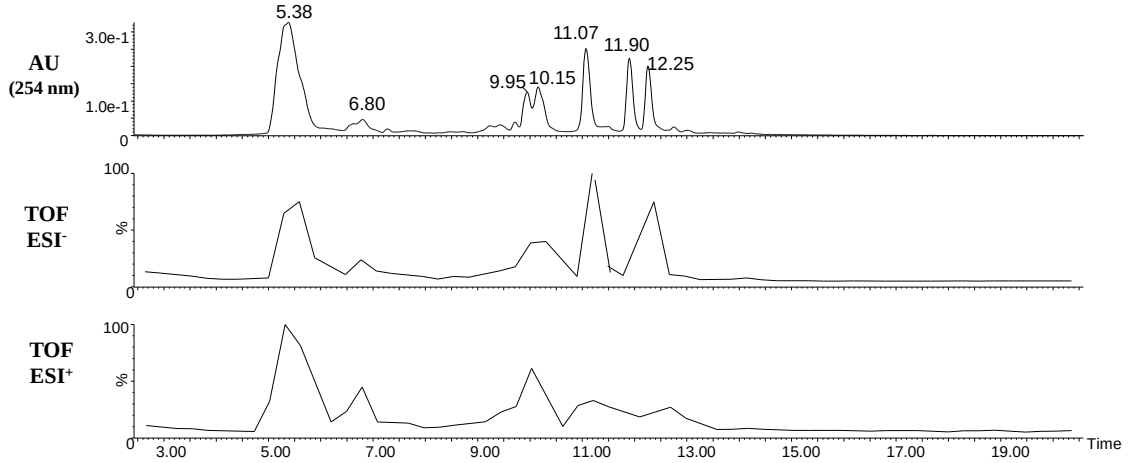


Figure S80. UPLC_ESI_TOF for **2c+2d+2e+3** (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2e$_2$] *(Previously described)*
Retention time: 5.38 min.
Chemical Formula: $C_{40}H_{60}N_{12}O_8S_4$
Exact Mass: 964.3540

[2d$_2$] *(Previously described)*
Retention time: 5.38 min.
Chemical Formula: $C_{36}H_{52}N_{12}O_8S_4$
Exact Mass: 908.2914

[2d-2e] *(Previously described)*
Retention time: 6.80 min.
Chemical Formula: $C_{38}H_{56}N_{12}O_8S_4$
Exact Mass: 936.3227

[2c$_3$]
Retention time: 9.95 min.
Chemical Formula: $C_{48}H_{66}N_{18}O_{12}S_6$
Exact Mass: 1278.3432

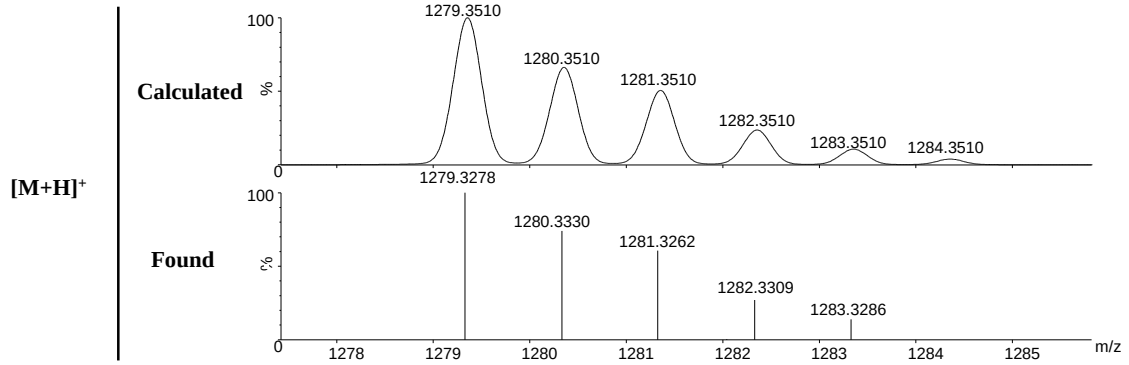


Figure S81. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2c$_3$]**.

[2c-2d$_2$]
Retention time: 10.15 min.
Chemical Formula: $C_{52}H_{74}N_{18}O_{12}S_6$
Exact Mass: 1334.4058

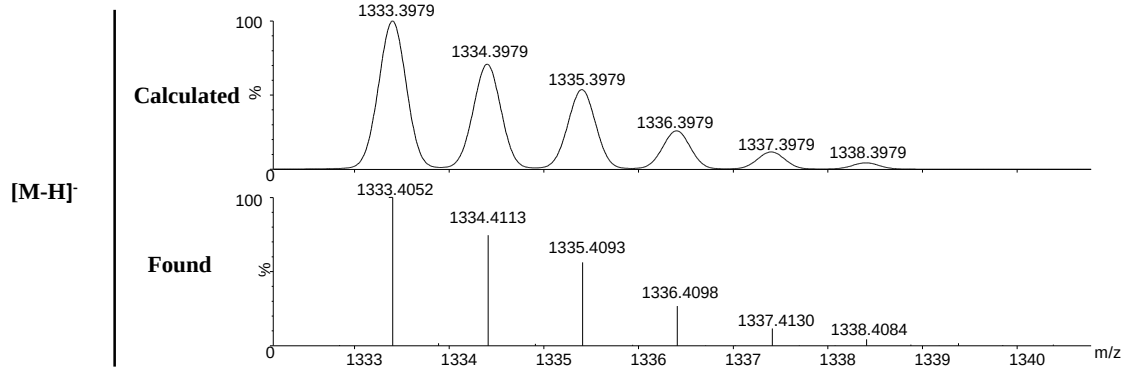


Figure S82. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M-H]^-$ of **[2c-2d$_2$]**.

[2c-3] *(Previously described)*
Retention time: 11.07 min.
Chemical Formula: $C_{30}H_{36}N_8O_{10}S_4$
Exact Mass: 796.1437

[2d-3] *(Previously described)*
Retention time: 11.90 min.
Chemical Formula: $C_{32}H_{40}N_8O_{10}S_4$
Exact Mass: 824.1750

[2e-3] *(Previously described)*
Retention time: 12.25 min.
Chemical Formula: $C_{34}H_{44}N_8O_{10}S_4$
Exact Mass: 852.2063

Mixture of 2c+2d+2e+3+ Cys (0.5+ 0.5+ 0.5+ 0.5+ 2.5 mM, 25% DMSO, pH=6.5)

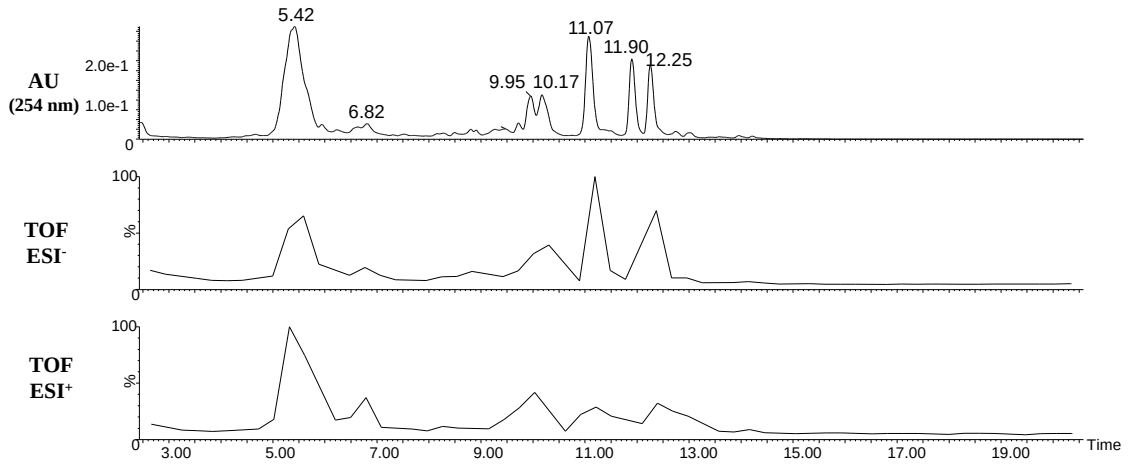


Figure S83. UPLC_ESI_TOF for **2c+2d+2e+3+Cys** (0.5+ 0.5+ 0.5+ 0.5+ 2.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2e$_2$] *(Previously described)*
Retention time: 5.42 min.
Chemical Formula: $C_{40}H_{60}N_{12}O_8S_4$
Exact Mass: 964.3540

[2d$_2$] *(Previously described)*
Retention time: 5.42 min.
Chemical Formula: $C_{36}H_{52}N_{12}O_8S_4$
Exact Mass: 908.2914

[2d-2e] *(Previously described)*
Retention time: 6.82 min.
Chemical Formula: $C_{38}H_{56}N_{12}O_8S_4$
Exact Mass: 936.3227

[2c$_3$] *(Previously described)*
Retention time: 9.95 min.
Chemical Formula: $C_{48}H_{66}N_{18}O_{12}S_6$
Exact Mass: 1278.3432

[2c-2d$_2$] *(Previously described)*
Retention time: 10.17 min.
Chemical Formula: $C_{52}H_{74}N_{18}O_{12}S_6$
Exact Mass: 1334.4058

[2c-3] *(Previously described)*
Retention time: 11.07 min.
Chemical Formula: $C_{30}H_{36}N_8O_{10}S_4$
Exact Mass: 796.1437

[2d-3] *(Previously described)*
Retention time: 11.90 min.
Chemical Formula: $C_{32}H_{40}N_8O_{10}S_4$
Exact Mass: 824.1750

[2e-3] *(Previously described)*
Retention time: 12.25 min.
Chemical Formula: $C_{34}H_{44}N_8O_{10}S_4$
Exact Mass: 852.2063

Mixture of 1+ 2c+ Cysteamine (CyA) (0.5 + 0.5 + 25 mM, 25% DMSO, pH=6.5)

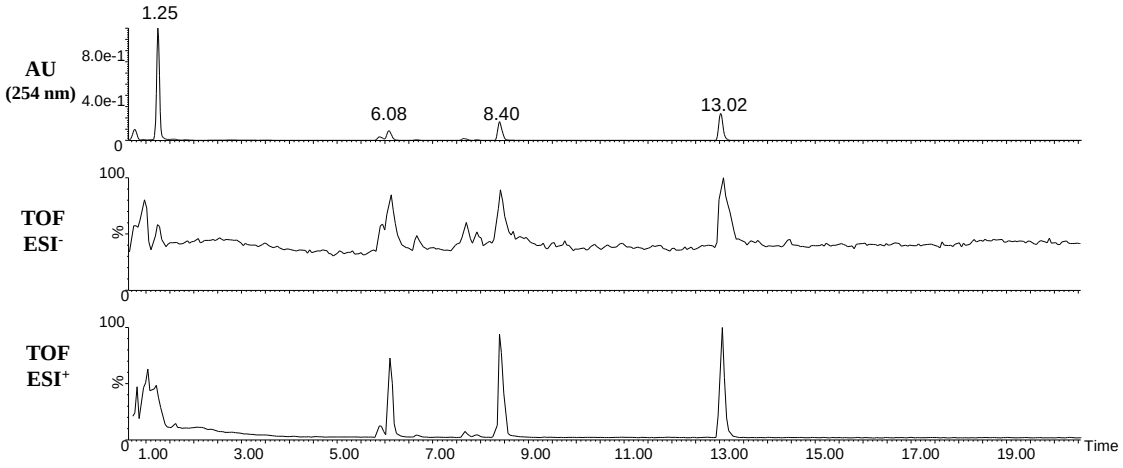


Figure S84. UPLC_ESI_TOF for **1+2c+CyA** (0.5+ 0.5+ 25 mM, pH.6.5, 25% DMSO).

Identification of the products:

[2c- CyA$_2$]
Retention time: 1.25 min.
Chemical Formula: $C_{20}H_{34}N_8O_4S_4$
Exact Mass: 578.1586

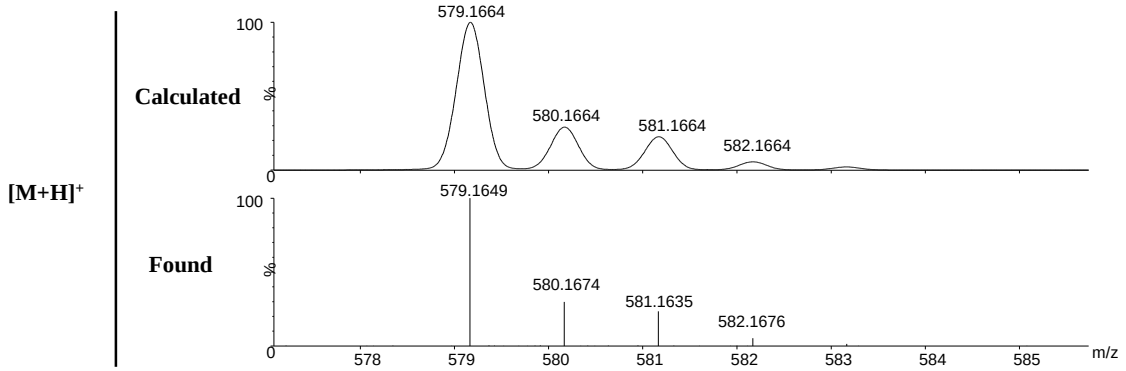


Figure S85. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2c-CyA$_2$]**.

[1-CyA$_3$]
Retention time: 6.08 min.
Chemical Formula: $C_{24}H_{36}N_6O_9S_6$
Exact Mass: 744.0868

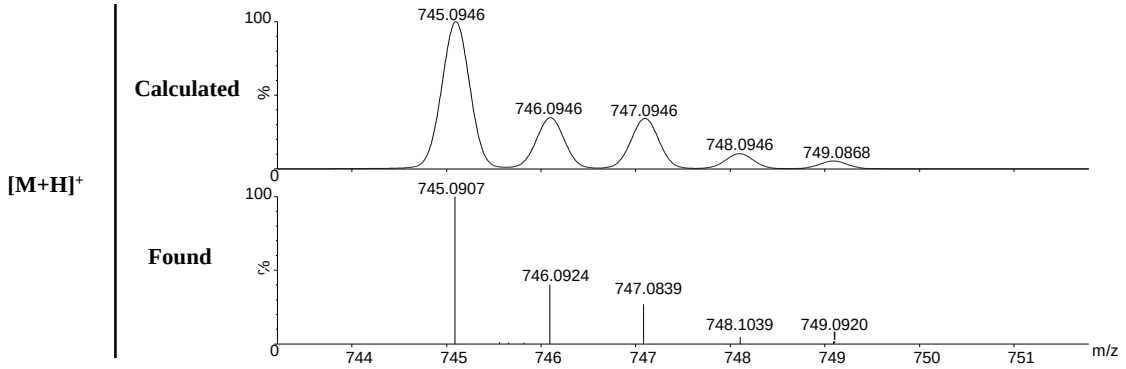


Figure S86. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-CyA$_3$]**.

[1-2c-CyA]
Retention time: 8.40 min.
Chemical Formula: $C_{36}H_{46}N_{10}O_{13}S_6$
Exact Mass: 1018.1570

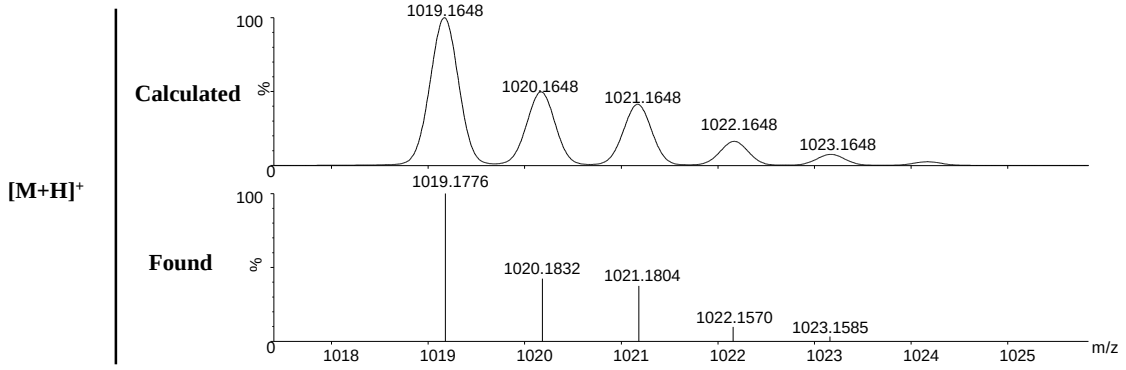


Figure S87. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2c-CyA]**.

[1_2-$2c_2$] *(Previously described)*
Retention time: 13.02 min.
Chemical Formula: $C_{68}H_{80}N_{18}O_{26}S_{10}$
Exact Mass: 1884.2698

Mixture of 1+ 2c at *t = 0 min* (0.5 mM each, 25% DMSO, pH=6.5)

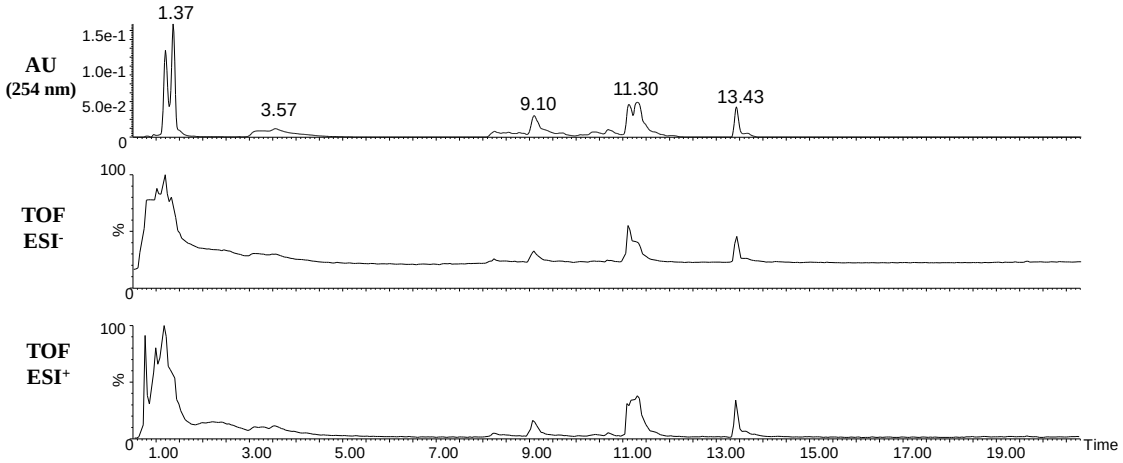


Figure S88. UPLC_ESI_TOF for **1+2c** at *t = 0 min* (0.5 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[$2c^{2-}$]
Retention time: 1.37 min.
Chemical Formula: $C_{16}H_{22}N_6O_4S_2$
Exact Mass: 426.1144

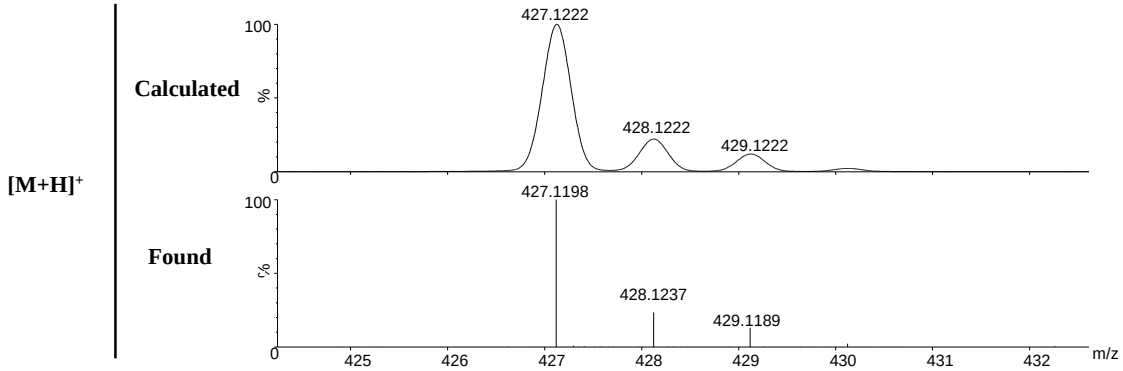


Figure S89. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[$2c^{2-}$]**.

[2c$_2$]
Retention time: 3.57 min.
Chemical Formula: $C_{32}H_{44}N_{12}O_8S_4$
Exact Mass: 852.2288

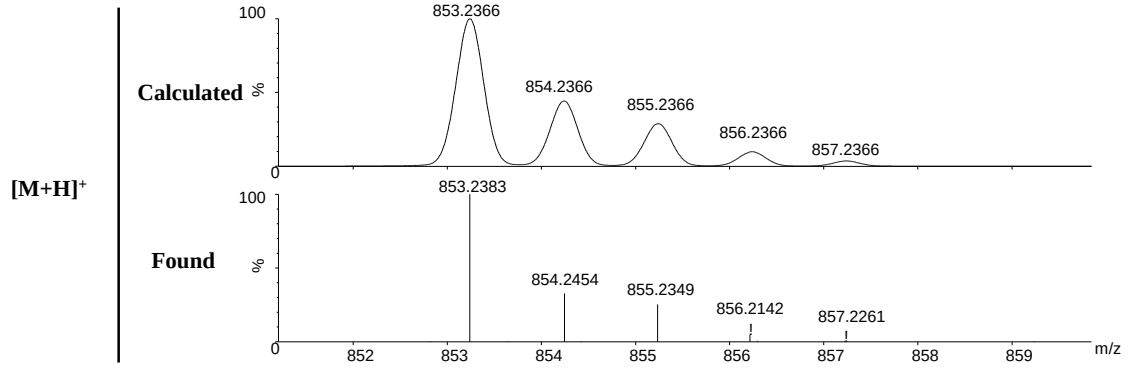


Figure S90. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2c$_2$]**.

[1$_2$-2c]
Retention time: 9.10 min.
Chemical Formula: $C_{52}H_{58}N_{12}O_{22}S_8$
Exact Mass: 1458.1554

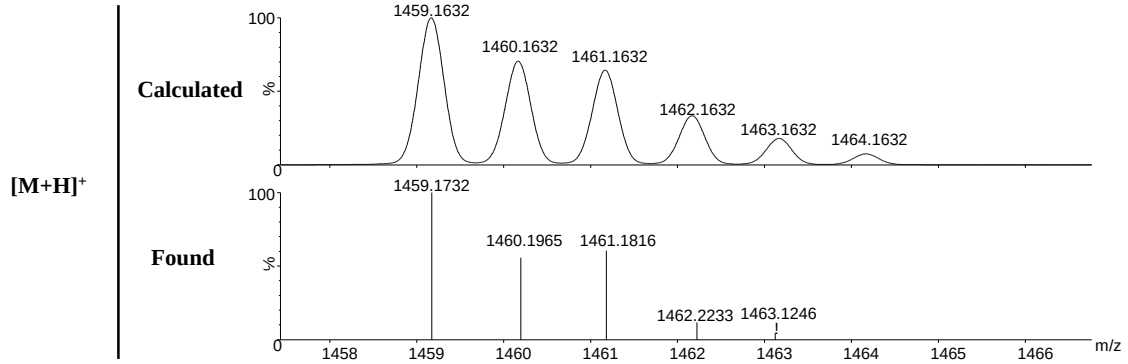


Figure S91. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1$_2$-2c]**.

[1-2c$^-$]
Retention time: 11.30 min.
Chemical Formula: $C_{34}H_{41}N_9O_{13}S_5$
Exact Mass: 943.1427

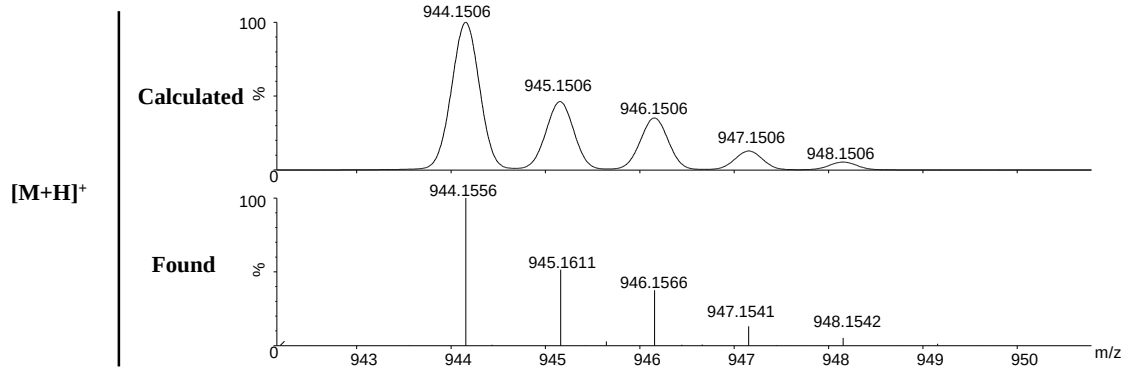
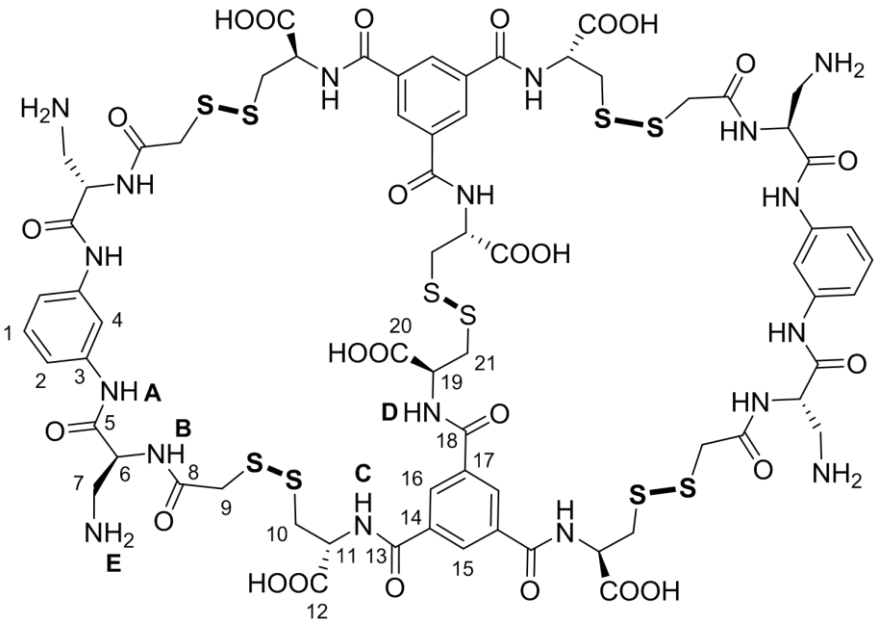


Figure S92. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2c$^-$].**

[1_2-2c_2] *(Previously described)*
Retention time: 13.43 min.
Chemical Formula: $C_{68}H_{80}N_{18}O_{26}S_{10}$
Exact Mass: 1884.2698

SYNTHESIS AND CHARACTERIZATION OF $\mathbf{1_2\text{-}2c_2}$ (**Ib**)



A 15 mM phosphate buffer was prepared by dissolving 0.091 g of Na_2HPO_4 and 0.136 g of NaH_2PO_4 in 100 mL of milli-Q water; the pH was adjusted to 6.5 by the addition of HCl (aq).
1 (23.38 mg, 0.045 mmol) and **2c** (29.55 mg, 0.045 mmol) were dissolved in 1 mL of DMSO; 900 µL of this mixture were added over 40 mL of 15 phosphate buffer. After 10 days under stirring at room temperature, complete oxidation of the starting reagents was observed by HPLC. Once completed, the aqueous solution was removed by lyophilization. The resulting white solid was purified by reversed-face flash chromatography (C_{18} column, gradient: from 3% to 15% CH_3CN in H_2O) and 19.5 mg of pure $\mathbf{1_2\text{-}2c_2}$ (46% yield) were obtained as a white solid.

NMR characterization of 1_2-$2c_2$

HRMS (ESI$^+$) calcd. for $C_{68}H_{80}N_{18}O_{26}S_{10}$ $[M+H]^{2+}$ (m/z): 943.1427, found: 943.1386.
^{1}H NMR (500 MHz, DMSO-d_6) δ = 10.23 (s, 4H, H_A), 9.16 (s, 4H, H_C), 9.09 (s, 2H, H_D), 8.68 (s, 4H, H_B), 8.54 (s, 2H, H_{15}), 8.52 (s, 4H, H_{16}), 7.94 (s, 2H, H_4 overlapped with the amine signal), 7.90 (s, 12H, H_D), 7.36 (m, 4H, H_2), 7.24 (t, J = 8.1 Hz, 2H, H_1), 4.79 (t, 4H, H_{11}), 4.76 (t, 2H, H_{19}), 4.69 (q, J = 7.6 Hz, 4H, H_6), 3.62 (s, 8H, H_9), 3.26 (dd, J = 12.7, 4.7 Hz, 8H, H_7), 3.20 (t, J = 11.6 Hz, 8H, H_{10}), 3.06 (m, 4H, H_{21}) ppm.
^{13}C NMR (125 MHz, DMSO-d_6): δ =169.4 (CO, C_8), 167.5 (CO, C_5), 166.0 (CO, C_{18}, C_{13}), 139.0 (C, C_3), 129.7 (C, C_{14}), 129.1 (CH, C_{16}, C_{15}), 128.8 (CH, C_1), 115.1 (CH, C_2), 111.3 (CH, C_4), 52.0 (CH, C_{11}, C_{19}), 51.3 (CH, C_6), 42.4 (CH_2, C_9), 39.6 (CH_2, C_7), 39.6 (CH_2, C_{21}), 39.0 (CH_2, C_{10}) ppm.

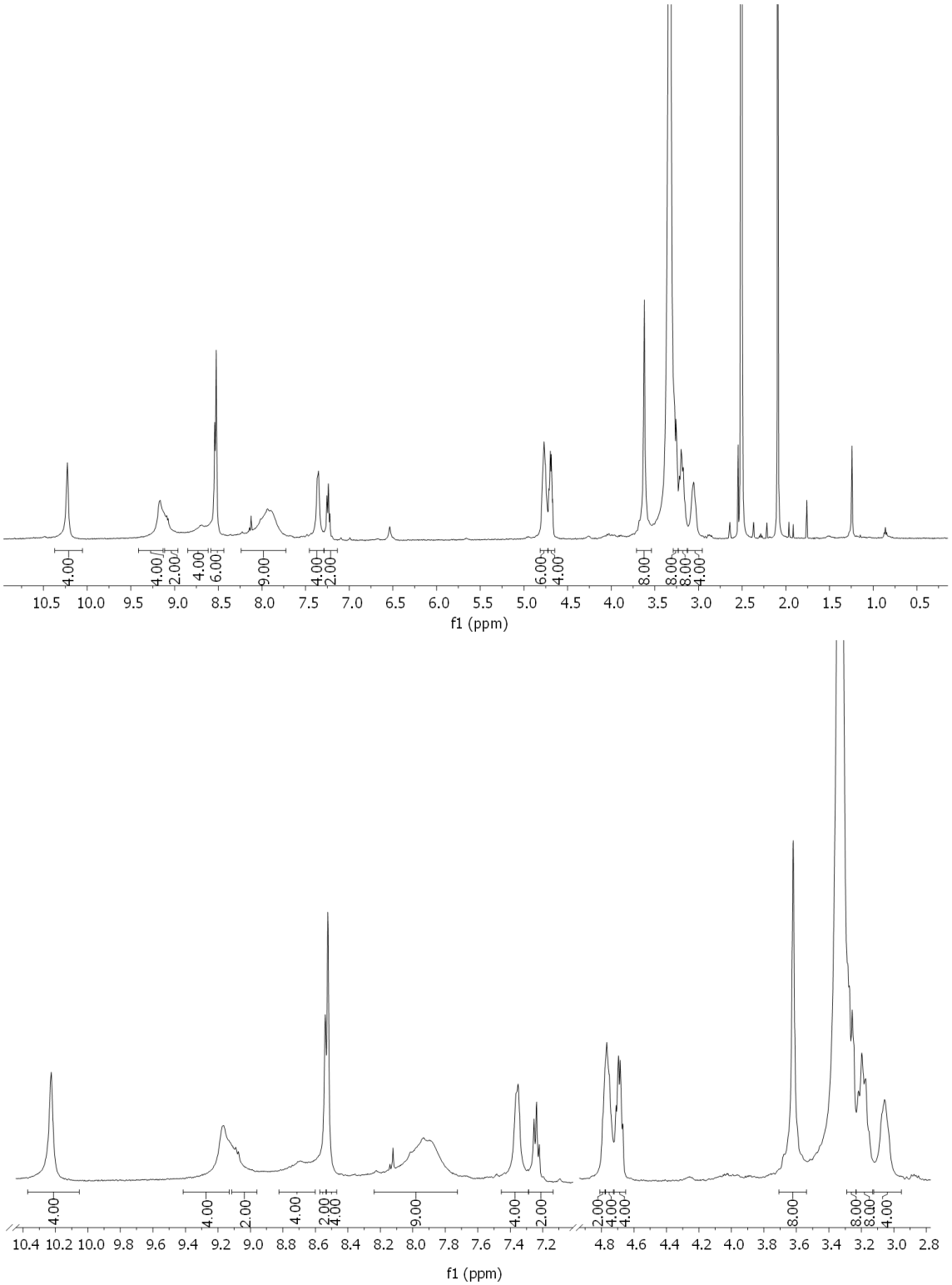


Figure S93. ^{1}H (500 MHz, 298 K in DMSO-d_6) spectrum of **1_2-$2c_2$**, and expansion of the aliphatic region (4.8 – 2.8 ppm) and of the amide region (10.4 – 7.2 ppm).

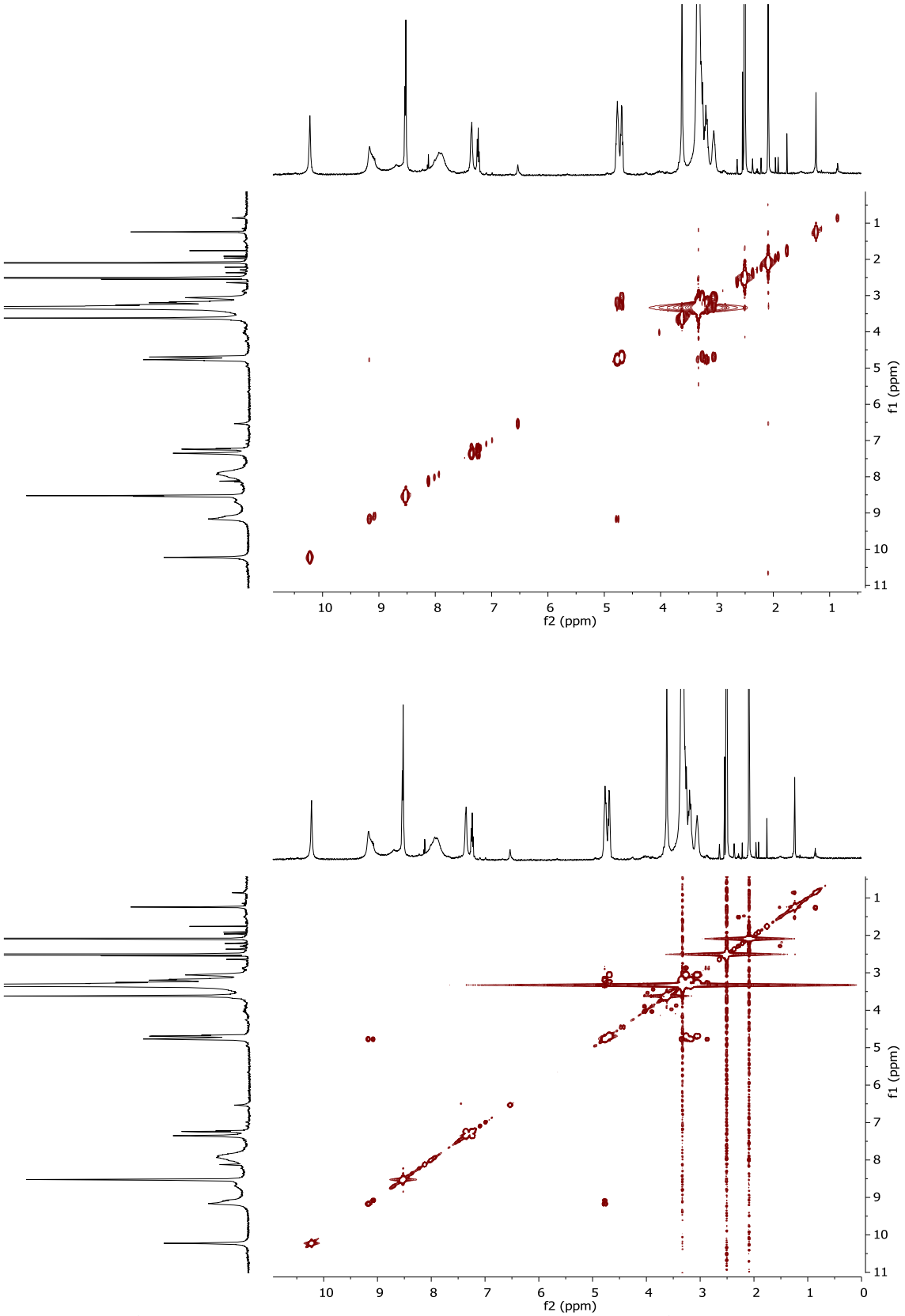


Figure S94. ^{1}H 2D gCOSY (500 MHz, 298 K in DMSO-d_6) and ^{1}H 2D TOCSY (500 MHz, 298 K in DMSO-d_6) spectra of **1_2-$2c_2$.**

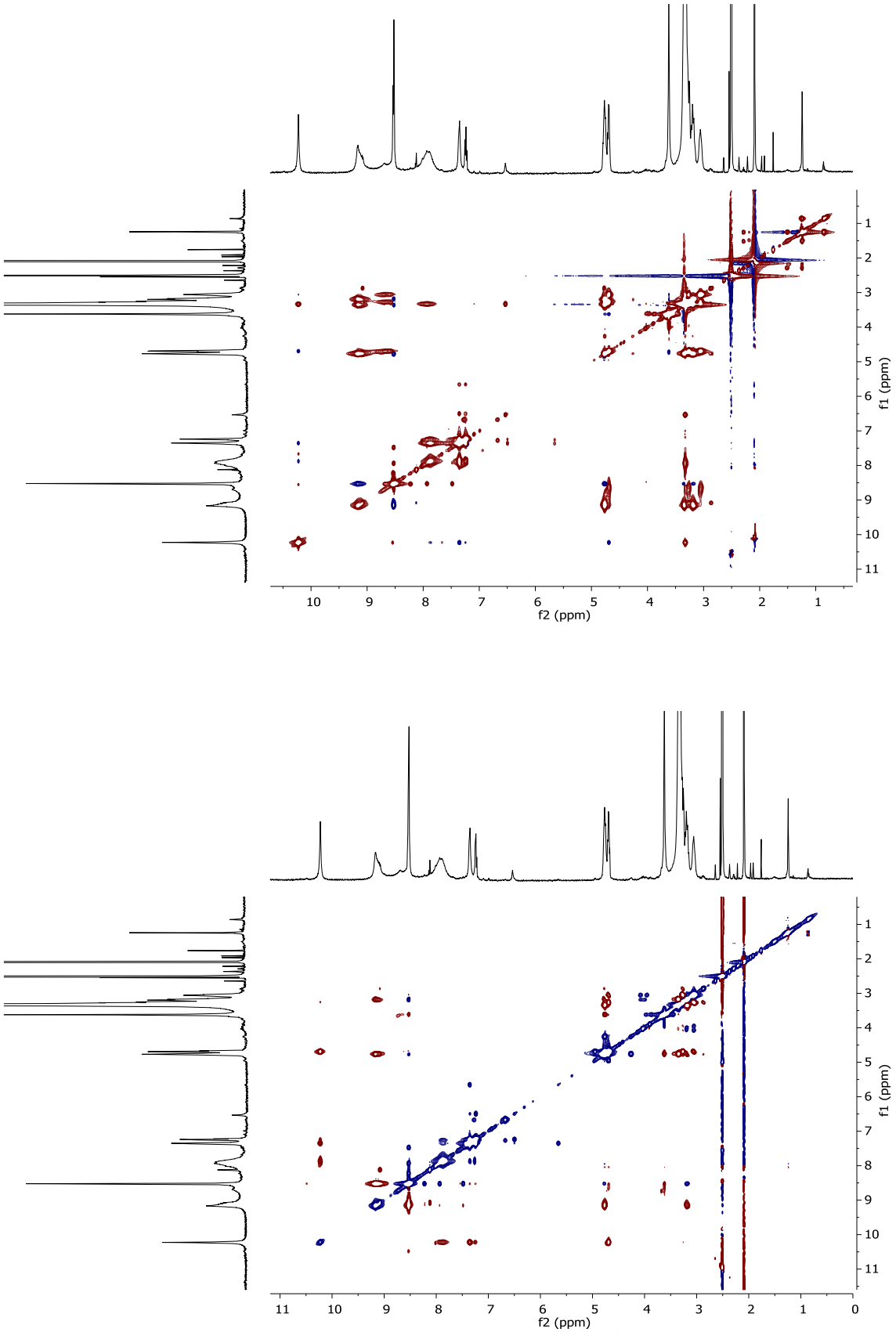


Figure S95. ^{1}H NOESY (500 MHz, 298 K DMSO-d_6) and ^{1}H ROESY (500 MHz, 298 K in DMSO-d_6) spectra of $\mathbf{1}_2$-$\mathbf{2c}_2$.

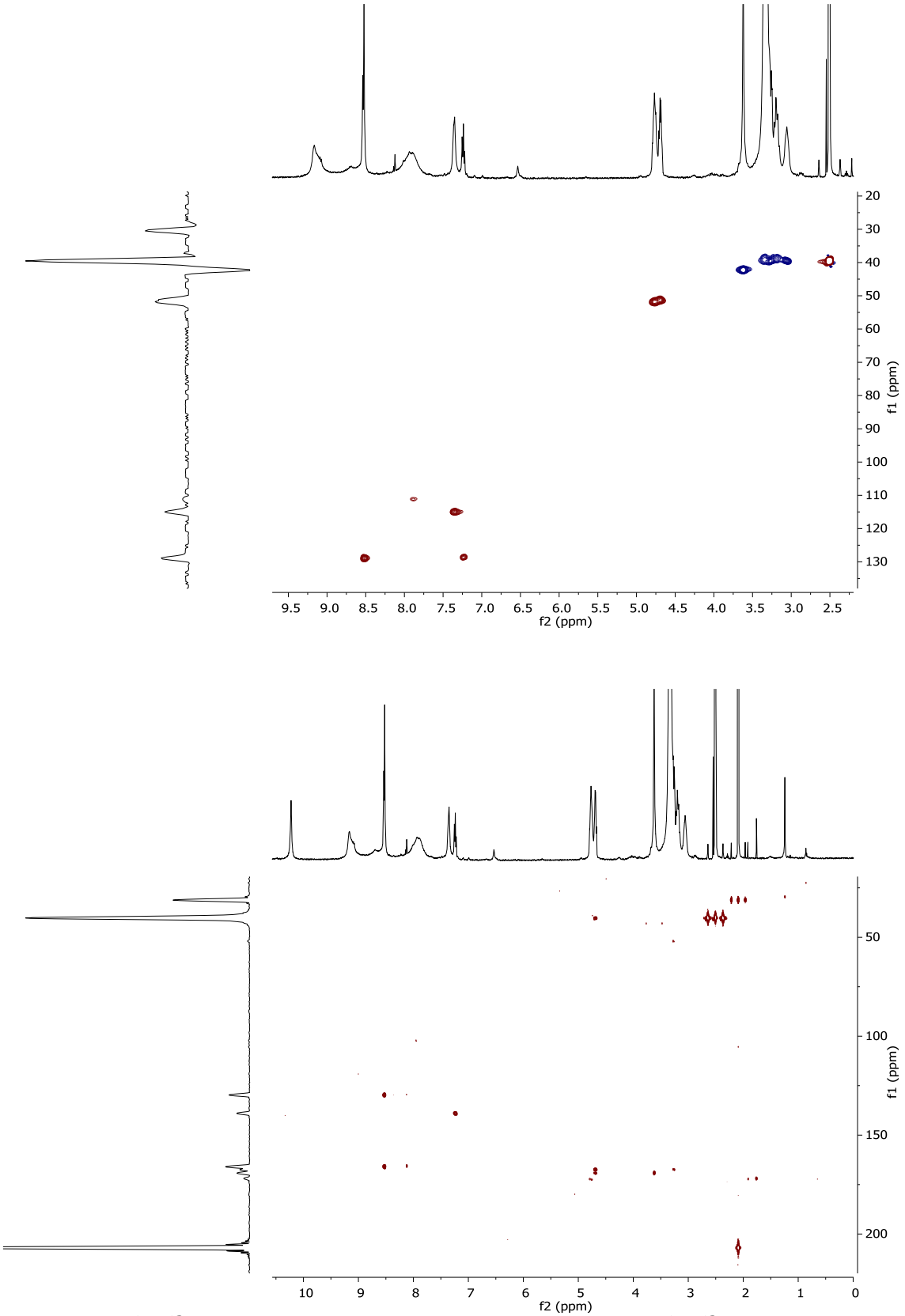


Figure S96. $^{1}H/^{13}C$ gHSQC (500 MHz, 298 K in DMSO-d_6) and $^{1}H/^{13}C$ gHMBC (500 MHz, 298 K in DMSO-d_6) spectra of **1_2-$2c_2$.**

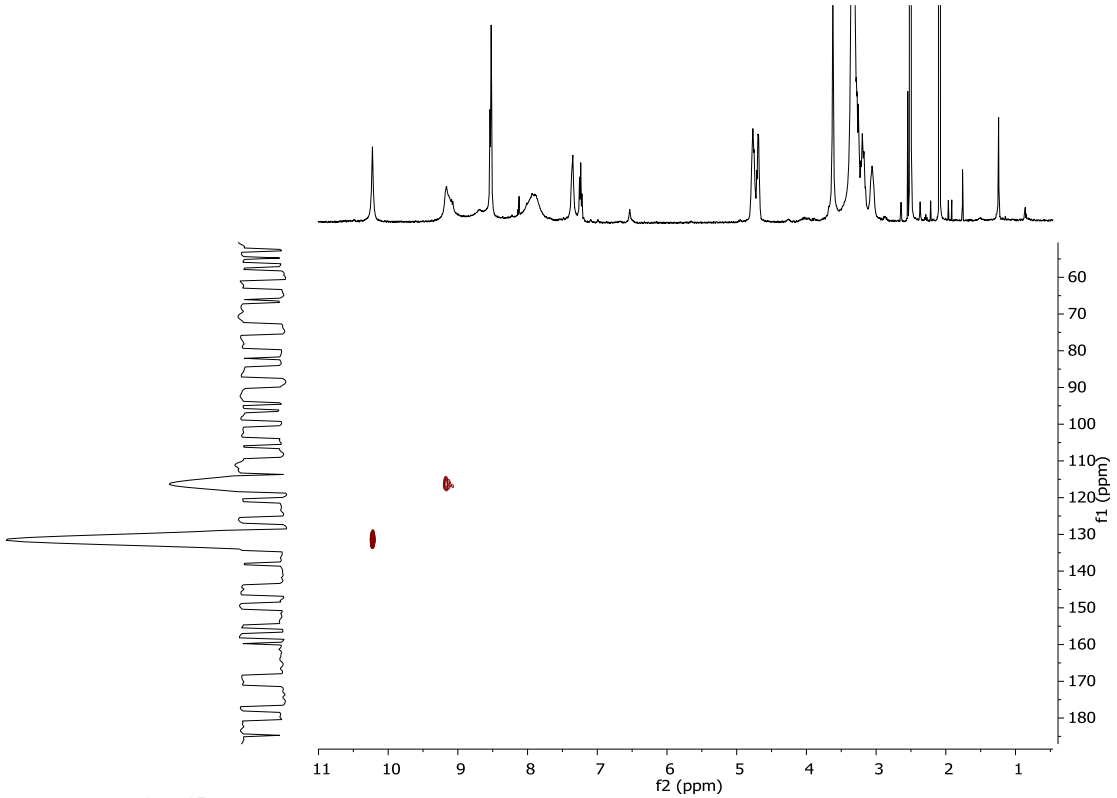


Figure S97. $^{1}H/^{15}N$ gHSQC (500 MHz, 298 K in DMSO-d_6) spectrum of $\mathbf{1_2}$-$\mathbf{2c_2}$.

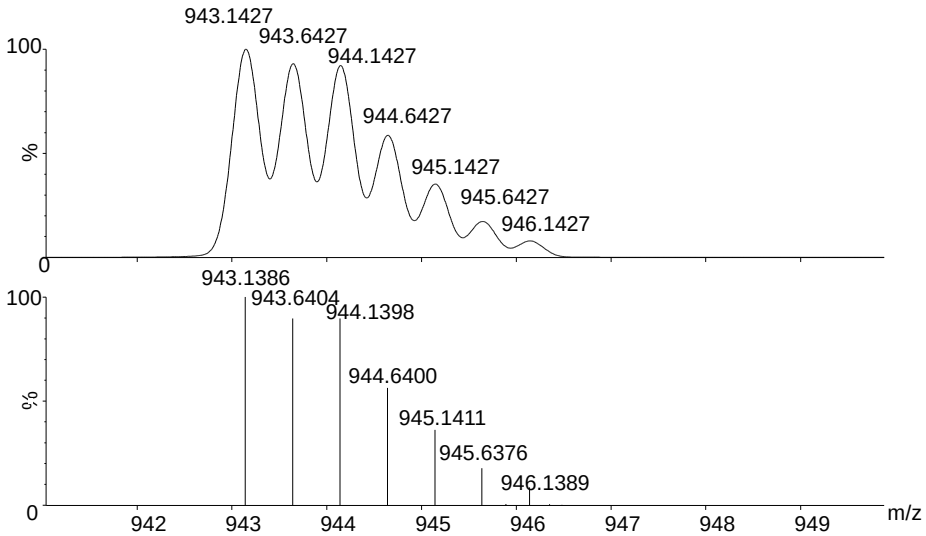


Figure S98. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^{2+}$ of $\mathbf{1_2}$-$\mathbf{2c_2}$.

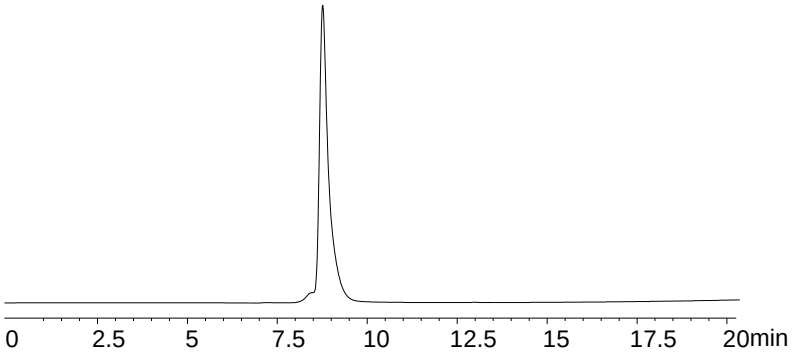


Figure S99. Analytical HPLC-trace of **1_2-2c_2** (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

Comparative ^{1}H spectra of 1_2-2c_2 (Ib) and 2c-3

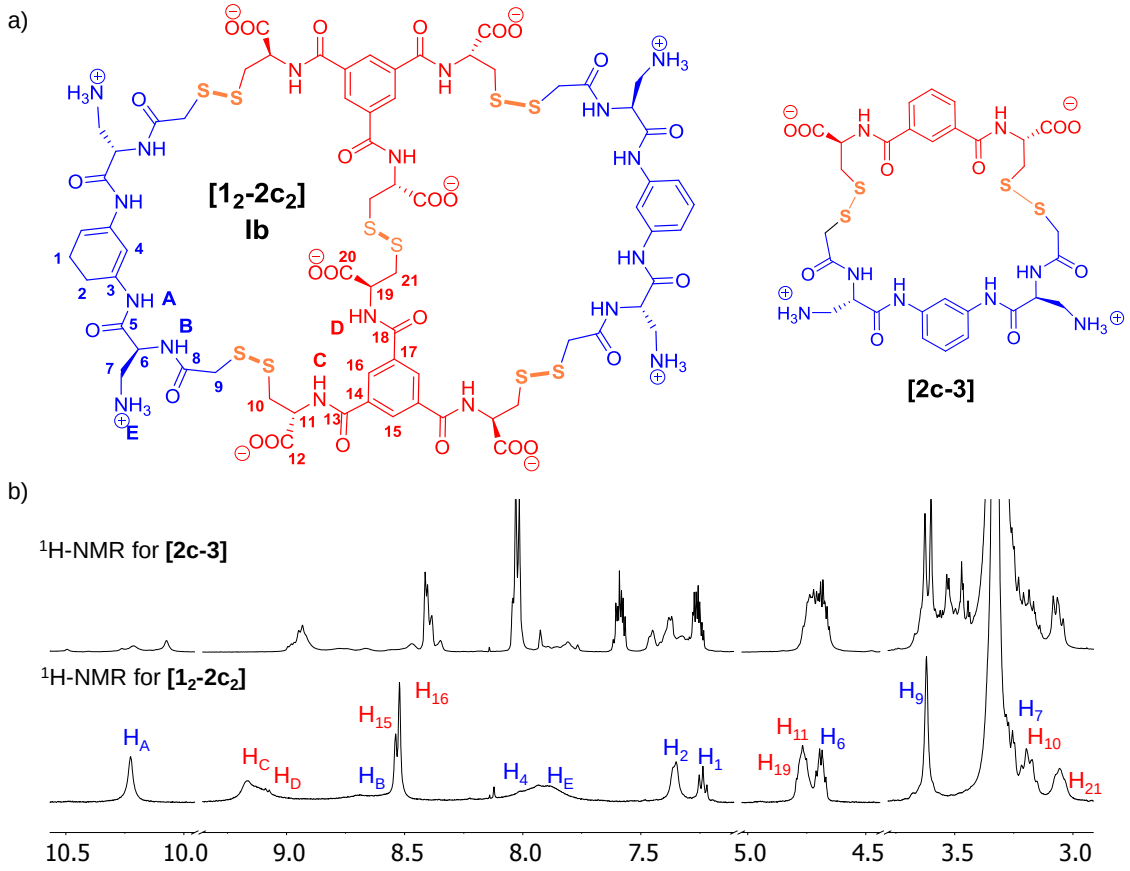


Figure S100. a) Structure for compounds **[1_2-2c_2]** (**Ib**) and **[2c-3]**. b) Comparative ^{1}H specta (500MHz, in DMSO-d_6, 298K) fot molecules **[1_2-2c_2]** and **[2c-3]**.

Analysis of 1_2-$2c_2$ (Ib) and 2c-3 by MALDI-TOF Mass Spectrometry

The samples were analyzed by MALDI-TOF MS in positive and negative reflectron modes and the positive mode was finally chosen. The MS/MS spectra were obtained in the collision-induced dissociation (CID) mode using argon as the collision gas. The samples were prepared with a Matrix α-Cyano-4-hydroxycinnamic acid using the dried droplet method.

[1_2-$2c_2$]
Chemical Formula: $C_{68}H_{80}N_{18}O_{26}S_{10}$
Exact Mass: 1884.2698

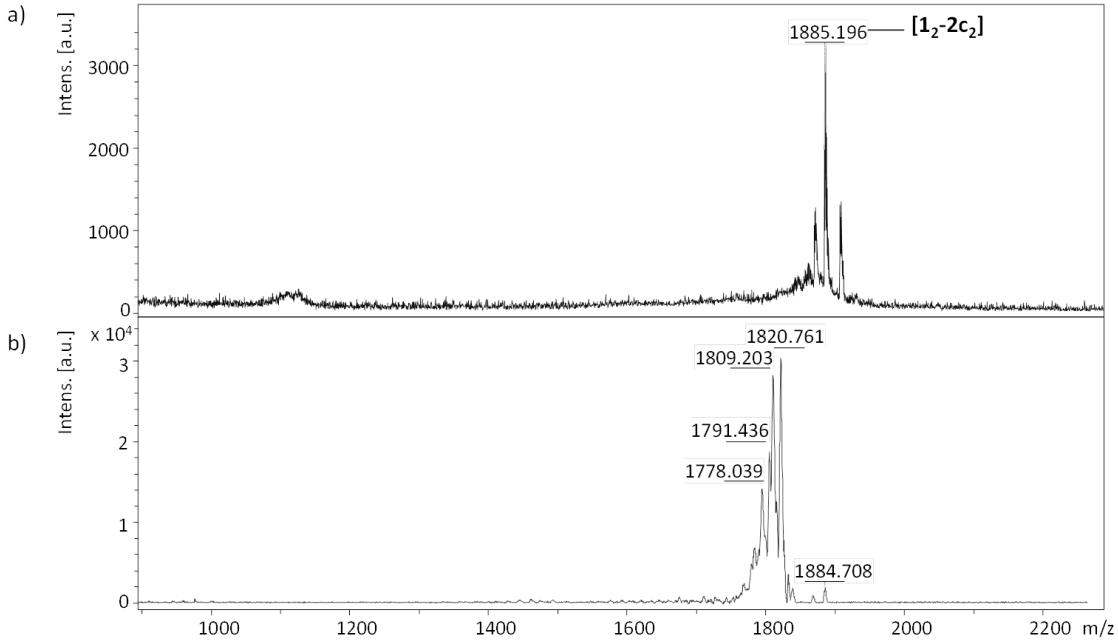


Figure S101. a) Mass spectrum by MALDI-TOF MS of [**1_2-$2c_2$**] in the *m/z* range 900-2300. b) Product ion mass spectra by MALDI-TOF MS/MS of [**1_2-$2c_2$**]: *m/z* 1820.761 (correspond to [**1_2-$2c_2$**] minus (-SS-) unity), *m/z* 1809.203 (correspond to [**1_2-$2c_2$**] minus (-CH_2SS-) unity), *m/z* 1791.436 (correspond to [**1_2-$2c_2$**] minus (-CH_2SSCH_2-) unity), *m/z* 1778.039 (correspond to [**1_2-$2c_2$**] minus (-$SSCH_2COOH$-) unity).

[3-2c]
Chemical Formula: $C_{30}H_{36}N_8O_{10}S_4$
Exact Mass: 796.1437

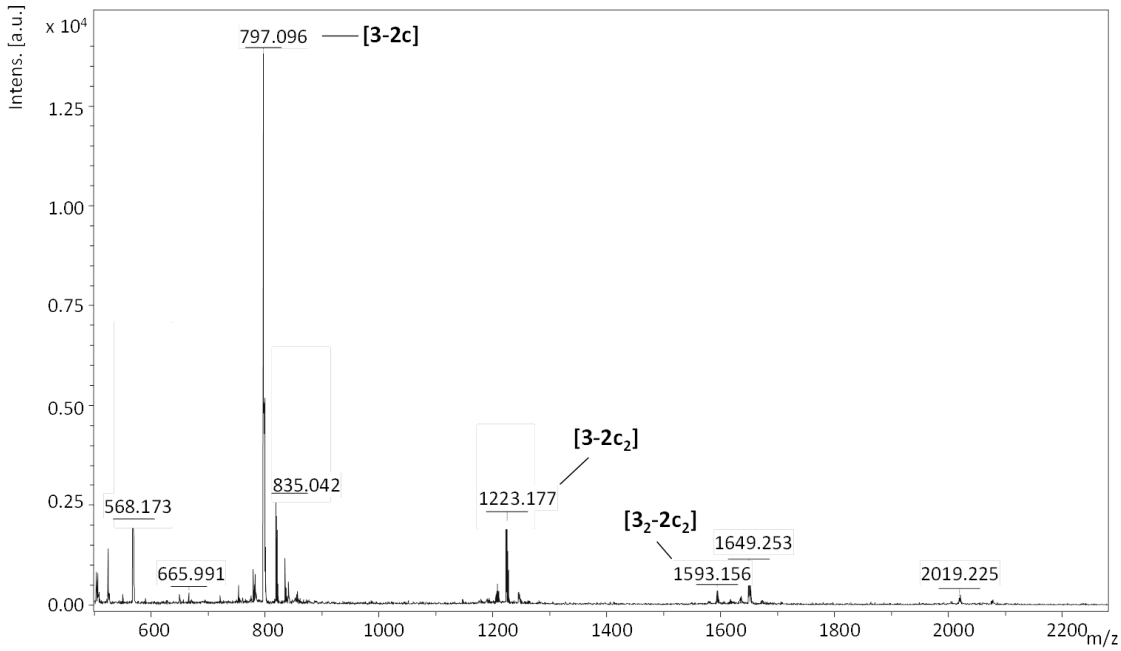


Figure S102. Mass spectrum by MALDI-TOF MS of [**3-2c**] in the *m/z* range 500-2300.

CHAPTER III

Potential applications of DCLs

CHAPTER III: Potential applications of DCLs

3.1. Introduction

3.1.1. Dynamic systems for sensing applications

As stated in Chapter 2, the final distribution of concentrations of the compounds generated in a DCL depends on their chemical structure and on the physical and chemical environment of the system (pH, solvent, concentration, presence of guests molecules, etc.). Therefore, the library composition is a dependent on the environment. Thereby, to use dynamic systems for sensing applications,[1-4] the preparation of dynamic networks that transfer the chemical information and produce a response as a readable signal is required.[5-7]

In sensors based on dynamic systems the information about the analyte is distributed over the entire spectrum. The spectrum therefore represents a "fingerprint" of the analyte. To correlate the spectral changes with the analyte properties of interest (identity, quantity, purity), it is advantageous to use multivariate analyses techniques.[8-9] In this regard, a dynamic sensor is related to sensor array but contrary to these which use independent sensors units, a dynamic sensor is comprised of compounds that are connected by exchange reactions. Furthermore, the various sensors of an array have to be addressed separately, whereas for a dynamic system, a single UV-VIS or fluorescence measurement is sufficient to obtain read-out.

As described in the previous Chapter, Lehn's group was pioneering in the study of the adaptive effects of DCLs. Over the recent years, they have studied constitutionally dynamic polymers ("dynamers") which result from the connection of monomers via reversible covalent bonds[10] that can reorganize the final distribution of the products as a response to the presence of an analyte in the medium of the DCL. In this context, they reported the condensation reaction of a 1:1:1 mixture of 2,7-diaminofluorene (**1**), *trans*-1,4-diaminocyclohexane (**2**), and 2,7-fluorene-bis-carboxaldehyde (**3**) in the presence of different amounts of zinc [as $Zn(BF_4)_2$ $(H_2O)_8$ salt] (Figure 3.1).[11-12] In the absence of zinc, polymer **A** containing the aliphatic monomer **2** is the main product; the addition of increasing amounts of zinc, however, shifted the equilibrium toward polymer **B** until it is the dominant species in the presence of two equivalents of Zn^{2+}. This change in the distribution of the final products of the DCL was explained by the preferential complexation of Zn^{2+} to the more nucleophilic diamine **2**. Polymers **A** and **B** show different absorption spectra and, as a result, the fluorescence emission maximum was shifted from 370 to 493 nm with an increase in intensity. This dynamic polymer constitutes one of the few examples of DCL as a system that is able to sense an analyte that induced constitutional rearrangement.

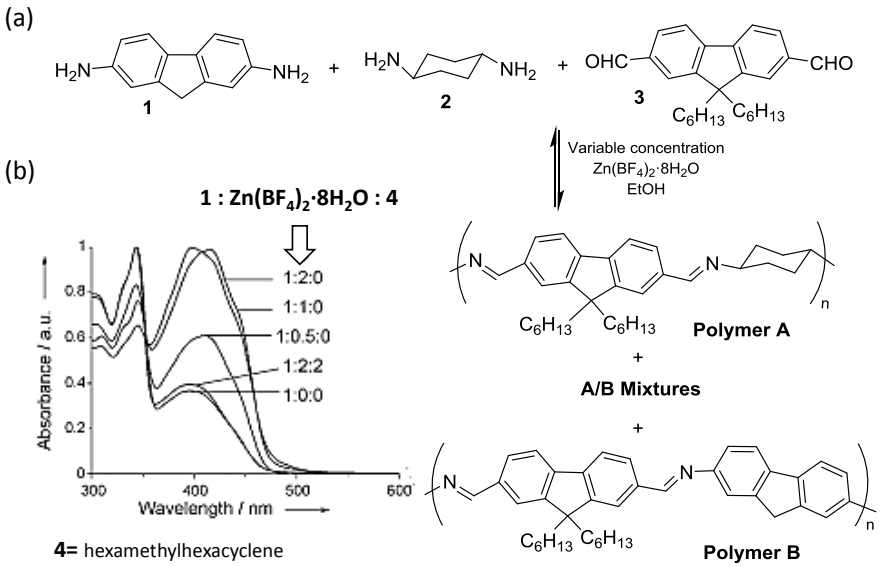


Figure 3.1. (a) Condensation of the diamines **1** and **2** with dialdehyde **3** results in the formation of a dynamic mixture of polymers. The addition of Zn^{2+} favors the formation of polymer **B**. The rearrangement is accompanied by a change of color and fluorescence. (b) Absorption spectra of the component exchange (Polymer **A** ⟶ Polymer **B**) promoted by Zn^{2+} recorded at 298 K in $CDCl_3$ for various amounts of $Zn(BF_4)2{\cdot}8H_2O$ solubilized in CD_3CN. **4** = Hexamethylhexacyclen, used as a zinc trapping agent (figure modified from reference [12]).

If we extend the approach and focus not only on DCL based sensors but in dynamic systems used for sensing applications, following the prior contributions of Lehn's group supramolecular chemistry has made a significant impact in the past decade. The integration of analytical and supramolecular chemistry promotes the birth of a new research area: “supramolecular analytical chemistry”, first termed by Prof. Eric Anslyn.[13] In Prof. Anslyn’s original definition, the field involves analytical chemistry applications of synthetic chemical structures that undergo molecular recognition and self-assembly. In detail, “supramolecular analytical chemistry” exploits the dynamic exchange of synthetic chemical structures that create assemblies which result in signal modulations upon addition of analytes.

In one of the last works of Anslyn’s group, they reported a combination of DCvC and axial chirality for the development of a versatile platform for the binding and sensing the chirality of multiple classes of mononucleophiles. Thought, the modulation of the reactivity within equilibrating open and cyclic biphenyl derivatives, dynamic covalent reactions of alcohols, thiols, primary amines, and secondary amines as well as the control of selectivity were accomplished. Central chirality was transferred to an axially chiral unit and the resulting preferential helical twist was successfully employed to differentiate the handedness and quantify the ee values of chiral analytes that were reported by circular dichroism (CD). The chiral substrates studied are CD inactive above 230 nm and, as a result, any CD signals above 230 nm are indicative of the induced helicity and directly correlate with the chirality of the analytes. A series of chiral alcohols and amines were investigated, and they all afforded significant CD responses. Therefore, the group achieved to differentiate enantiomers of chiral alcohols and amines, demonstrating the successfully of the approach.

In summary, in order to get functional sensors, a measurable signal must be generated upon the addition of an analyte, and hence, molecular recognition (i.e., binding) of the analyte and the mechanism of associated signal transduction must be taken into consideration into sensor design. To achieve sensitivity and selectivity dedicate design and optimization is generally requiered and, as a result, many specific receptors have complicated structures and require multistep synthesis, thereby limiting their availability and practicality. Despite the few examples mentioned above, sensors based on dynamic systems are a promising tool that shown the potential of chemical networks to achieve advanced functions. Specifically, in DCC, if we understand the forces that drive the formation of a particular compound in a chemical dynamic library, we can rationally design molecular networks that give a readable signal in the presence of our analyte.

Typically, DCLs are analyzed by spectroscopic and/or chromatographic methods. For sensing purposes, however, faster and cheaper analysis methods such as fluorescence or ultraviolet - visible (UV - Vis) spectroscopy are preferred. There are different types of sensors based on DCLs such as colorimetric sensors,[14-17] molecular timers[18] and fluorescent sensors.[11-12, 19] Despite that only a few reports about fluorescent sensors based on DCLs that change their constitution upon addition of an analyte had been described, in our group we decided to explore the options of fluorescent sensors.

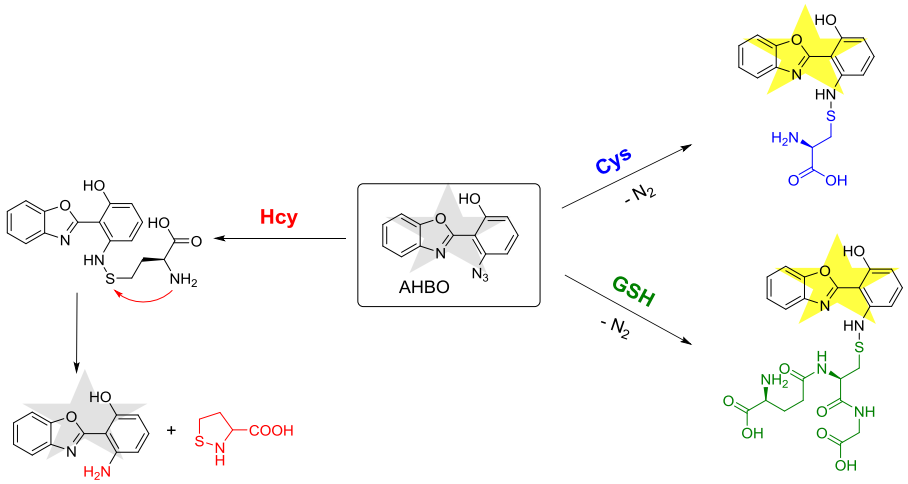
3.1.2. Precedents of cysteine fluorescence sensors

Biological thiols, such as cysteine (Cys) or glutathione (GSH), play essential roles in the cell cycle.[20-21] Owing to their characteristic redox properties and nucleophilicity, they play major unique roles in human health. For example, GSH is responsible for maintaining the redox cycle[22-24] and Cys is involved in protein synthesis and metabolism.[25-26] Thus, abnormal levels of biological thiols are involved in a variety of diseases, for example elevated levels of Cys are considered to be linked with neurotoxicity, and its deficiency is involved in slow growth in children, hair depigmentation, loss of muscle and fat, skin lesion, liver damage and edema ,[27] GSH deficiency is associated with various diseases such as leukocyte loss, cancer, AIDS and neurodegenerative diseases.[28] Consequently, the selective detection of these thiols is of great importance for biomedical research and diagnostics.[29]

Typically, different analytical methods have been used for the detection and determination of specific biothiols, these include high-performance liquid chromatography (HPLC),[30] capillary electrophoresis,[31-32] spectrophotometry,[33] voltammetry,[34] mass spectrometry (MS),[35-36] and HPLC-MS/MS.[37] Fluorescence detection offers several advantages over these techniques including high sensitivity, speed, low cost, simplicity of operation and non-invasiveness.[38-39]

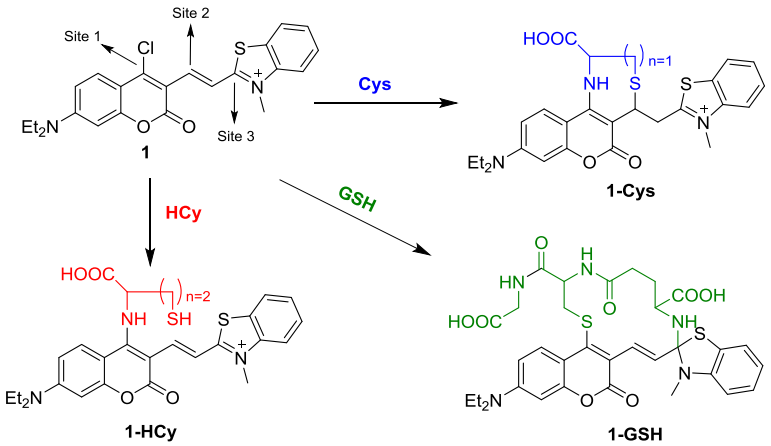
Despite the considerable number of Cys sensors that have been reported, selectivity against other biothiols is still an issue: these sensors usually form covalent adducts with any sulfhydryl-containing molecules. In this section we focus on Cys sensors,[38, 40-42] the selective detection of this molecule is of great importance for clinical diagnosis. However, specific and selective methods for the detection of Cys are still scarce and most of these sensors require the presence of the free thiol, precluding the detection of cystine.[43-47]

Prior studies have reported fluorescent Cys sensors from different perspectives, a summary of them are described hereafter. In the past years, Xu and co-workers developed a simple azido-substituted hydroxyphenylbenzoxazole derivative (AHBO) as a biothiol sensor.[48] In comparison with the reported azido-substituted fluorescent sensors for the detection of thiol, AHBO shows no obvious fluorescence response to sulfide but a selective turn-on fluorescence response to Cys and GSH, which is based on the mechanism of selective nucleophilic substitution–rearrangement reactions (Scheme 3.1). The fluorescence intensities of AHBO in the absence and presence of GSH are pH independent in the range of 7–8, demonstrating that AHBO can detect GSH in the biological environment. AHBO was the first fluorescent sensor based on azide recognition for the selective sensing of Cys and GSH over homocysteine (Hcy) and other sulfides. The selective fluorescence response of AHBO to GSH and Cys over Hcy is attributed to the nucleophilic substitution–rearrangement reactions that take place in the case of Hcy but not in the case of GSH and Cys. The exceptional and interesting response derived from thiol–azide reaction provides a new strategy for biothiol detection, although this system is not capable to give a selective response between Cys and GSH; thus, it is not able to selectively detect Cys.



Scheme 3.1. Reaction mechanism for AHBO sensor that shows selective turn-on fluorescence response to Cys and GSH over HCy. AHBO is a selective sensor of Cys versus HCy but not against GSH (figure modified from reference [48]).

Another example was provided by Guo's group who presented a novel discrimination strategy for biothiols based on a chlorinated coumarin-hemicyanine dye with three potential reaction sites.[49] The probe could not only discriminate Cys (or GSH) from Hcy/GSH (or Cys/Hcy) but also simultaneously detect Cys and GSH from different emission channels and thus holds great potential in biological applications. The Cys-induced substitution−rearrangement−cyclization, Hcy-induced substitution− rearrangement, and GSH-induced substitution−cyclizatioin cascades lead to the corresponding aminocoumarin, amino-coumarin-hemicyanine and thiol-coumarin with distinct photophysical properties, enabling Cys and GSH to be selectively detected from different emission channels at two different excitation wavelengths. The design rationale is depicted in Scheme 3.2.



Scheme 3.2. Reaction mechanism of **1** with Cys, Hcy and GSH (figure modified from reference [49]).

Despite this and other relevant advances[42, 50] there is a need for new, more efficient sensing tools for relevant biothiols. For example Guo's sensor needs 10 eq of added thiol for a reliable response. Therefore, the new systems to sense biothiols require improved sensitivity, greater fluorescence turn-on response and decreased response time. Moreover, most of the sensors described require the presence of the free thiol, precluding the detection of cystine.

With these challenges in mind, we designed a molecular network able to give a selective and sensitive fluorescence response for the detection of Cys and its oxidized form, cystine in biological samples. To the best of our knowledge, our system is the first fluorescent sensor-based on a dynamic covalent library of pseudopeptidic BBs upon disulfide exchange for the selective detection Cys and cystine.

3.1.3. Cystinuria disease

As explained in the previous section, biological thiols, such as Cys, play essential roles in the cell cycle and abnormal levels of Cys are involved in a variety of diseases. One of the diseases in which Cys is involved is Cystinuria. Cystinuria is an autosomal recessive disease,[51] which means that the defective gene responsible for the disease is located on an autosome, and two copies of the defective gene (one inherited from each parent) are required in order to be born with the disease. The parents of an individual with an autosomal recessive disease both carry one copy of the defective gene, but usually do not experience any signs or symptoms of the disease. This means that it is a disease from which no premature or preventive control is carried out, normally it is not discovered until symptoms are felt in the individual. The overall prevalence of cystinuria is approximately 1 in 7,000 neonates.

Cystinuria is caused by mutations in the SLC3A1 and SLC7A9 genes. These defects prevent proper reabsorption of basic amino acids: cysteine, lysine, ornithine and arginine.[52] Under normal circumstances, these amino acids, including Cys, are reabsorbed into the blood from the filtered fluid that will become urine. Mutations in either of these genes disrupt the ability of reabsorb these amino acids, allowing them to become concentrated in the urine and leading to the formation of stones in kidneys, bladder, and ureters due to the insolubility of the Cys oxidized dimer, cystine.[51] The other amino acids that are not reabsorbed do not create crystals in urine.

The symptoms are the usual ones of kidney colic: strong pain in the back and the side, as well as the appearance of blood in the urine. If not treated properly, the disease could cause serious damage to the kidneys and surrounding organs and, in some rare cases, death. At present, this disease has no cure and its treatment is usually based on the ingestion of water in abundance,[f] alkalization of the urine with citrate supplementation or acetazolamide, and dietary modification to reduce salt and protein intake (especially methionine). If this fails then patients are usually started on chelation therapy with an agent such as penicillamine.[g,53] Once renal stones have formed, however, surgery can be necessary. Therefore, currently the best strategy is solubilizing cystine to avoid kidney stones.

For the diagnosis of cystinuria some methods have been used. The oldest one but still used, is sodium-nitroprusside test, effective in about 70% of the cases but need concentrations of cystine above 315 mM and false positives can be obtained for example in the case of patients with homocystinuria or ketonuria.[54] The simplest diagnostic process is the observation of hexagonal crystals of cystine in urine but is only possible in a small group of patients between 19-26% of the cases. At the present, the ion-exchange chromatographic quantitation methods provide the best results of the analysis of kidney stones, which requires derivatization and calibration. However, still no specific and selective method to diagnose Cys so it is necessary the development of sensors.

It is important to keep in mind that Cystinuria disease is considered for cystine concentrations above 0.4 mM, and stones in the kidney are not formed until the concentrations are above 0.8 mM,[54] therefore, cystinuria is usually asymptomatic when no stones are formed. This is an important issue in the detection of the disease. Despite cystinuria is not a very common disease, it's a disease that can be easily complicated due to the few studies that exist about it and the lack of adequate diagnosis.

With these challenges in mind, our goal was to design a molecular network that yields a fluorescence response in the presence of Cys or cystine. The network is based on a dynamic covalent library of pseudopeptidic BBs[55] upon disulfide exchange.

[f] Adequate hydration is the foremost aim of treatment to prevent cystine stones, increase the urine volume because the concentration of cystine in the urine is reduced which prevents cystine from precipitating from the urine and forming stones.

[g] Penicillamine is a drug that acts to form a complex with cystine that is 50 times more soluble than cystine itself.

3.1.4. References

[1] S. Otto, K. Severin, in *Creative Chemical Sensor Systems* (Ed.: T. Schrader), Springer Berlin Heidelberg, Berlin, Heidelberg, **2007**, pp. 267-288.
[2] A. Buryak, A. Pozdnoukhov, K. Severin, *Chem. Commun.* **2007**, 2366-2368.
[3] B. Andrey, S. Kay, *Angew. Chem. Int. Ed.* **2005**, *44*, 7935-7938.
[4] A. Buryak, K. Severin, *J. Comb. Chem.* **2006**, *8*, 540-543.
[5] E. V. Anslyn, *J. Org. Chem.* **2007**, *72*, 687-699.
[6] L. You, D. Zha, E. V. Anslyn, *Chem. Rev.* **2015**, *115*, 7840-7892.
[7] M. H. Lee, J. S. Kim, J. L. Sessler, *Chem. Soc. Rev.* **2015**, *44*, 4185-4191.
[8] A. T. Wright, E. V. Anslyn, *Chem. Soc. Rev.* **2006**, *35*, 14-28.
[9] K. J. Albert, N. S. Lewis, C. L. Schauer, G. A. Sotzing, S. E. Stitzel, T. P. Vaid, D. R. Walt, *Chem. Rev.* **2000**, *100*, 2595-2626.
[10] J.-M. Lehn, *Prog. Polym. Sci.* **2005**, *30*, 814-831.
[11] N. Giuseppone, J.-M. Lehn, *J. Am. Chem. Soc.* **2004**, *126*, 11448-11449.
[12] N. Giuseppone, G. Fuks, J. M. Lehn, *Chem. Eur. J.* **2006**, *12*, 1723-1735.
[13] J. W. Steed, J. L. Atwood, *Supramolecular chemistry*, John Wiley & Sons, **2013**.
[14] B. T. Nguyen, E. V. Anslyn, *Coord. Chem. Rev.* **2006**, *250*, 3118-3127.
[15] L. Fabbrizzi, M. Licchelli, A. Taglietti, *Dalton Trans.* **2003**, 3471-3479.
[16] J. J. Lavigne, E. V. Anslyn, *Angew. Chem. Int. Ed.* **1999**, *38*, 3666-3669.
[17] Z. Zhong, E. V. Anslyn, *J. Am. Chem. Soc.* **2002**, *124*, 9014-9015.
[18] A. Buryak, F. Zaubitzer, A. Pozdnoukhov, K. Severin, *J. Am. Chem. Soc.* **2008**, *130*, 11260-11261.
[19] Y. Ruff, J. M. Lehn, *Angew. Chem. Int. Ed.* **2008**, *47*, 3556-3559.
[20] Z. A. Wood, E. Schröder, J. Robin Harris, L. B. Poole, *Trends Biochem. Sci.* **2003**, *28*, 32-40.
[21] R. O. Ball, G. Courtney-Martin, P. B. Pencharz, *J.Nutr.* **2006**, *136*, 1682S-1693S.
[22] A. Meister, M. E. Anderson, *Annu. Rev. Biochem.* **1983**, *52*, 711-760.
[23] C. Perricone, C. De Carolis, R. Perricone, *Autoimmun. Rev.* **2009**, *8*, 697-701.
[24] R. Franco, O. J. Schoneveld, A. Pappa, M. I. Panayiotidis, *Arch. Physiol. Biochem.* **2007**, *113*, 234-258.
[25] N. M. Giles, G. I. Giles, C. Jacob, *Biochem. Biophys. Res. Commun.* **2003**, *300*, 1-4.
[26] V. Gazit, R. Ben-Abraham, R. Coleman, A. Weizman, Y. Katz, *Amino Acids* **2004**, *26*, 163-168.
[27] S. Shahrokhian, *Anal. Chem.* **2001**, *73*, 5972-5978.
[28] H. Tapiero, D. M. Townsend, K. D. Tew, *Biomed. Pharmacother.* **2003**, *57*, 134-144.
[29] L. A. Pham-Huy, H. He, C. Pham-Huy, *Int. J. Biomed. Sci.* **2008**, *4*, 89-96.
[30] J. Vacek, B. Klejdus, J. Petrlova, L. Lojkova, V. Kuban, *Analyst* **2006**, *131*, 1167-1174.
[31] J. S. Stamler, J. Loscalzo, *Anal. Chem.* **1992**, *64*, 779-785.
[32] T. Inoue, J. R. Kirchhoff, *Anal. Chem.* **2002**, *74*, 1349-1354.
[33] I. Rahman, A. Kode, S. K. Biswas, *Nat. Protoc.* **2007**, *1*, 3159-3165.
[34] A. Kaczyñska, M. M. A. L. Pelsers, A. Bochowicz, M. Kostrubiec, J. F. C. Glatz, P. Pruszczyk, *Clin. Chim. Acta* **2006**, *371*, 117-123.
[35] M. J. MacCoss, N. K. Fukagawa, D. E. Matthews, *Anal. Chem.* **1999**, *71*, 4527-4533.
[36] A. P. Vellasco, R. Haddad, M. N. Eberlin, N. F. Hoehr, *Analyst* **2002**, *127*, 1050-1053.
[37] B. Seiwert, U. Karst, *Anal. Chem.* **2007**, *79*, 7131-7138.
[38] X. Chen, Y. Zhou, X. Peng, J. Yoon, *Chem. Soc. Rev.* **2010**, *39*, 2120-2135.
[39] Y. Yang, Q. Zhao, W. Feng, F. Li, *Chem. Rev.* **2013**, *113*, 192-270.
[40] H. S. Jung, X. Chen, J. S. Kim, J. Yoon, *Chem. Soc. Rev.* **2013**, *42*, 6019-6031.
[41] L.-Y. Niu, Y.-Z. Chen, H.-R. Zheng, L.-Z. Wu, C.-H. Tung, Q.-Z. Yang, *Chem. Soc. Rev.* **2015**, *44*, 6143-6160.
[42] C. X. Yin, K. M. Xiong, F. J. Huo, J. C. Salamanca, R. M. Strongin, *Angew. Chem. Int. Ed.* **2017**, *56*, 13188-13198.
[43] H. Maeda, H. Matsuno, M. Ushida, K. Katayama, K. Saeki, N. Itoh, *Angew. Chem. Int. Ed.* **2005**, *44*, 2922-2925.
[44] S. Sreejith, K. P. Divya, A. Ajayaghosh, *Angew. Chem. Int. Ed.* **2008**, *47*, 7883-7887.
[45] X. Sun, E. V. Anslyn, *Angew. Chem. Int. Ed.* **2017**, *56*, 9522-9526.
[46] M. Zhang, M. L. Saha, M. Wang, Z. Zhou, B. Song, C. Lu, X. Yan, X. Li, F. Huang, S. Yin, P. J. Stang, *J. Am. Chem. Soc.* **2017**, *139*, 5067-5074.
[47] F. Y. Thanzeel, C. Wolf, *Angew. Chem. Int. Ed.* **2017**, *56*, 7276-7281.

[48] D. Zhang, Z. Yang, H. Li, Z. Pei, S. Sun, Y. Xu, *Chem. Commun.* **2016**, *52*, 749-752.
[49] J. Liu, Y.-Q. Sun, Y. Huo, H. Zhang, L. Wang, P. Zhang, D. Song, Y. Shi, W. Guo, *J. Am. Chem. Soc.* **2014**, *136*, 574-577.
[50] G. X. Yin, T. T. Niu, Y. B. Gan, T. Yu, P. Yin, H. M. Chen, Y. Y. Zhang, H. T. Li, S. Z. Yao, *Angew. Chem. Int. Ed.* **2018**, *57*, 4991-4994.
[51] E. Fjellstedt, L. Harnevik, J.-O. Jeppsson, H.-G. Tiselius, P. Söderkvist, T. Denneberg, *Urol. Res.* **2003**, *31*, 417-425.
[52] K. Ahmed, P. Dasgupta, M. S. Khan, *Postgrad. Med. J.* **2006**, *82*, 799-801.
[53] K. Ahmed, M. S. Khan, K. Thomas, B. Challacombe, M. Bultitude, J. Glass, R. Tiptaft, P. Dasgupta, *Urol. Int.* **2008**, *80*, 141-144.
[54] C. S. Biyani, J. J. Cartledge, *EAU-EBU Update Series* **2006**, *4*, 175-183.
[55] J. Atcher, J. Sola, I. Alfonso, *Org. Biomol. Chem.* **2017**, *15*, 213-219.

3.2. Specific objectives and hypothesis

The main objective of the present Chapter is to apply all the knowledge obtained in the study of pseudopeptidic DCLs (described in the previous Chapter) to perform dynamic systems with a practical application. We focused on the design and preparation of specific amino acid sensors in aqueous media. The specific aim is to prepare a sensor that selectively responds to the presence of Cys and L-cystine through self-recognition processes and gives a fluorescent response. With this purpose, we hypothesized that the design of a Cys sensor that works in biological fluids as urine could be a novel tool for cystinuria diagnosis.

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

1) To prepare monothiol BBs with suitable fluorescent aromatic groups capable of generate dynamic combinatorial libraries and fluoresce readable signal.
    a. To synthetize and characterize the candidates.
    b. To evaluate and analyze the fluorescence and solubility of each candidate.
2) To generate DCLs with the new monopodal constituent and find suitable conditions for the development of the sensor in aqueous media.
3) To check the individual and crossed responses, and to study the interference from other amino acids with the sensor.
4) To analyze biological samples (urine) with the new sensor developed.

3.3. PUBLICATION D:

A dynamic chemical network for cystinuria diagnosis

Angew. Chem. Int. Ed. **2018**

Publiation Date (online): 12 April 2018

Dynamic Networks

International Edition: DOI: 10.1002/anie.201802189
German Edition: DOI: 10.1002/ange.201802189

A Dynamic Chemical Network for Cystinuria Diagnosis

Maria Lafuente, Jordi Solà, and Ignacio Alfonso**

Abstract: *The study of molecular networks represents a conceptual revolution in chemistry. Building on previous knowledge and after understanding the rules of non-covalent interactions, the design of stimulus-responsive chemical systems is possible. Herein we report a new strategy, based on the reorganization of a dynamic chemical network that generates new fluorescent associations in the presence of cysteine or cystine. The binding and sensing units are encoded in the components that dynamically assemble and disassemble responding to external stimuli as a successful tool to detect both cysteine and cystine in aqueous media. Moreover, the dynamic sensing system works in human urine, as a prospective application for cystinuria diagnosis.*

Traditionally studying pure compounds, chemists have a new methodology to face their challenges.[1] The understanding of the emergent properties in complex molecular systems opens opportunities to work with complexity.[2] The use of dynamic, reversible reactions has led to the preparation of complex exchanging systems[3] while the current analytical tools allow complex mixtures to be dealt with.[4] Inspired by the cellular signaling,[5] we envisioned the preparation of artificial networks that mimic the transfer of information and produce a chemical response,[6] such as a readable signal.[7] Specifically, we present a dynamic system able to selectively sense a biologically relevant analyte, cysteine, in its reduced or its oxidized forms (cystine) in aqueous media and in a biofluid (urine) for diagnostic application.

Biological thiols, such as cysteine (Cys) or glutathione (GSH), play essential roles in the cell cycle.[8] Thus, GSH is responsible for maintaining the redox cycle[9] and Cys is involved in protein synthesis and metabolism.[10] We focused on cystinuria, a disorder of cysteine and dibasic amino acids transport that produces stones in kidneys, bladder, and ureters due to the insolubility of its oxidized dimer, cystine.[11] Diagnosis of cystinuria may be done by the observation of crystals in urine, or by the sodium-nitroprusside test, which leads to false-positives in patients taking acetylcysteine.[12] An alternative is ion-exchange chromatographic quantitation, which requires derivatization and calibration.[13] Despite the molecular sensors reported for Cys, selectivity against other biothiols is still an issue.[14] Moreover, most of these sensors require the presence of the free thiol, precluding the detection of cystine.[15] With these challenges in mind, we designed a molecular network that yields a fluorescence response in the presence of cysteine or cystine. The network is based on a dynamic covalent library of pseudopeptidic building blocks (BBs)[16] upon disulfide exchange.[17] We discovered that mixing equimolar amounts of BB **1**, **2**, and **3a** (Figure 1) at pH close to neutrality in aqueous media resulted in the almost exclusive assembly of the heterotrimer **4a**. This selectivity is due to the self-recognition mediated by non-covalent interactions between the different functionalities of the constituents (Figure 1B).[17]

We anticipated that by replacing cysteine (**3a**) with a different monothiol containing a fluorophore with a lesser tendency to form the corresponding heterotrimer (**3b**), a complex mixture would occur. Addition of **3a** to this library would lead to the most stable **4a**, with the concomitant redistribution of the dynamic mixture, releasing the fluorophore as its disulfide homodimer. This conversion could be read in the fluorescence spectra due to the formation of excimers[18] (Figure 1B).

After exploring different candidates (Figure S35 in the Supporting Information), we selected **3b** with suitable fluorescence spectra of its reduced and oxidized ([$\mathbf{3b_2}$]) forms: [$\mathbf{3b_2}$] presents an excimer emission band at about 500 nm that is absent in the monomer, which is characterized by a low-wavelength band with fine structure (c.a. 350–450 nm). After optimization, we found that a library containing **1** and **2** (0.1 mM each), and **3b** (0.05 mM) minimized the formation of [$\mathbf{3b_2}$] homodimer, with almost no excimer emission (Figure 2, blue trace). When cysteine was added to the reaction the excimer band increased (Figure 2, red) owing to the formation of [$\mathbf{3b_2}$].

The emission at 501 nm was detected with a Cys concentration of 50 μM and increasing the concentration of **3a** resulted in an increase of the fluorescent response (Figure 2). This can be ratiometrically read by the excimer (501 nm) over monomer (385 nm) intensity ratio. Considering this ratio, the response to the presence of cysteine goes from 1.4 (50 μM) to 3.8 (1 mM) times the blank. Since the normal presence of Cys in urine is approximately 35 μM and the occurrence of stone-producing cystinuria starts from approximately 0.8 mM, our method fits within these values.

Thanks to the dynamic nature of the sensing system based on disulfide formation and exchange, [$\mathbf{3a_2}$] behaved similar to **3a** (Figure 3A, red and green traces). This ability allows detecting cysteine in its reduced or oxidized forms in aqueous media without an extra preparation step.

Regarding the selectivity against different biothiols, the system gives almost no response to the presence of other biologically relevant cysteine derivatives, such as GSH (Figure 3A, blue) or N-acetylcysteine (Figure 3A, orange), over-

[*] M. Lafuente, Dr. J. Solà, Dr. I. Alfonso
Department of Biological Chemistry and Molecular Modeling
IQAC-CSIC
Jordi Girona 18–26, 08034 Barcelona (Spain)
E-mail: jordi.sola@iqac.csic.es
ignacio.alfonso@iqac.csic.es

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201802189.

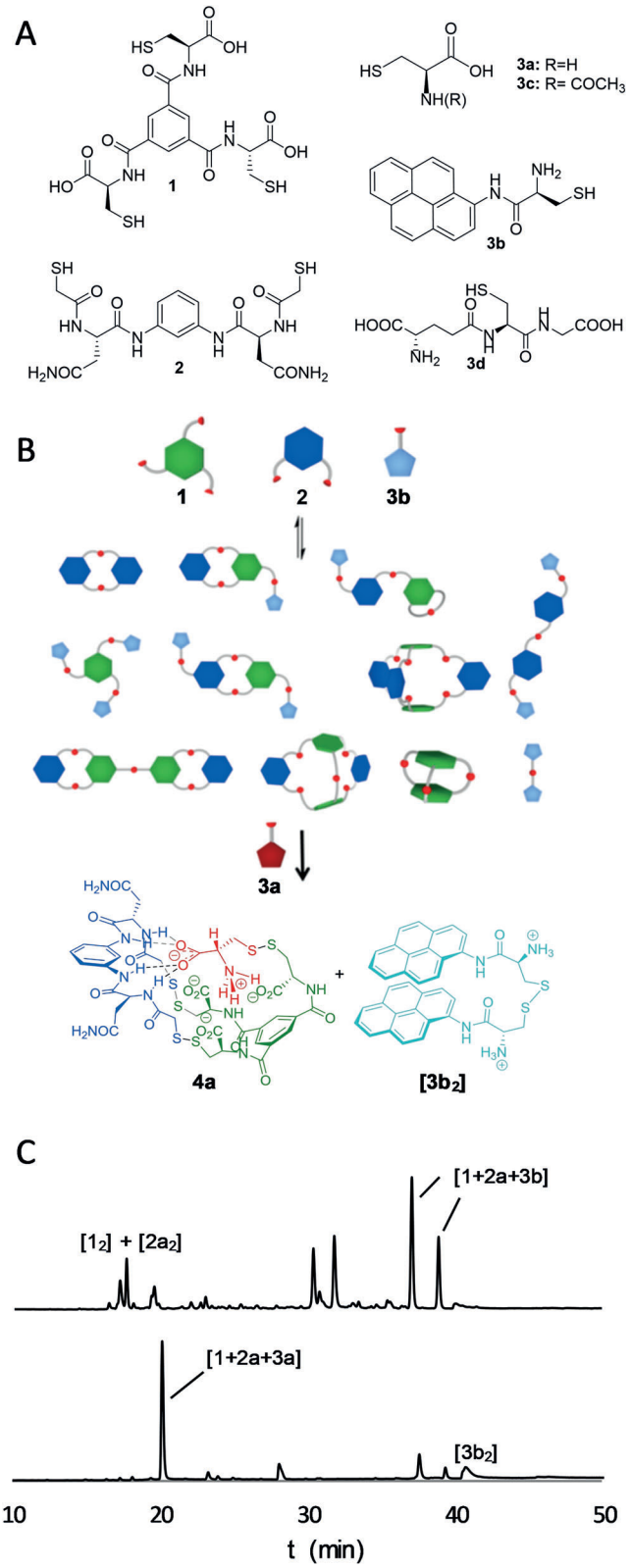


Figure 1. A) Chemical structures of the BBs. B) Mixing BBs **1**–**2**–**3b** produces a complex mixture. The addition of **3a** leads to the formation of heterotrimer **4a** and the homodimer $[\mathbf{3b}]_2$. C) HPLC traces for the oxidized mixtures of **1**–**2**–**3b** (top) and **1**–**2**–**3b**–**3a** (bottom).

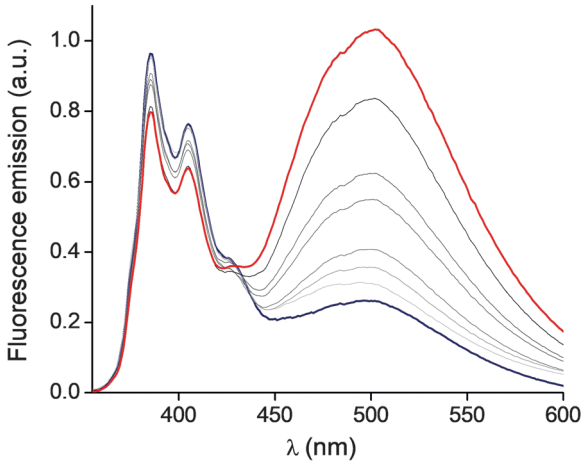


Figure 2. Fluorescence emission spectra of the dynamic library (blue trace; 25% DMSO in aqueous bis-tris buffer at pH 6.5) and under increasing concentrations of Cys (0.025–2.5 mM). Samples were 34-times diluted in 1:1 H_2O:DMSO for the measurements. See text for details.

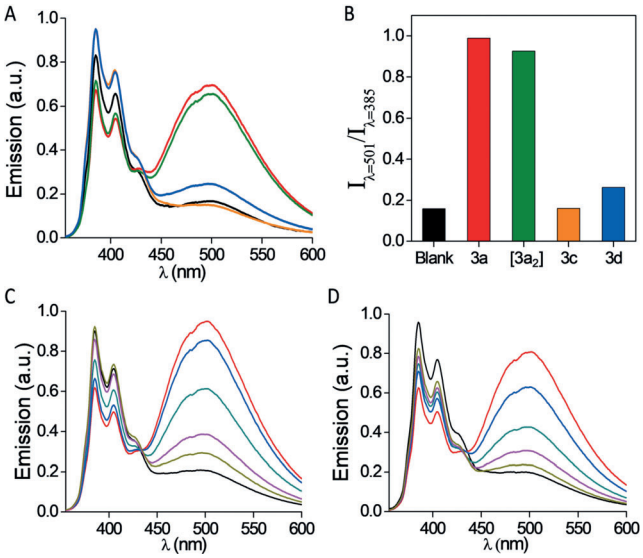


Figure 3. A) Fluorescence emission spectra of the library alone (black) and in the presence of 0.5 mM Cys (red), 0.25 mM cystine (green), 0.5 mM N-acetyl-Cys (orange) and 0.5 mM GSH (blue). B) Plot of the excimer/monomer ratio for the different biothiols. C) Sensing of Cys (0.05–1.0 mM, green to red traces) in the presence of the three basic amino acids (Lys, Orn, and Arg) at a typical concentration found in urine. D) As in (C) but in a buffer containing (Lys, Orn, Arg, Asn, Gly, Ala, His, Asp, β-Ala, Ser, Tyr, and Met). Samples were 34-times diluted in 1:1 H_2O:DMSO for the measurements.

coming false-positives. The selectivity for **3a** and $[\mathbf{3a}_2]$ is illustrated in Figure 3B by the excimer/monomer ratio plot, and can be explained by the specific recognition motif within the folded conformation of **4a** that is precluded with GSH and acetylcysteine. Additionally, we tested the behavior of the sensing network in the presence of the amino acids that can be found in urine. We started with the basic amino acids (Lys, Orn, Arg), because they also show high values in cystinuria patients. Gratifyingly, the dynamic sensor is able to detect **3a**

(Figure 3C) and [**3a**$_2$] (Figures S71) in the presence of these basic amino acids at concentrations in the upper limit of the expected in regular urine samples. More importantly, the sensor also provides a clear readout of **3a** or cystine in the presence of pathological concentrations of these amino acids (Figure S72). This highlights the robustness of the network sensor, which is able to detect the analytes in the presence of high concentrations of basic amino acids.

The sensor also responded to minute concentrations of both **3a** and [**3a**$_2$] in very competitive media containing most of the amino acids that can be found in urine (Lys, Orn, Arg, Asn, Gly, Ala, His, Asp, β-Ala, Ser, Tyr, Met) at their normal concentrations (Figure 3D and Figures S73–75).

Once we had evaluated the sensitivity and the selectivity of the method, we tested the system in urine samples from healthy volunteers (Supporting Information). We measured the fluorescence spectra of the urine without sensor to confirm that no other metabolites in this fluid could interfere (Figure 4A, black). We also measured the positive response of the sensor to the naturally excreted cysteine in the urine samples (light blue spectrum in Figure 4A and U samples in Figure 4B). Moreover, the addition of cysteine into these samples produced the increase of the band at 501 nm, with a detection range that goes from normally occurring Cys in urine (blue points in Figure 4B) to pathological concentrations (red in Figure 4B). Thus, we could also sense abnormal concentrations of cysteine that would not yet cause calculi (green in Figure 4B).

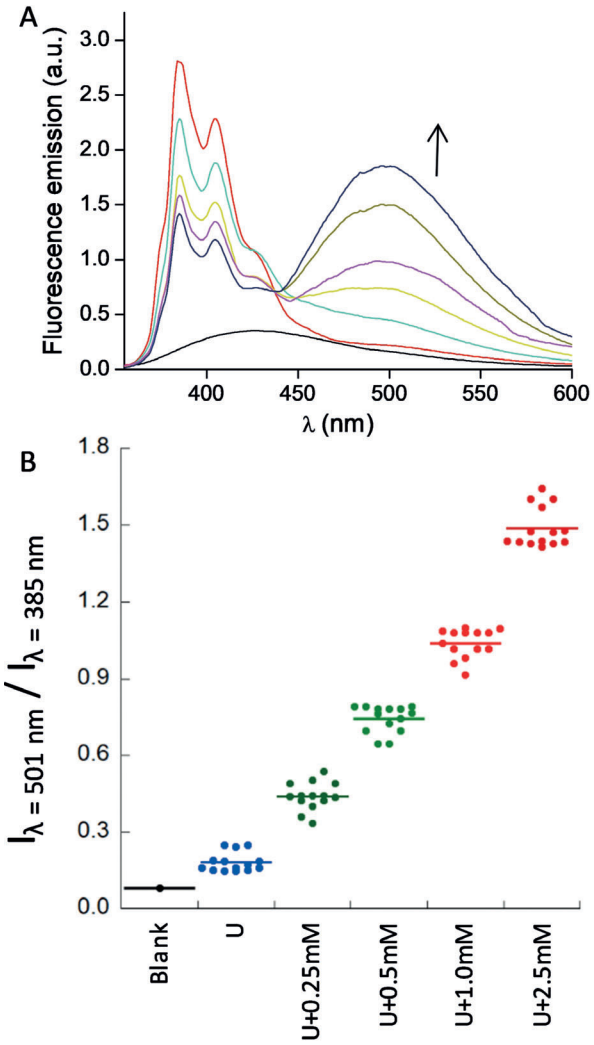


Figure 4. A) Selected normalized fluorescence spectra of urine (black), sensor alone (red), sensor + urine (light blue) and after increasing concentrations of added Cys (0.25–2.5 mM). B) Scatter plot of the excimer/monomer emission of the library alone (Blank), for different urine (U) samples of healthy volunteers and for the urine samples plus additional Cys. Samples were 34-times diluted in 1:1 H_2O:DMSO for the measurements.

In conclusion, we described a rationally designed dynamic molecular network able to selectively sense biologically relevant molecules. Instead of a discrete fluorescent probe, the sensor comprises an ensemble of species that rearranges and releases a fluorescent reporter in response to the analyte. This system works in aqueous media to selectively sense cysteine and cystine against other biothiols and amino acids, even in human urine. Our work shows the potential of chemical networks to achieve advanced functions, on the road towards "artificial" life.[19]

Acknowledgements

Financial Support from MINECO/FEDER (CTQ2015-70117-R and BES-2013-063128), AGAUR (2014 SGR 231) and European Union (COST CM1304) are gratefully acknowledged.

Conflict of interest

The authors declare no conflict of interest.

Keywords: cystinuria · dynamic networks · pseudopeptides · sensors · systems chemistry

[1] a) P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652; b) R. F. Ludlow, S. Otto, *Chem. Soc. Rev.* **2008**, *37*, 101; c) E. Mattia, S. Otto, *Nat. Nanotechnol.* **2015**, *10*, 111; d) G. Ashkenasy, T. M. Hermans, S. Otto, A. F. Taylor, *Chem. Soc. Rev.* **2017**, *46*, 2543.

[2] a) J. R. Nitschke, *Nature* **2009**, *462*, 736; b) J.-M. Lehn, *Angew. Chem. Int. Ed.* **2013**, *52*, 2836; *Angew. Chem.* **2013**, *125*, 2906; c) I. Alfonso, *Chem. Commun.* **2016**, *52*, 239.

[3] a) M. Barboiu, *Constitutional Dynamic Chemistry: Bridge from Supramolecular Chemistry to Adaptive Chemistry*, Springer, Berlin, Heidelberg, **2012**; b) J.-M. Lehn, A. V. Eliseev, *Science* **2001**, *291*, 2331; c) J. Li, P. Nowak, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 9222; d) D. Komáromy, M. Tezcan, G. Schaeffer, I. Marić, S. Otto, *Angew. Chem. Int. Ed.* **2017**, *56*, 14658; *Angew. Chem.* **2017**, *129*, 14850.

[4] a) S. Otto, *Acc. Chem. Res.* **2012**, *45*, 2200; b) S. N. Semenov, L. J. Kraft, A. Ainla, M. Zhao, M. Baghbanzadeh, V. E. Campbell, K. Kang, J. M. Fox, G. M. Whitesides, *Nature* **2016**, *537*, 656; c) A. M. Valdivielso, F. Puig-Castellvi, J. Atcher, J. Solà, R. Tauler, I. Alfonso, *Chem. Eur. J.* **2017**, *23*, 10789.

[5] a) M. I. Simon, M. P. Strathmann, N. Gautam, *Science* **1991**, *252*, 802; b) J. Schlessinger, *Cell* **2000**, *103*, 211.
[6] a) A. P. de Silva, H. Q. N. Gunaratne, T. Gunnlaugsson, A. M. Huxley, C. P. McCoy, J. T. Rademacher, T. E. Rice, *Chem. Rev.* **1997**, *97*, 1515; b) E. V. Anslyn, *J. Org. Chem.* **2007**, *72*, 687; c) L. You, D. Zha, E. V. Anslyn, *Chem. Rev.* **2015**, *115*, 7840; d) M. H. Lee, J. S. Kim, J. L. Sessler, *Chem. Soc. Rev.* **2015**, *44*, 4185.
[7] Selected examples: a) E. G. Shcherbakova, T. Minami, V. Brega, T. D. James, P. Anzenbacher, Jr., *Angew. Chem. Int. Ed.* **2015**, *54*, 7130; *Angew. Chem.* **2015**, *127*, 7236; b) R. Berardozzi, E. Badetti, N. A. C. dos Santos, K. Wurst, G. Licini, G. Pescitelli, C. Zonta, L. Di Bari, *Chem. Commun.* **2016**, *52*, 8428; c) P. Zardi, K. Wurst, G. Licini, C. Zonta, *J. Am. Chem. Soc.* **2017**, *139*, 15616; d) G. Gasparini, F. Bettin, P. Scrimin, L. J. Prins, *Angew. Chem. Int. Ed.* **2009**, *48*, 4546; *Angew. Chem.* **2009**, *121*, 4616.
[8] a) Z. A. Wood, E. Schröoder, J. R. Harris, L. B. Poole, *Trends Biochem. Sci.* **2003**, *28*, 32; b) R. O. Ball, G. Courtney-Martin, P. B. Pencharz, *J. Nutr.* **2006**, *136*, 1682S.
[9] a) A. Meister, M. E. Anderson, *Annu. Rev. Biochem.* **1983**, *52*, 711; b) C. Perricone, C. De Carolis, R. Perricone, *Autoimmun. Rev.* **2009**, *8*, 697; c) R. Franco, O. Schoneveld, A. Pappa, M. Panayiotidis, *Arch. Physiol. Biochem.* **2007**, *113*, 234.
[10] a) N. M. Giles, G. I. Giles, C. Jacob, *Biochem. Biophys. Res. Commun.* **2003**, *300*, 1; b) V. Gazit, R. Ben-Abraham, R. Coleman, A. Weizman, Y. Katz, *Amino Acids* **2004**, *26*, 163.
[11] E. Fjellstedt, L. Harnevik, J.-O. Jeppsson, H.-G. Tiselius, P. Söderkvist, T. Denneberg, *Urol. Res.* **2003**, *31*, 417.
[12] C. S. Biyani, J. J. Cartledge, *EAU-EBU Update Series* **2006**, *4*, 175.
[13] R. E. Grier, W. A. Gahl, T. Cowan, I. Bernardini, G. A. McDowell, P. Rinaldo, *Genet. Med.* **2004**, *6*, 66.
[14] a) X. Chen, Y. Zhou, X. Peng, J. Yoon, *Chem. Soc. Rev.* **2010**, *39*, 2120; b) H. S. Jung, X. Chen, J. S. Kim, J. Yoon, *Chem. Soc. Rev.* **2013**, *42*, 6019; c) L.-Y. Niu, Y.-Z. Chen, H.-R. Zheng, L.-Z. Wu, C.-H. Tung, Q.-Z. Yang, *Chem. Soc. Rev.* **2015**, *44*, 6143; d) C.-X. Yin, K.-M. Xiong, F.-J. Huo, J. C. Salamanca, R. M. Strongin, *Angew. Chem. Int. Ed.* **2017**, *56*, 13188; *Angew. Chem.* **2017**, *129*, 13368.
[15] Seleted examples: a) H. Maeda, H. Matsuno, M. Ushida, K. Katayama, K. Saeki, N. Itoh, *Angew. Chem. Int. Ed.* **2005**, *44*, 2922; *Angew. Chem.* **2005**, *117*, 2982; b) S. Sreejith, K. P. Divya, A. Ajayaghosh, *Angew. Chem. Int. Ed.* **2008**, *47*, 7883; *Angew. Chem.* **2008**, *120*, 8001; c) J. Liu, Y.-Q. Sun, Y. Huo, H. Zhang, L. Wang, P. Zhang, D. Song, Y. Shi, W. Guo, *J. Am. Chem. Soc.* **2014**, *136*, 574; d) D. Zhang, Z. Yang, H. Li, Z. Pei, S. Sun, Y. Xu, *Chem. Commun.* **2016**, *52*, 749; e) X. Sun, E. V. Anslyn, *Angew. Chem. Int. Ed.* **2017**, *56*, 9522; *Angew. Chem.* **2017**, *129*, 9650; f) M. Zhang, M. L. Saha, M. Wang, Z. Zhou, B. Song, C. Lu, X. Yan, X. Li, F. Huang, S. Yin, P. J. Stang, *J. Am. Chem. Soc.* **2017**, *139*, 5067; g) F. Y. Thanzeel, C. Wolf, *Angew. Chem. Int. Ed.* **2017**, *56*, 7276; *Angew. Chem.* **2017**, *129*, 7382.
[16] J. Atcher, J. Solà, I. Alfonso, *Org. Biomol. Chem.* **2017**, *15*, 213.
[17] M. Lafuente, J. Atcher, J. Solà, I. Alfonso, *Chem. Eur. J.* **2015**, *21*, 17002.
[18] S. Nishizawa, Y. Kato, N. Teramae, *J. Am. Chem. Soc.* **1999**, *121*, 9463.
[19] a) A. de la Escosura, C. Briones, K. Ruiz-Mirazo, *J. Theor. Biol.* **2015**, *381*, 11; b) K. Ruiz-Mirazo, C. Briones, A. de la Escosura, *Chem. Rev.* **2014**, *114*, 285.

Manuscript received: February 19, 2018
Version of record online: ■■ ■■, ■■■■

3.3.1. Electronic Supplementary Information (ESI) for:

PUBLICATION D

A dynamic chemical network for cystinuria diagnosis

Maria Lafuente, Jordi Solà and Ignacio Alfonso**

Department of Biological Chemistry and Molecular Modelling.
Institute of Advanced Chemistry of Catalonia, IQAC-CSIC
Jordi Girona 18-26, 08034, Barcelona, Spain. Fax: (+34)932045904
E-mail: ignacio.alfonso@iqac.csic.es
jordi.sola@iqac.csic.es

Table of Contents

GENERAL METHODS

General: Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka or Merck) and were used without further purification. Flash chromatographic purifications and preparative reversed-phase purifications were performed on a Biotage® Isolera Prime™ equipment. TLCs were performed using 6x3 cm SiO_2 pre-coated aluminium plates (ALUGRAM® SIL G/UV_{254}).

Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC): Analyses were performed on a Hewlett Packard Series 1100 (UV detector 1315A) modular system using:

i) For the analysis of building blocks **3b** and **3e-h**: a reversed-phase X-Terra C_{18} (15 x 0.46 cm, 5 µm) column. (CH_3CN + 0.07% TFA and H_2O + 0.1% TFA) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelengths were set at 220 nm.

ii) For the analysis of the DCLs: a reversed-phase kromaphase C_{18} (25 x 0.46 cm, 5µm) column. (CH_3CN + 20 mM HCOOH and H_2O + 20 mM HCOOH) mixtures at 1 mL/min were used as mobile phase and the monitoring wavelength was set at 254 nm.

Nuclear Magnetic Resonance (NMR): Spectroscopic experiments were carried out on a Varian Mercury 400 instrument (400 MHz for ^1H and 101 MHz for ^{13}C). Chemical shifts (δ_H) are quoted in parts per million (ppm) and referenced to the appropriate NMR solvent peak(s). 2D-NMR experiments COSY, HSQC and HMBC were used where necessary in assigning NMR spectra. Spin-spin coupling constants (*J*) are reported in Hertz (Hz).

pH measurements were performed at room temperature on a Crison GLP21 pH-meter with the electrode Crison 50 14T.

High Resolution Mass Spectrometry (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.

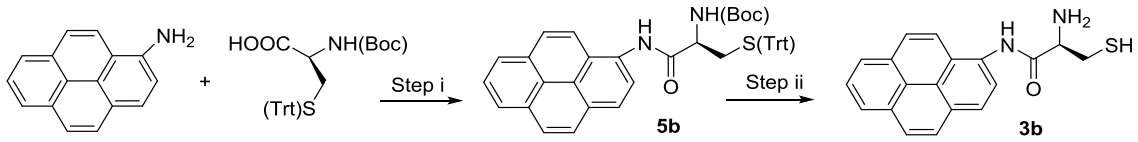
Absorbance measurements were performed on a Molecular Devices SpectraMax M5 microplate reader, at room temperature, and the monitoring wavelength was set at 412 nm. The 96 well microplates, PS, F-bottom, 655101 were used to place the samples when performing the absorbance measurements.

Fluorescence spectroscopy: Fluorescence emission and excitation spectra were collected on a Photon Technology International Instrument, the Fluorescence Master Systems, using the Software Felix32 and cuvettes with 10 mm path length.

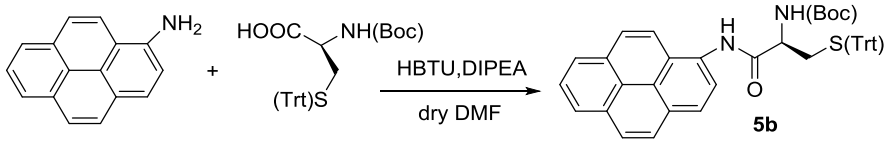
SYNTHESIS OF THE BUILDING BLOCKS

The trithiol [**1**] was synthesized as previously described[1] and the dithiol [**2**] was synthesized as we previously reported.[2] The monothiols [**3a, c, d**] were purchased from commercial suppliers (Sigma-Aldrich and Iris Biotech). Also the tritylsulfanyl acetic acid was prepared as previously described[3].

Synthetic scheme for the preparation of 3b



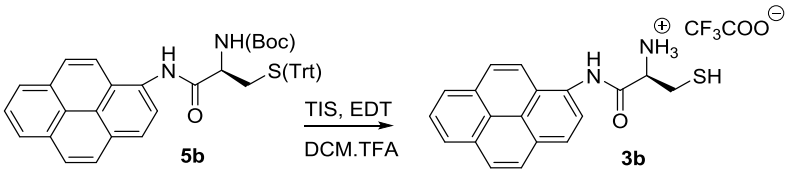
Step i: Experimental procedure for the synthesis of 5b



Boc-Cys(Trt)-OH (2.134 g, 4.60 mmol) was dissolved in dry DMF (6 mL) and then HBTU (2.094 g, 5.52 mmol) and HOBt (0.746 g, 5.52 mmol) were added. The reaction mixture was stirred during 2 min and after 1-aminopyrene (200 mg, 0.92 mmol) was added. The solution was stirred at room temperature under an inert atmosphere of Ar for 5 hours (TLC monitoring). The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, dried over MgSO4 and concentrated under reduced pressure. The crude product was purified by flash chromatography using EtOAc/hexane/toluene as eluent (from 20/60/20 to 40/40/20) to give 311 mg (51% yield) of **5b** as a light pink solid.

Rf of **5b** in EtOAc/hexane, 3:7 (v/v)= 0.42.
HRMS (ESI$^-$) calcd. for $C_{43}H_{38}N_2O_3S$ $[M-H]^-$ (m/z): 661.2575, found: 661.2525.
^{1}H NMR (400 MHz, $CDCl_3$-*d*): δ = 8.90 (s, 1H, NHCOC*H), 8.46 (d, *J* = 8.1 Hz, 1H, CH_{Ar}), 8.19 – 8.11 (m, 3H, CH_{Ar}), 8.09 – 7.96 (m, 4H, CH_{Ar})), 8.00 (s, 1H, CH_{Ar}), 7.52 (dd, *J* = 7.4, 1.7 Hz, 6H, CH_{Ar}), 7.33 (dd, *J* = 8.4, 6.8 Hz, 6H, CH_{Ar}), 7.26 (d, *J* = 3.7 Hz, 3H, CH_{Ar}), 4.98 (d, *J* = 7.1 Hz, 1H, C*HNHCO), 4.24 – 4.11 (m, 1H, C*H), 2.89 (ddd, *J* = 71.3, 13.1, 6.2 Hz, 2H, C*HCH_2), 1.53 (s, 9H, CH_3) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ=169. 6 (CO), 156.2 (CO), 144.6 (3 x C_{Ar}), 131.5 (C_{Ar}), 130.9 (C_{Ar}), 129.8 (6 x C_{Ar}), 128.3 (8 x C_{Ar}), 128.1 (C_{Ar}), 127.9 (C_{Ar}), 127.5 (C_{Ar}), 127.1 (3 x C_{Ar}), 126.9 (C_{Ar}), 126.2 (C_{Ar}), 125.5 (C_{Ar}), 125.3 (C_{Ar}), 125.2 (C_{Ar}), 125.1 (C_{Ar}), 124.9 (C_{Ar}), 121.6 (C_{Ar}), 120.5 (C_{Ar}), 81.1 (C(CH_3)$_3$), 67.7 (CPh_3), 54.6 (C*H) , 33.2 (CHCH_2S), 28.5 (3 x CH_3) ppm.

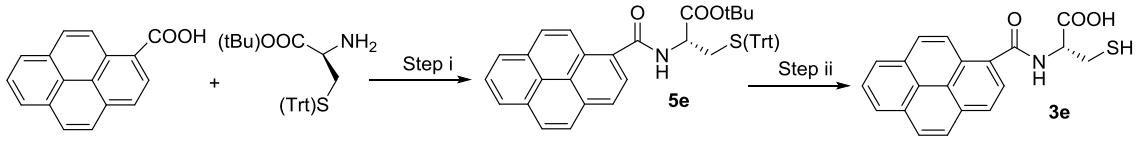
Step ii: Experimental procedure for the synthesis of 3b



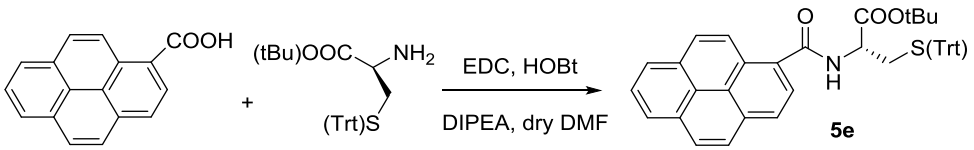
To a solution of **5b** (280 mg, 0.42 mmol) in DCM (1.5 mL), 1 mL of trifluoroacetic acid (TFA), 1,2-Ethanedithiol (EDT, 0.64 mL, 7.61 mmol) and triisopropylsilane (TIS, 1.30 mL, 6.34 mmol) were added rapidly and under stirring. The reaction mixture was stirred at room temperature for 2 hours, after 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. The product was purified by reversed-phase flash chromatography using a mixture of MeCN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 0% to 10% MeCN in H_2O). After lyophilization 75 mg (55% yield) of **3b**·1TFA were obtained as a white solid.

HRMS (ESI$^+$) calcd. for $C_{19}H_{16}N_2OS$ $[M+H]^+$ (m/z): 321.1061, found: 321.1041.
^{1}H NMR (400 MHz, MeOH-d_4): δ = 8.29 – 8.17 (m, 6H, H_5, H_6, H_8, H_{10}, H_{12}, H_{13}), 8.11 (q, *J* = 1.5 Hz, 2H, H_2, H_3), 8.05 (t, *J* = 7.6 Hz, 1H, H_9), 4.48 (dd, *J* = 6.9, 5.4 Hz, 1H, C*HCH_2), 3.38- 3.24 (dd, *J* = 14.6, 5.4 Hz, 2H, C*HCH_2) ppm.
^{13}C NMR (101 MHz, MeOH-d_4): δ = 168.2 (CO), 132.7 (C_{Ar}), 132.2 (C_{Ar}), 131.5 (C_{Ar}), 130.7 (C_{Ar}), 129.2 (C_{Ar}), 128.6 (C_{Ar}), 128.2 (C_{Ar}), 127.7 (C_{Ar}), 127.5 (C_{Ar}), 126.8 (C_{Ar}), 126.6 (C_{Ar}), 126.5 (C_{Ar}), 126.0 (C_{Ar}), 125.6 (C_{Ar}), 124.9 (C_{Ar}), 122.5 (C_{Ar}), 56.8 (C*H), 26.5 (CH_2SH) ppm.

Synthetic scheme for the preparation of 3e



Step i: Experimental procedure for the synthesis of 5e



To a solution of pyrene carboxylic acid (250mg, 1.06 mmol) in dry DMF (6 mL) EDC·HCl (681 mg, 3.55 mmol), HOBt (562 mg, 4.16 mmol) and DIPEA (1.25 mL, 9.13 mmol) were added. The reaction mixture was stirred during 2 min and H-Cys(Trt)-OtBu (468 mg, 1.12 mmol) was then added. The solution was stirred at room temperature under an inert atmosphere of Ar for 48 hours, and the formation of the product was followed by TLC. The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, and concentrated by distillation under reduced pressure. The crude product was purified by flash

chromatography using EtOAc/Hexane as eluent (from 10% to 30% EtOAc to give 470 mg (77% yield) of **5e** as a yellow solid.

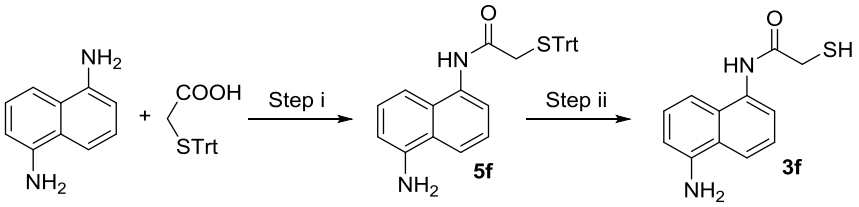
HRMS (ESI$^+$) calcd. for $C_{43}H_{37}NO_3S$ $[M+H]^+$ (m/z): 647.8330, found: 647.8310.
^{1}H NMR (400 MHz, $CDCl_3$-*d*): δ= 8.72 (d, *J* = 9.3 Hz, 1H, CONHC*H), 8.32 – 8.08 (m, 8H, CH_{Ar}), 7.47 (dd, *J* = 7.5, 1.8 Hz, 6H, CH_{Ar}), 7.32 – 7.15 (m, 9H, CH_{Ar}), 6.64 (d, *J* = 7.8 Hz, 1H, CH_{Ar}), 4.98 (dt, *J* = 7.9, 4.9 Hz, 1H, C*H), 2.96 – 2.81 (m, 2H, C*HCH_2), 1.58 (s, 9H, CH_3) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ = 169.6 (CO), 169.3 (CO), 144.5 (3 x C_{Ar}), 132.9 (C_{Ar}), 131.3 (C_{Ar}), 130.9 (C_{Ar}), 130.5 (C_{Ar}), 129.6 (6 x C_{Ar}), 129.0 (C_{Ar}), 128.9 (C_{Ar}), 128.9 (C_{Ar}), 128.1 (6 x C_{Ar}), 127.3 (C_{Ar}), 127.0 (3 x C_{Ar}), 126.5 (C_{Ar}), 126.0 (C_{Ar}), 125.9 (C_{Ar}), 125.1 (C_{Ar}), 124.8 (2 x C_{Ar}), 124.5 (2 x C_{Ar}), 83.0 (C(CH_3)$_3$), 66.9 ($CPhe_3$), 52.4 (C*HCH_2), 34.5 (CH2STrt), 28.2 (3 x CH_3) ppm.

Step ii: Experimental procedure for the synthesis of 3e

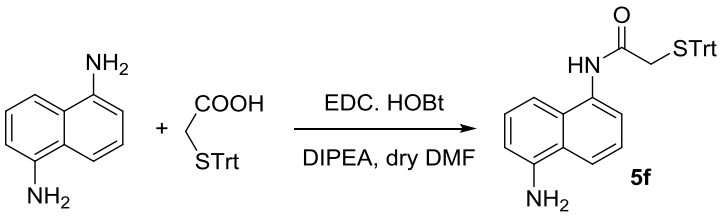
To a solution of **5e** (215 mg, 0.33 mmol) in DCM (1.5 mL), 1 mL of trifluoroacetic acid (TFA), 1,2-Ethanedithiol (EDT, 0.5 mL, 5.97 mmol) and triisobutylsilane (TIS, 1.30 mL, 4.98 mmol) were added rapidly and under stirring. The reaction mixture was stirred at room temperature for 2 hours, after which 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. The product was purified by reversed-phase flash chromatography using a mixture of MeCN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 0% to 10% MeCN in H_2O). After lyophilisation 67 mg (58% yield) of **3e** were obtained as a white solid.

HRMS (ESI$^+$) calcd. for $C_{20}H_{15}NO_3S$ $[M+H]^+$ (m/z): 350.0851, found: 350.0839.
^{1}H NMR (400 MHz, DMSO-d_6) δ = 13.04 (s, 1H, OH), 8.99 (d, *J* = 7.8 Hz, 1H, CONH), 8.60 (d, *J* = 9.2 Hz, 1H, CH_{Ar}), 8.36 (dd, *J* = 7.8, 2.8 Hz, 3H, CH_{Ar}), 8.30 – 8.22 (m, 3H, CH_{Ar}), 8.19 (d, *J* = 8.0 Hz, 1H, CH_{Ar}), 8.13 (t, *J* = 7.6 Hz, 1H, CH_{Ar}), 4.71 (td, *J* = 8.4, 4.3 Hz, 1H, C*H), 3.14 – 2.89 (m, 2H, CH_2SH), 2.70 (s, 1H, SH) ppm.
^{13}C NMR (101 MHz, DMSO-d_6) δ = 171.8 (COOH), 169.2 (CONH), 131.7 (C_{Ar}), 131.4 (C_{Ar}), 130.7 (C_{Ar}), 130.2 (C_{Ar}), 128.4 (C_{Ar}), 128.1 (C_{Ar}), 127.9 (C_{Ar}), 127.2 (C_{Ar}), 126.6 (C_{Ar}), 125.8 (C_{Ar}), 125.7 (C_{Ar}), 125.4 (C_{Ar}), 124.7 (C_{Ar}), 124.4 (C_{Ar}), 123.7 (C_{Ar}), 123.6 (C_{Ar}), 55.7 (C*H), 25.3 (CH_2SH) ppm.

Synthetic scheme for the preparation of 3f



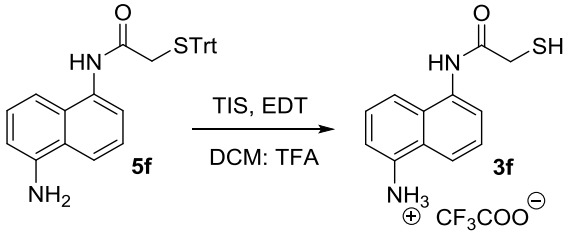
Step i: Experimental procedure for the synthesis of 5f



Tritylsulfanyl acetic acid (500mg, 1.50 mmol) was dissolved in dry DMF (6 mL) and EDC·HCl (383 mg, 2 mmol), HOBt (304 mg, 2.25 mmol) and DIPEA (0.87 mL, 5 mmol) were added over the solution. The reaction mixture was stirred during 2 min and 1,5-diaminonaphthalene (158 mg, 1 mmol) was added over the mixture. The solution was stirred at room temperature under an inert atmosphere of Ar for 48 hours, and the formation of the product was followed by TLC. The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, and dried under reduced pressure. The crude product was purified by flash chromatography using EtOAc/Hexane as eluent (from 30% to 50% EtOAc) to give 287 mg (61% yield) of **5f** as a pink solid.

Rf of 5f in EtOAc/hexane 2:3 (v/v): 0.46.
HRMS (ESI$^+$) calcd. for $C_{31}H_{26}N_2OS$ $[M+H]^+$ (m/z): 475.1844, found: 475.1815.
^{1}H NMR (400 MHz, CDCl3-*d*): δ= 8.44 (s, 1H, CH2CONH), 7.83 (d, *J* = 7.4 Hz, 1H, CH_{Ar}), 7.61 (d, *J* = 8.5 Hz, 1H, CH_{Ar}), 7.48 (d, *J* = 7.7 Hz, 6H, CH_{Ar}), 7.37 (t, *J* = 8.0 Hz, 1H, CH_{Ar}), 7.30–7.23 (m, 7H, CH_{Ar}), 7.18 (t, *J* = 7.3 Hz, 3H, CH_{Ar}), 7.06 (d, *J* = 8.5 Hz, 1H, CH_{Ar}), 6.78 (d, *J* = 7.4 Hz, 1H, CH_{Ar}), 3.46 (s, 2H, CH_2CONH) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ= 166.3 (CONH), 143.9 (2 x C_{Ar}), 142.8 (C_{Ar}), 132.5 (C_{Ar}), 129.6 (6 x C_{Ar}), 128.5 (6 x C_{Ar}), 127.4 (5 x C_{Ar}), 126.8 (C_{Ar}), 124.6 (C_{Ar}), 124.2 (C_{Ar}), 119.4 (C_{Ar}), 118.0 (C_{Ar}), 111.2 (C_{Ar}), 110.0 (C_{Ar}) , 68.6 ($CPhe_3$), 37.1 (CH_2STrt) ppm.

Step ii: Experimental procedure for the synthesis of 3f

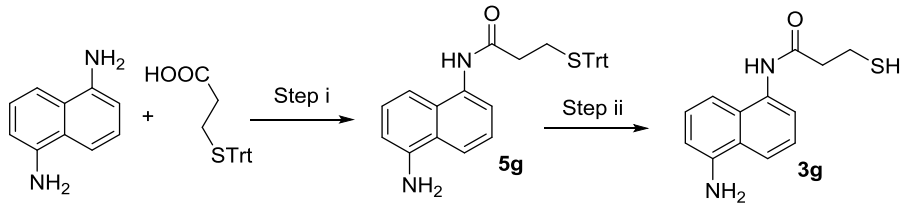


To a solution of **5f** (166 mg, 0.35 mmol) in DCM (1 mL), 1 mL of trifluoroacetic acid (TFA), 1,2-Ethanedithiol (EDT, 0.53 mL, 6.30 mmol) and triisopropylsilane (TIS, 1.07 mL, 5.24 mmol) were added rapidly and under stirring. The reaction mixture was stirred

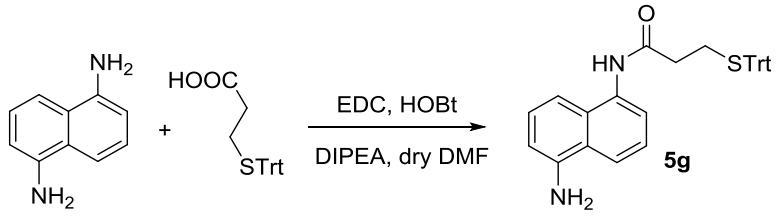
at room temperature for 2 hours, after which the solvents were partially evaporated using a N_2 flow. Diethyl ether was added over the reaction mixture and the product was filtered off and washed with diethyl ether. The product was purified by reversed-phase flash chromatography using a mixture of MeCN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 0% to 8% MeCN in H_2O). After lyophilisation 60 mg (51% yield) of **3f**·1TFA were obtained as a white solid.

HRMS (ESI$^+$) calcd. for $C_{12}H_{12}N_2OS$ $[M+H]^+$ (m/z): 233.3035, found: 233.3015.
^{1}H NMR (400 MHz, MeOH-d_4): δ = 7.96 (d, *J* = 8.6 Hz, 1H, CH_{Ar}), 7.88 (d, *J* = 8.5 Hz, 1H, CH_{Ar}), 7.74 (d, *J* = 7.4 Hz, 1H, CH_{Ar}), 7.63 (t, *J* = 7.9 Hz, 1H, CH_{Ar}), 7.53 (t, *J* = 8.1 Hz, 1H, CH_{Ar}), 7.41 (d, *J* = 7.3 Hz, 1H, CH_{Ar}), 3.48 (s, 2H, $C\underline{H_2}SH$) ppm.
^{13}C NMR (101 MHz, MeOH-d_4): δ= 172.7 (CONH), 134.6 (C_{Ar}), 133.1 (C_{Ar}), 130.8 (C_{Ar}), 127.8 (C_{Ar}), 127.4 (C_{Ar}), 127.0 (C_{Ar}), 124.7 (C_{Ar}), 121.5 (C_{Ar}), 120.3 (C_{Ar}), 118.8 (C_{Ar}), 28.8 (CH_2SH) ppm.

Synthetic scheme for the preparation of 3g



Step i: Experimental procedure for the synthesis of 5g

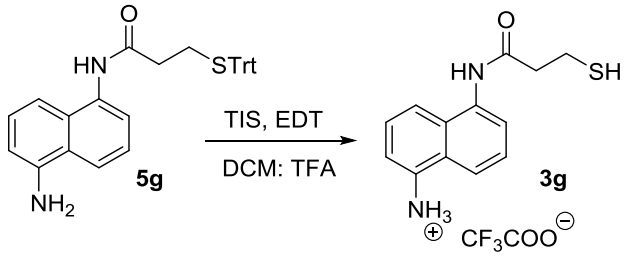


To a solution of 3-(Tritylsulfanyl)propanoic acid (500mg, 1.43 mmol) in dry DMF (10 mL), EDC·HCl (383 mg, 2 mmol), HOBt (304 mg, 2.25 mmol) and DIPEA (0.87 mL, 5 mmol) were added. The reaction mixture was stirred during 2 min and 1,5-diaminonaphthalene (158 mg, 1 mmol) was added over the mixture. The solution was stirred at room temperature under an inert atmosphere of Ar for 48 hours, and the formation of the product was followed by TLC. The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, and dried under reduced pressure. The crude product was purified by flash chromatography using EtOAc/Hexane as eluent (from 30% to 50% EtOAc) to give 301 mg (62% yield) of **5g** as a dark pink solid.

Rf of **5g** in EtOAc/Hexane, 1:1 (v/v): 0.42
HRMS (ESI$^+$) calcd. for $C_{32}H_{28}N2OS$ $[M+H]^+$ (m/z): 489.2001, found: 489.2003.
^{1}H NMR (400 MHz, DMSO-*d*): δ= 9.70 (s, 1H, CON$\underline{H}$), 7.90 (d, *J* = 8.4 Hz, 1H, CH_{Ar}), 7.54 (d, *J* = 7.2 Hz, 1H, CH_{Ar}), 7.38 – 7.32 (m, 12H, CH_{Ar}), 7.30 – 7.17 (m, 6H, CH_{Ar}), 6.69 (dd, *J* = 6.8, 1.7 Hz, 1H, CH_{Ar}), 5.73 (s, 2H, $N\underline{H_2}$), 2.56 (t, *J* = 7.3 Hz, 2H, $COC\underline{H_2}CH_2$), 2.39 (t, *J* = 7.2 Hz, 2H, $CH_2C\underline{H_2}STrt$) ppm.
^{13}C NMR (101 MHz, DMSO-d_6): δ = 169.6 (CONH), 145.0 (C_{Ar}), 144.4 (3 x C_{Ar}), 133.2 (C_{Ar}), 129.1 (6 x C_{Ar}), 128.0 (6 x C_{Ar}), 126.7 (3 x C_{Ar}), 126.5 (C_{Ar}), 124.3 (C_{Ar}),

123.3 (C_{Ar}), 122.9 (C_{Ar}), 121.6 (C_{Ar}), 119.7 (C_{Ar}), 110.1 (C_{Ar}), 107.6 (C_{Ar}), 66.0 ($CPhe_3$), 34.6 ($CO\underline{C}H_2CH_2$), 27.6 ($CH_2\underline{C}H_2STrt$) ppm.

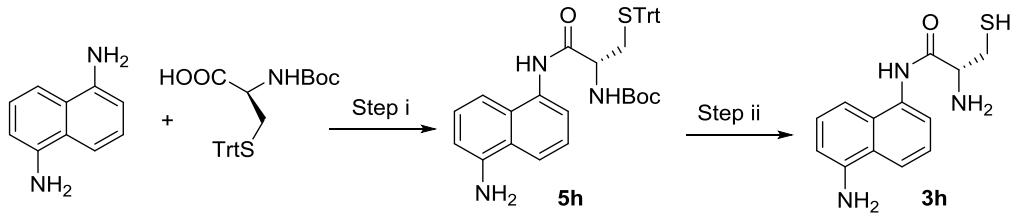
Step ii: Experimental procedure for the synthesis of 3g



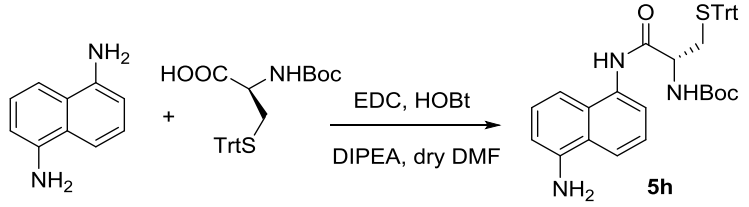
To a solution of **5g** (240 mg, 0.49 mmol) in DCM (1 mL), 1 mL of trifluoroacetic acid (TFA), 1,2-Ethanedithiol (EDT, 0.74 mL, 8.84 mmol) and triisobutylsilane (TIS, 1.52 mL, 5.89 mmol) were added rapidly and under stirring. The reaction mixture was stirred at room temperature for 2 hours, after which 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. The product was purified by reversed-phase flash chromatography using a mixture of MeCN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 0% to 10% MeCN in H_2O). After lyophilisation 107 mg (61% yield) of **3g**·1TFA were obtained as a white solid.

HRMS (ESI$^+$) calcd. for $C_{13}H_{14}N_2OS$ $[M+H]_+$ (m/z): 233.3035, found: 233.3015.
^{1}H NMR (400 MHz, MeOH-d_4): δ = 8.01 (d, *J* = 8.5 Hz, 1H, CH_{Ar}), 7.87 (d, *J* = 8.3 Hz, 1H, CH_{Ar}), 7.69 (d, *J* = 7.2 Hz, 1H, CH_{Ar}), 7.63 (t, *J* = 8.1, 7.6 Hz, 1H, CH_{Ar}), 7.52 (t, *J* = 8.3, 7.6 Hz, 1H, CH_{Ar}), 7.43 (d, *J* = 7.3 Hz, 1H, CH_{Ar}), 3.28 (dt, *J* = 3.5, 1.7 Hz, 2H, $COC\underline{H}_2CH_2$), 2.86 (d, *J* = 4.8 Hz, 2H, $CH_2C\underline{H}_2SH$) ppm.
^{13}C NMR (101 MHz, MeOH- d_4): δ= 173.6 (CONH), 135.1 (C_{Ar}), 132.2 (C_{Ar}), 131.2 (C_{Ar}), 128.2 (C_{Ar}), 127.8 (C_{Ar}), 126.9 (C_{Ar}), 125.4 (C_{Ar}), 122.9 (C_{Ar}), 120.4 (C_{Ar}), 119.6 (C_{Ar}), 41.3 ($CO\underline{C}H_2CH_2$), 21.0 ($CH_2\underline{C}H_2SH$) ppm.

Synthetic scheme for the preparation of 3h



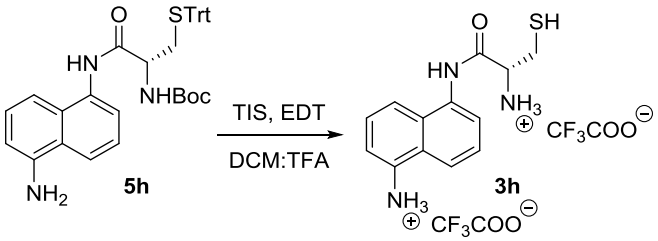
Step i: Experimental procedure for the synthesis of 5h



To a solution of Boc-Cys(Trt)-OH (1 g, 2.15 mmol) in dry DMF (10 mL) and EDC·HCl (0.723 g, 3.77 mmol), HOBt (0.597 g, 4.42 mmol) and DIPEA (1.32 mL, 9.7 mmol) were added. The reaction mixture was stirred during 2 min and 1,5-diaminonaphthalene (170 mg, 1.08 mmol) was added over the mixture. The solution was stirred at room temperature under an inert atmosphere of Ar for 48 hours, and the formation of the product was followed by TLC. The mixture was diluted with DCM, washed with saturated aqueous $NaHCO_3$ and saturated aqueous NaCl, and dried under reduced pressure. The crude product was purified by flash chromatography using EtOAc/Hexane as eluent (from 30% to 50% EtOAc) to give 356 mg (55% yield) of **5h** as a dark pink solid.

Rf of 5h in EtOAc/Hexane, 1:1 (v/v): 0.43.
HRMS (ESI$^+$) calcd. for $C_{37}H_{37}N_3O_3S$ $[M+H]^+$ (m/z): 604.2634, found: 604.2637.
^{1}H NMR (400 MHz, $CDCl_3$-*d*): δ= 8.51 (s, 1H, CONH), 7.99 (d, *J* = 7.5 Hz, 1H, CH_{Ar}), 7.62 (d, *J* = 8.5 Hz, 1H, CH_{Ar}), 7.46 (d, *J* = 7.5 Hz, 5H), 7.39 (t, *J* = 7.8 Hz, 1H), 7.32 – 7.19 (m, 12H, CH_{Ar}, CONH), 6.80 – 6.73 (m, 2H, CH_{Ar}), 4.90 (d, *J* = 7.1 Hz, 1H, C*HCH$_2$), 2.80 (ddd, *J* = 72.7, 13.3, 6.2 Hz, 2H, CH$_2$STrt), 1.46 (s, 9H, CH$_3$) ppm.
^{13}C NMR (101 MHz, $CDCl_3$-*d*): δ = 169.2 (CONHC$_{Ar}$), 144.5 (COOC(CH$_3$)$_3$), 142.8 (C$_{Ar}$), 129.7 (6 x C$_{Ar}$), 128.3 (6 x C$_{Ar}$), 127.1 (3 x C$_{Ar}$), 126.8 (C$_{Ar}$), 125.4 (2 x C$_{Ar}$), 124.7 (C$_{Ar}$), 124.3 (C$_{Ar}$), 120.0 (C$_{Ar}$), 118.2 (C$_{Ar}$), 111.7 (2 x C$_{Ar}$), 111.5 (C$_{Ar}$), 110.0 (2 x C$_{Ar}$), 81.0 (OC(CH$_3$)$_3$), 67.6 (CPhe$_3$), 54.5 (C*HCH$_2$), 33.2 (CH$_2$STrt), 28.5 (3 x CH$_3$) ppm.

Step ii: Experimental procedure for the synthesis of 3h



To a solution of **5h** (200 mg, 0.191 mmol) in DCM (1 mL), 1 mL of trifluoroacetic acid (TFA), 1,2-Ethanedithiol (EDT, 0.3 mL, 3.44 mmol) and triisopropylsilane (TIS, 0.6 mL, 2.29 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. The product was purified by reversed-phase flash chromatography using a mixture of MeCN + 0.07% (v/v) TFA and H_2O + 0.1% (v/v) TFA as mobile phase (gradient: from 0% to 10% MeCN in H_2O). After lyophilisation 77 mg (83% yield) of **3h**·2TFA were obtained as a white solid.

HRMS (ESI$^+$) calcd. for $C_{31}H_{15}N_3OS$ $[M+H]^+$ (m/z): 262.1014, found: 262.1001.
^{1}H NMR (400 MHz, MeOH-d_4): δ = 7.92 (d, *J* = 8.5 Hz, 2H, CH_{Ar}), 7.75 (d, *J* = 7.4 Hz, 1H, CH_{Ar}), 7.61 (t, *J* = 8.0 Hz, 1H, CH_{Ar}), 7.50 (t, *J* = 8.0 Hz, 1H, CH_{Ar}), 7.36 (d, *J* = 7.4 Hz, 1H, CH_{Ar}), 4.38 (t, *J* = 6.1 Hz, 1H, C*$\underline{H}CH_2$), 3.30 – 3.12 (dd, 2H, $C\underline{H_2}SH$) ppm.
^{13}C NMR (101 MHz, MeOH-d_4): δ = 168.2 (CONH), 134.5 (C_{Ar}), 133.9 (C_{Ar}), 130.8 (C_{Ar}), 127.9 (C_{Ar}), 127.4 (C_{Ar}), 127.2 (C_{Ar}), 125.0 (C_{Ar}), 121.2 (C_{Ar}), 120.8 (C_{Ar}), 118.4 (C_{Ar}), 56.7 ($\underline{C}$*HCH_2), 26.4 ($\underline{C}H_2SH$) ppm.

NMR spectra, HRMS (ESI+) spectrum and HPLC trace of 3b

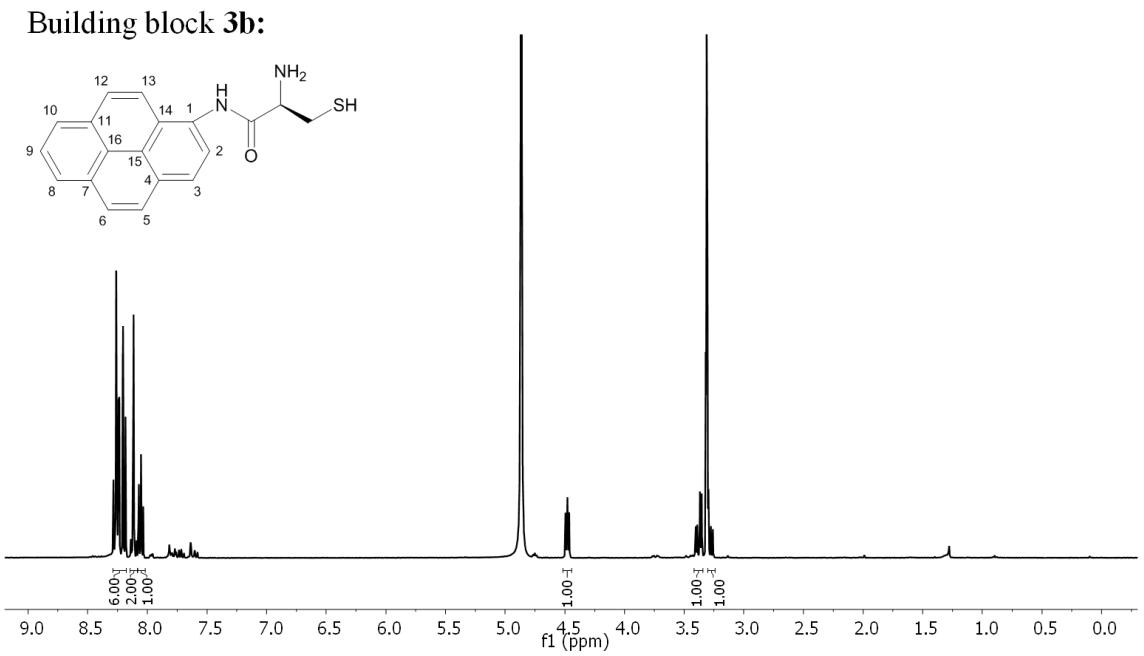


Figure S1. ^{1}H (400 MHz, 298 K in MeOD-d_4) spectrum of **3b**.

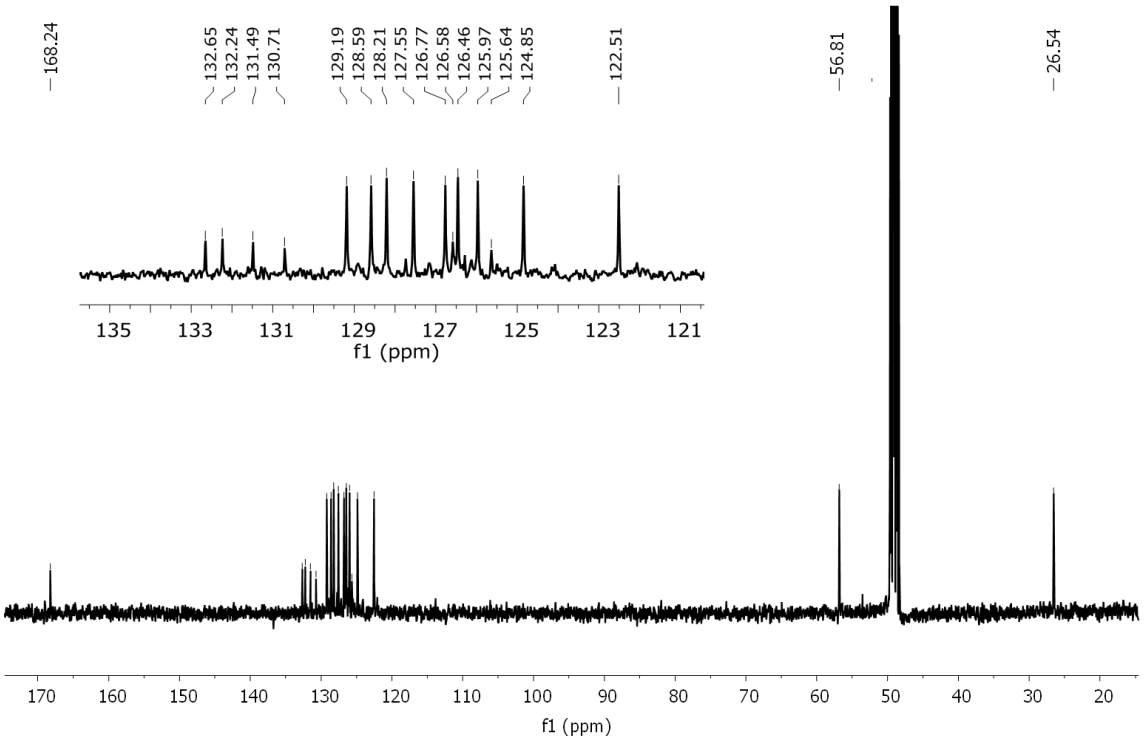


Figure S2. ^{13}C (101 MHz, 298 K in MeOD-d_4) spectrum of **3b**.

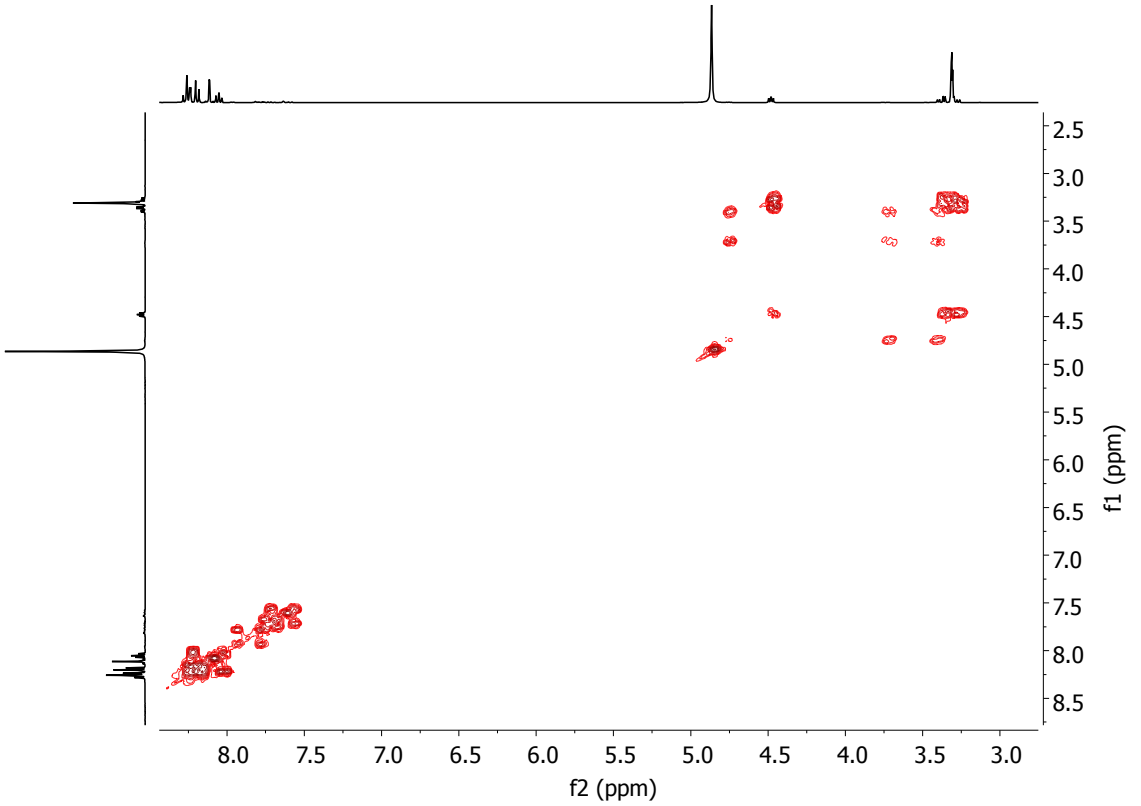


Figure S3. ^{1}H-^{1}H gCOSY (400 MHz, 298 K in MeOD-d_4) spectrum of **3b**.

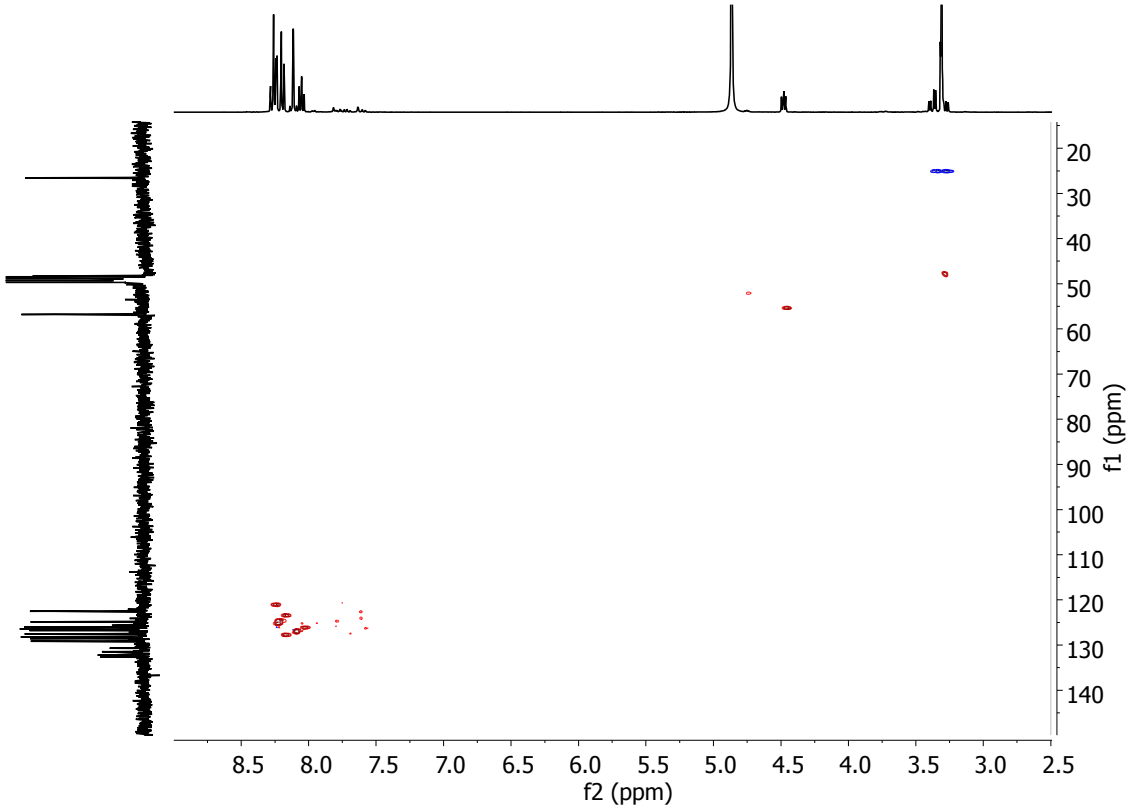


Figure S4. ^{1}H-^{13}C gHSQC (400 MHz, 298 K in MeOD-d_4) spectrum of **3b**.

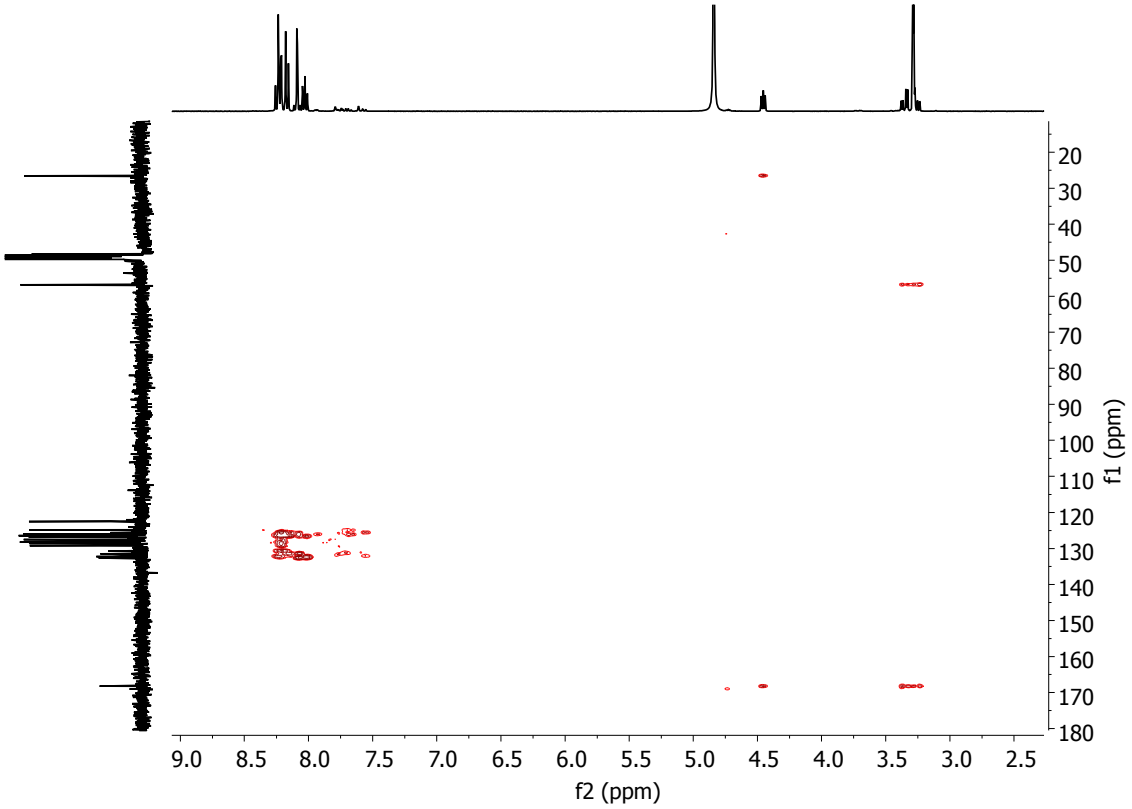


Figure S5. ^{1}H-^{13}C gHMBC (400 MHz, 298 K in MeOD-d_4) spectrum of **3b**.

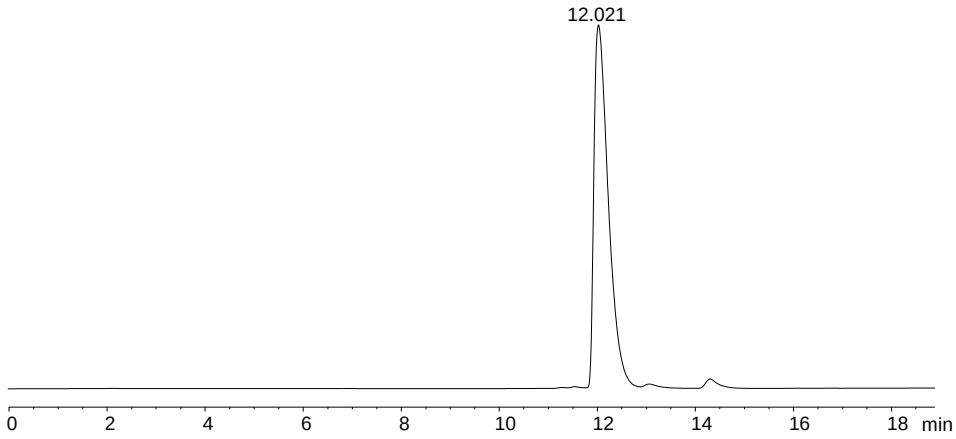


Figure S6. Analytical HPLC-trace of **3b** (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

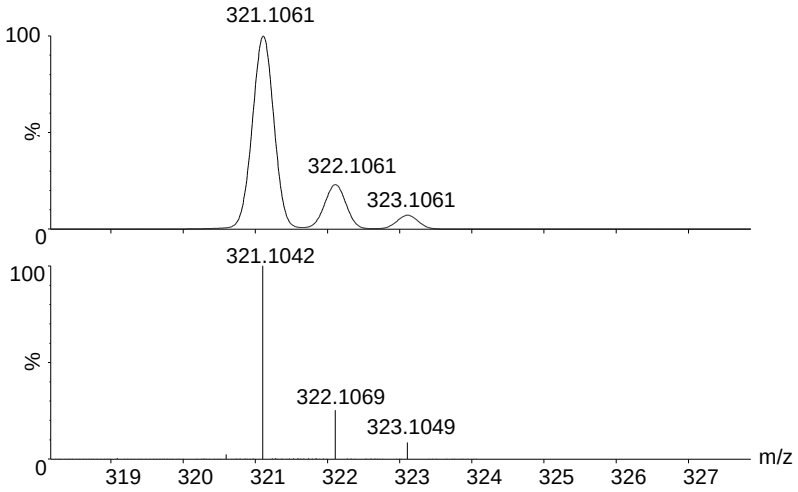
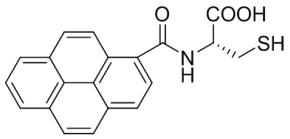


Figure S7. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **3b**.

NMR spectra, HRMS (ESI+) spectrum and HPLC trace of 3e

Building block **3e:**



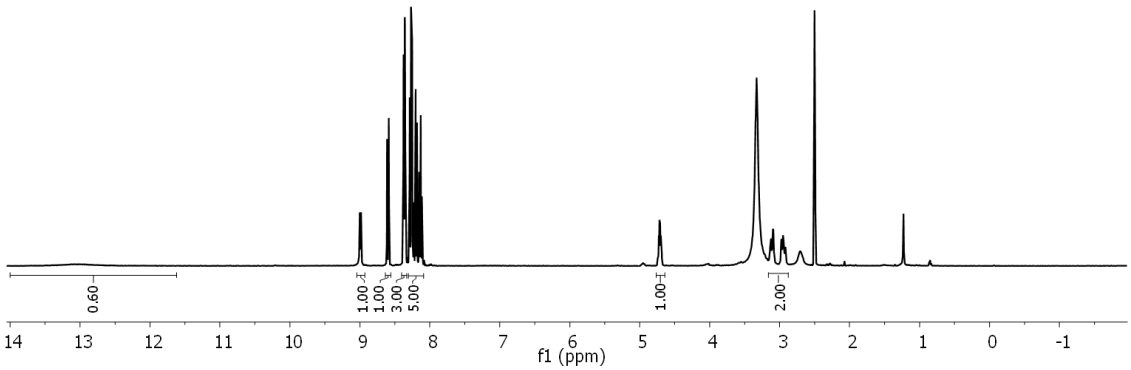


Figure S8. ^{1}H (400 MHz, 298 K in $(CD_3)_2SO$-d_6) spectrum of **3e**.

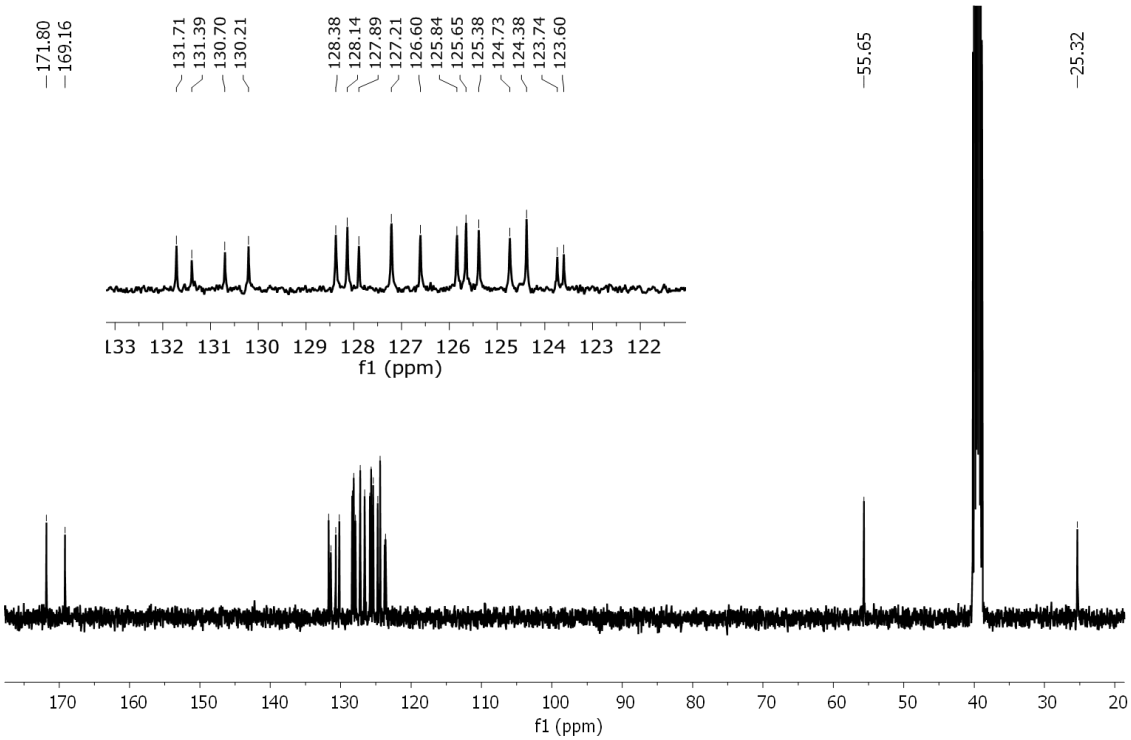


Figure S9. ^{13}C (101 MHz, 298 K in $(CD_3)_2SO$-d_6) spectrum of **3e**.

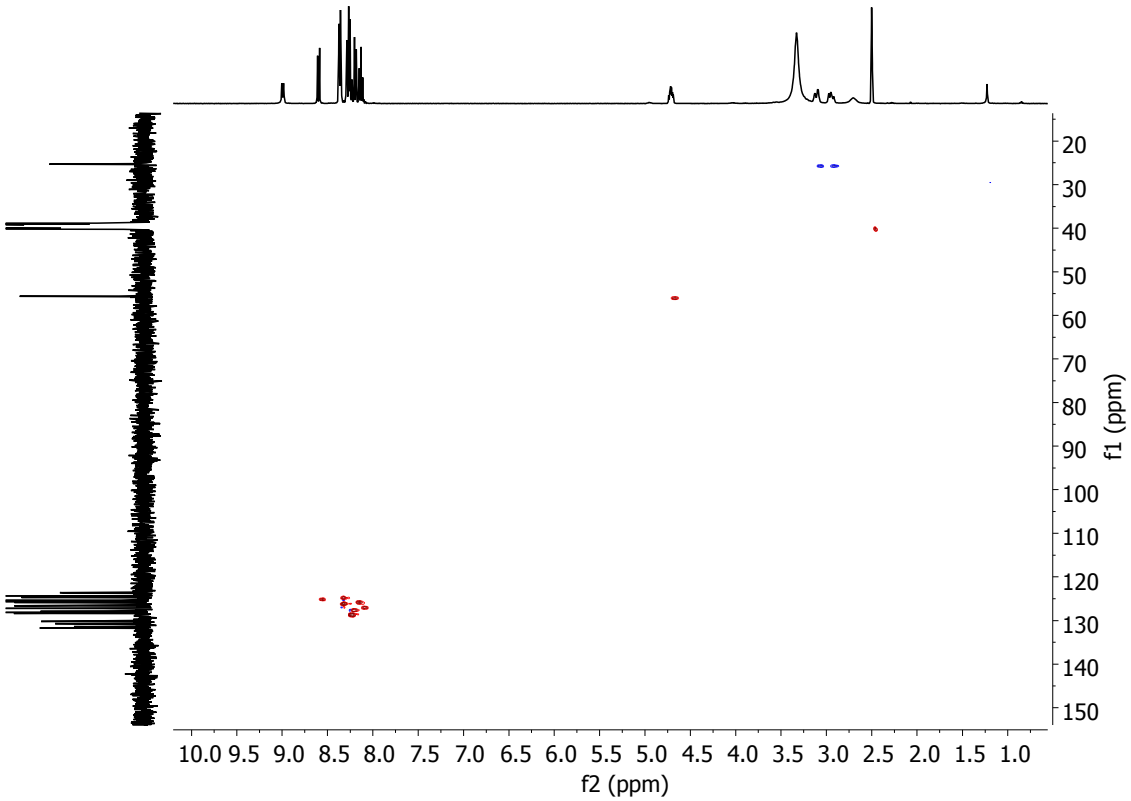


Figure S10. ^{1}H-^{13}C gHSQC (400 MHz, 298 K in $(CD_3)_2SO$-d_6) spectrum of **3e**.

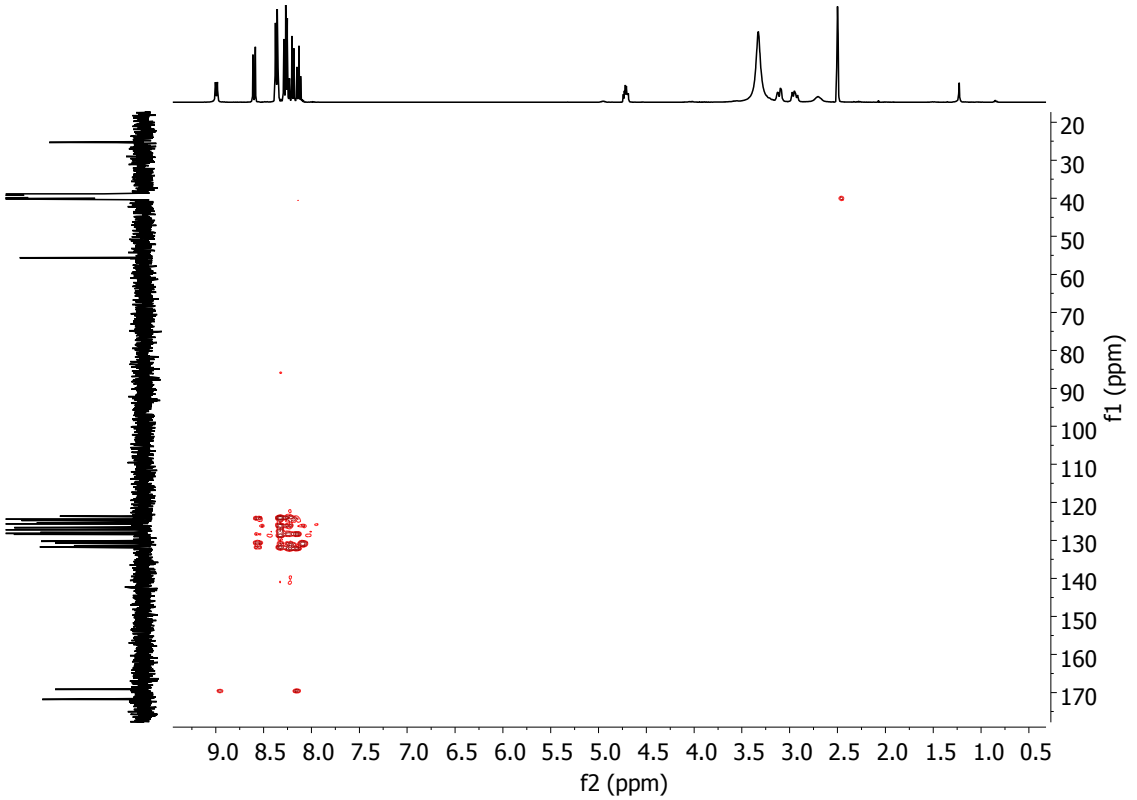


Figure S11. ^{1}H-^{13}C gHMBC (400 MHz, 298 K in $(CD_3)_2SO$-d_6) spectrum of **3e**.

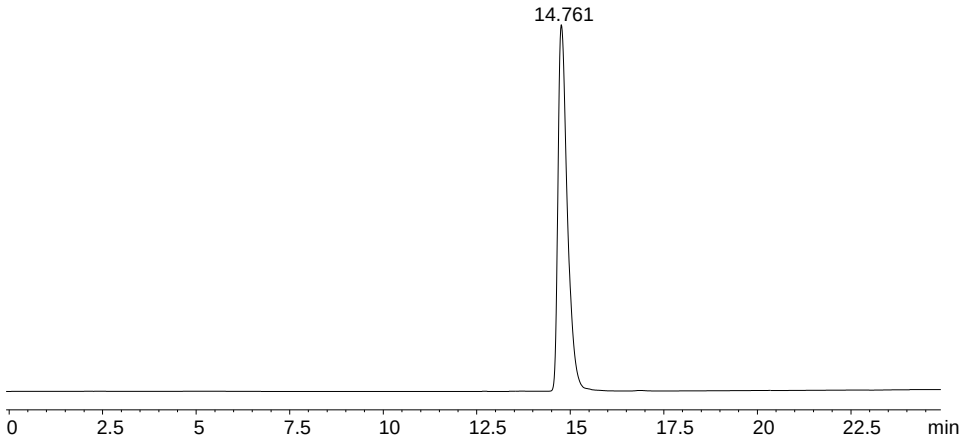


Figure S12. Analytical HPLC traces of **3e** (2 min at 1% CH_3CN in H_2O, then linear gradient from 1% to 100% CH_3CN over 20 min).

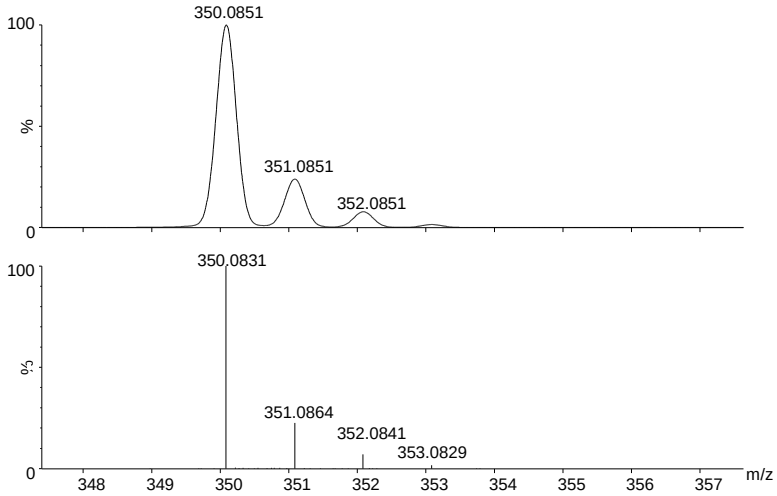


Figure S13. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **3e**.

NMR spectra, HRMS (ESI+) spectrum and HPLC trace of 3f

Building block **3f:**

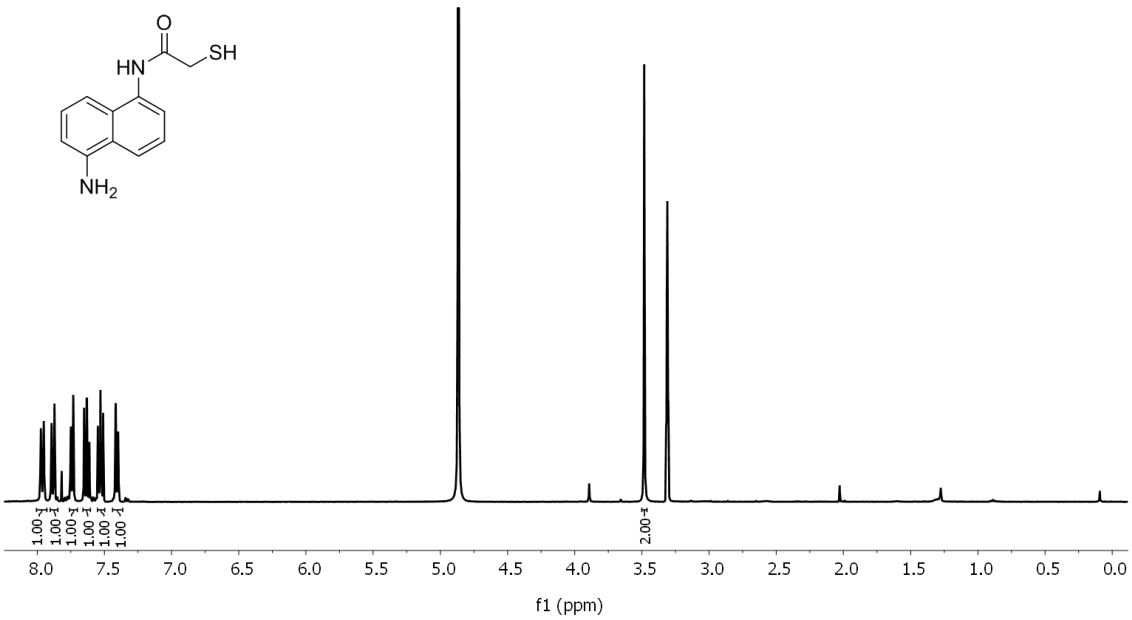


Figure S14. ^{1}H (400 MHz, 298 K in MeOD-d_4) spectrum of **3f**.

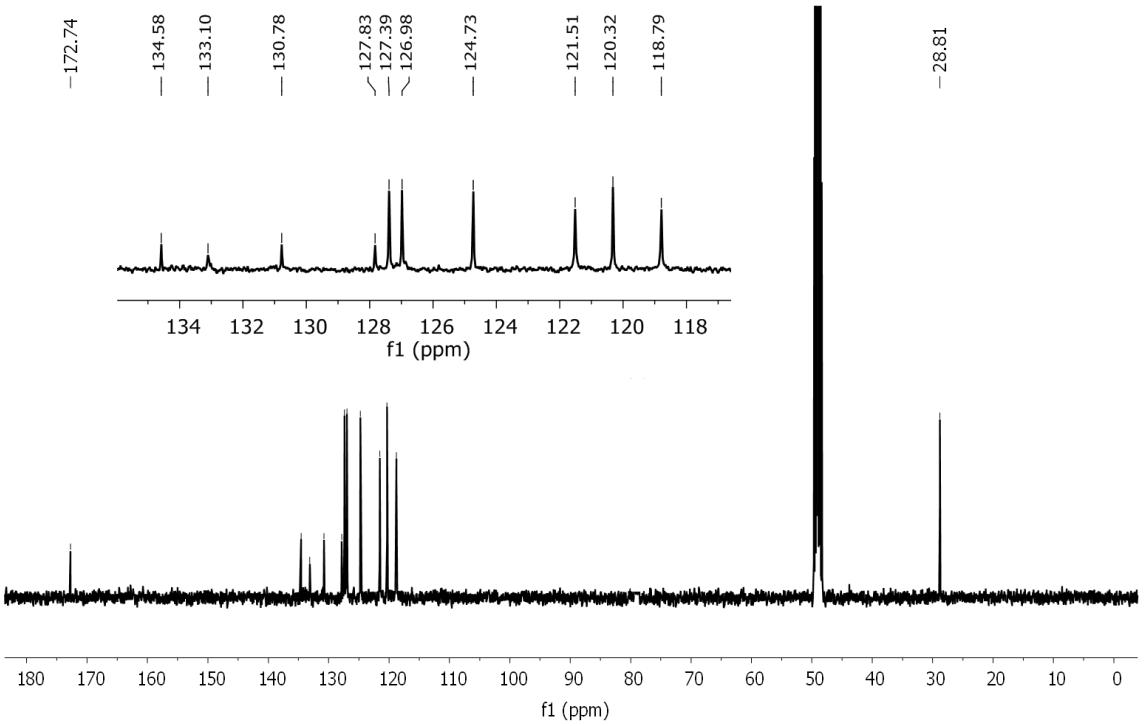


Figure S15. ^{13}C (101 MHz, 298 K in MeOD-d_4) spectrum of **3f**.

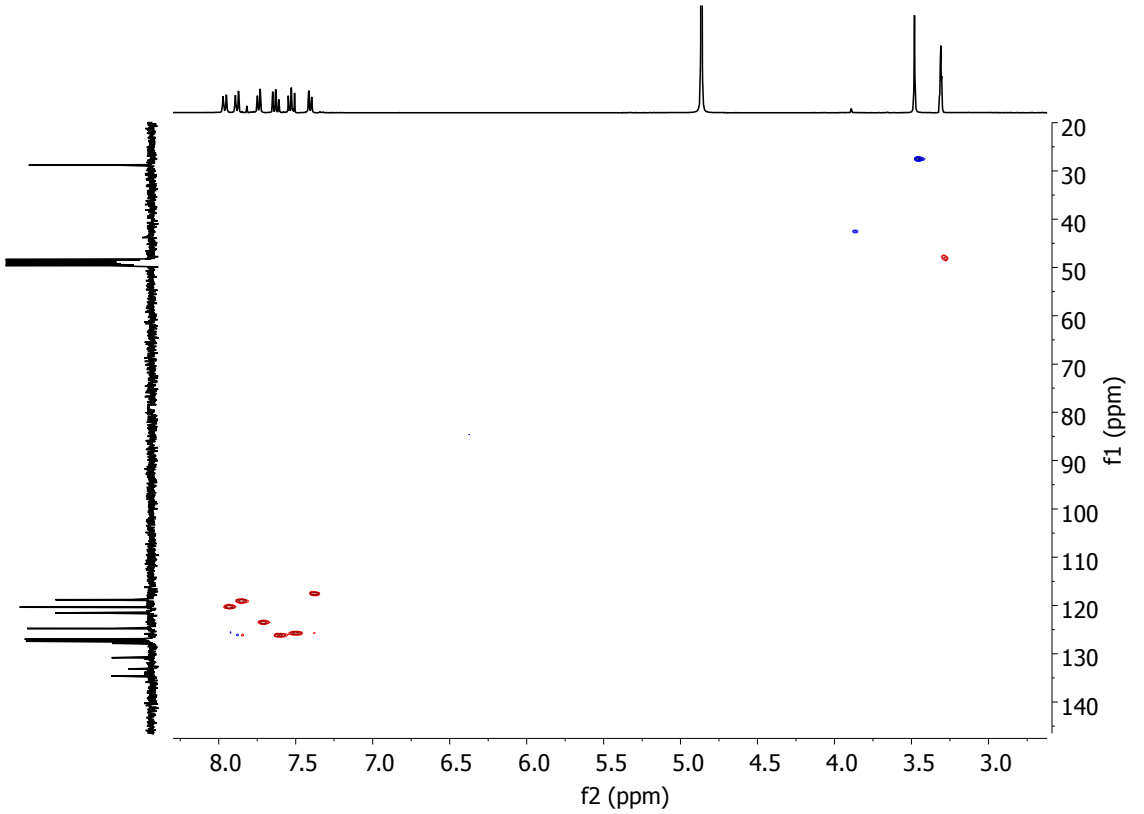


Figure S16. ^{1}H-^{13}C gHSQC (400 MHz, 298 K in MeOD-d_4) spectrum of **3f**.

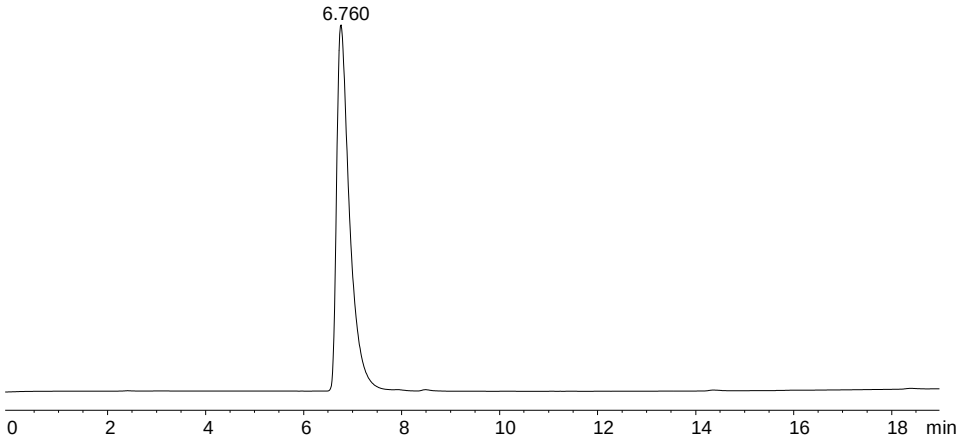


Figure S17. HPLC of **3f** (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

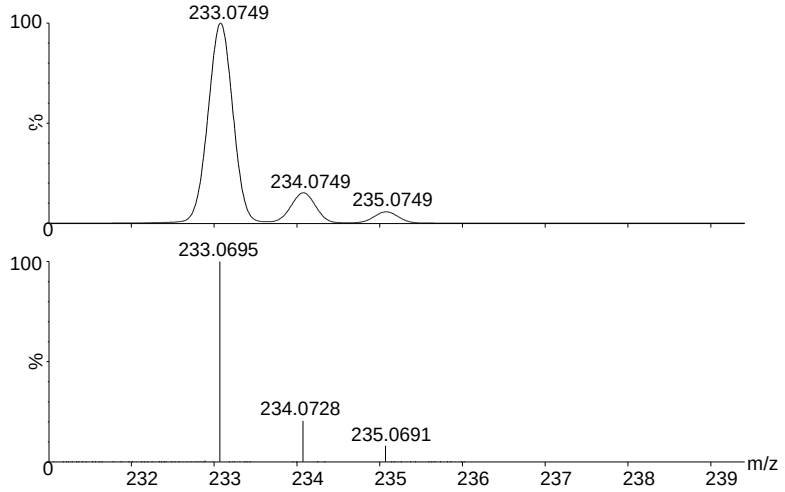


Figure S18. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **3f**.

NMR spectra, HRMS (ESI+) spectrum and HPLC trace of 3g

Building block **3g:**

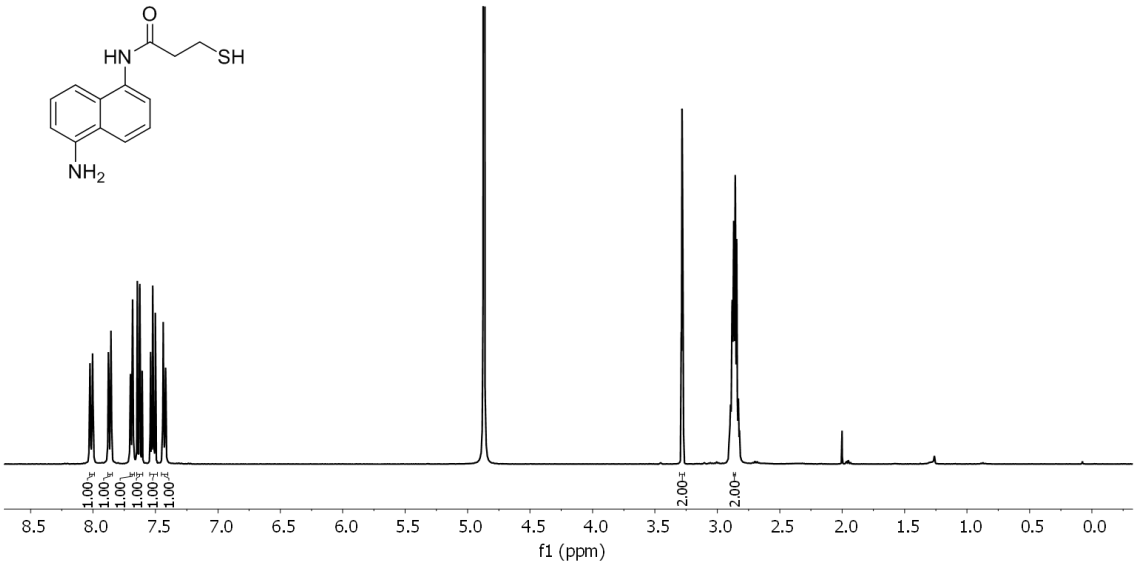


Figure S19. ^{1}H (400 MHz, 298 K in MeOD-d_4) spectrum of **3g**.

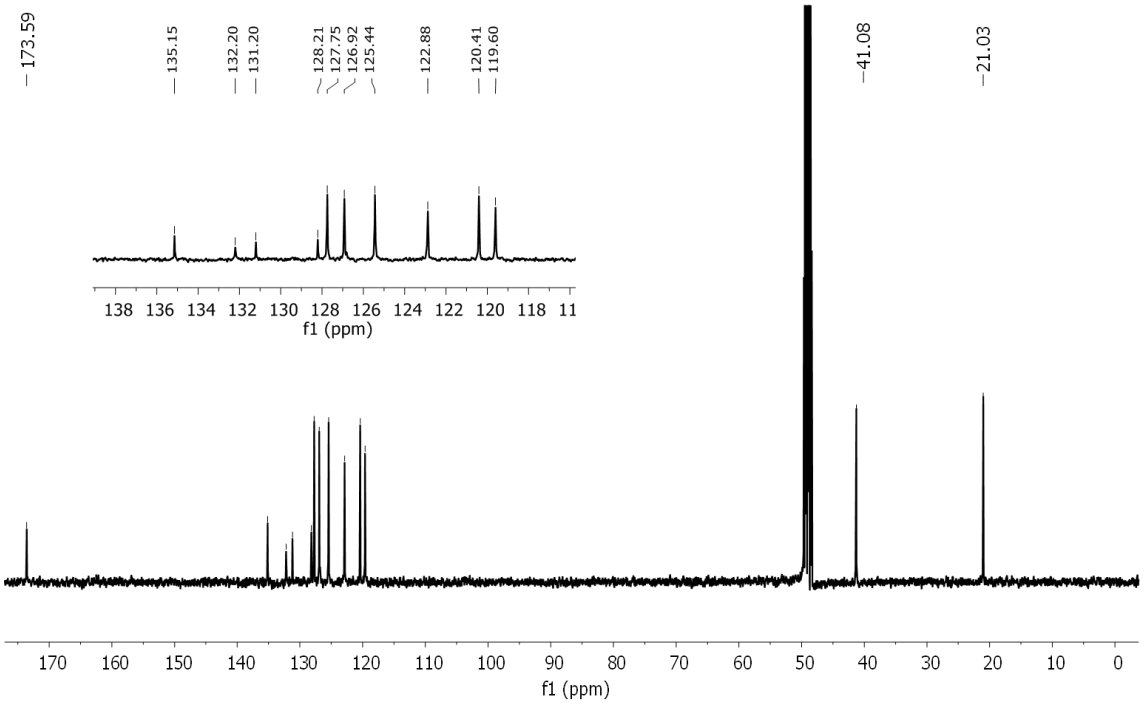


Figure S20. ^{13}C (101 MHz, 298 K in MeOD-d_4) spectrum of **3g**.

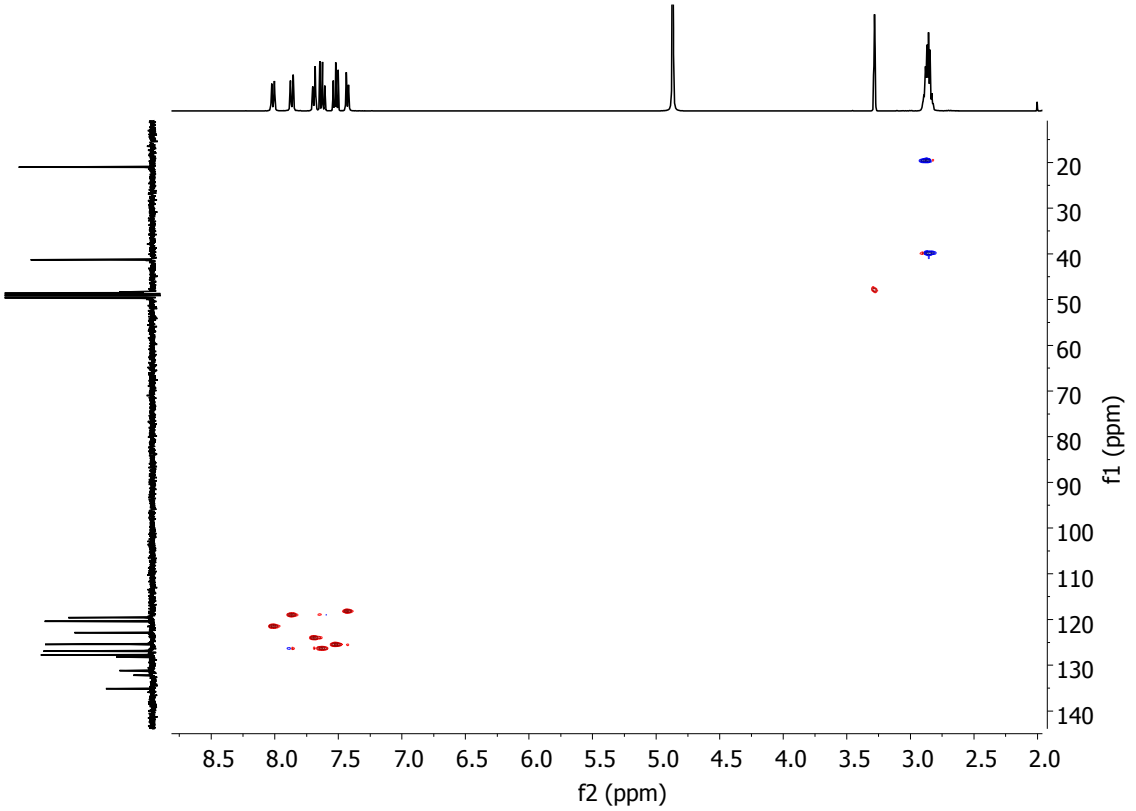


Figure S21. ^{1}H-^{13}C gHSQC (400 MHz, 298 K in MeOD-d_4) spectrum of **3g**.

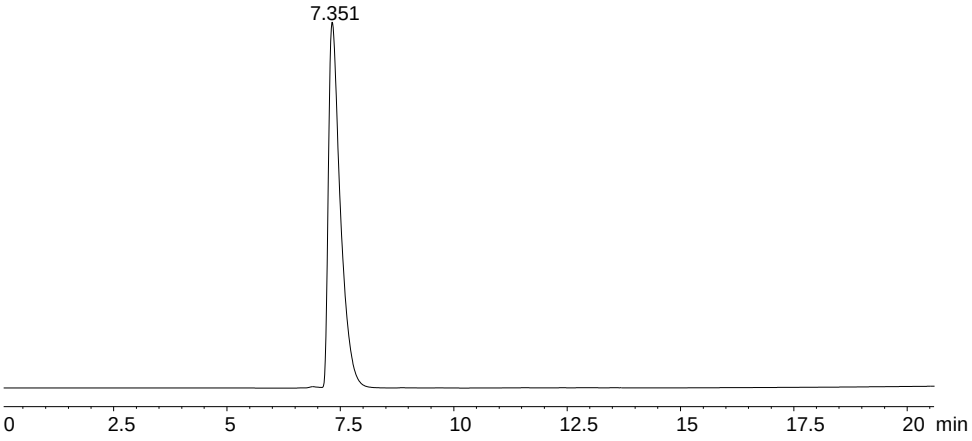


Figure S22. Analytical HPLC trace of **3g** (2 min at 1% CH_3CN in H_2O, then linear gradient from 1% to 100% CH3CN over 20 min).

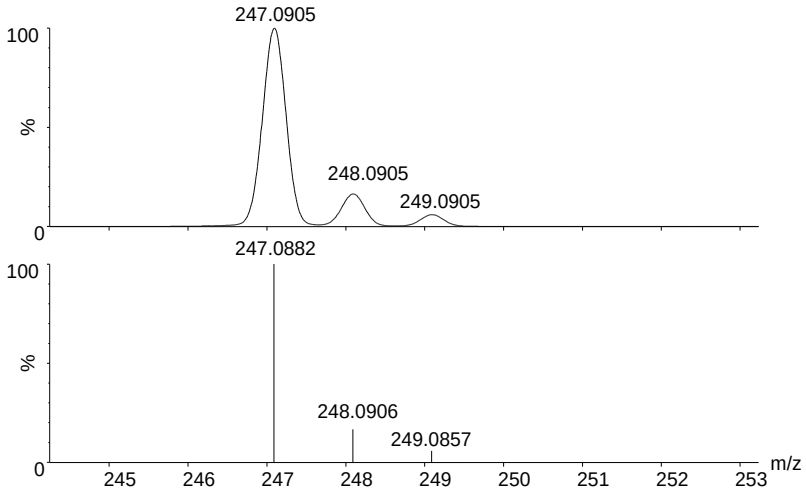


Figure S23. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **3g**.

NMR spectra, HRMS (ESI+) spectrum and HPLC trace of 3h

Building block **3h:**

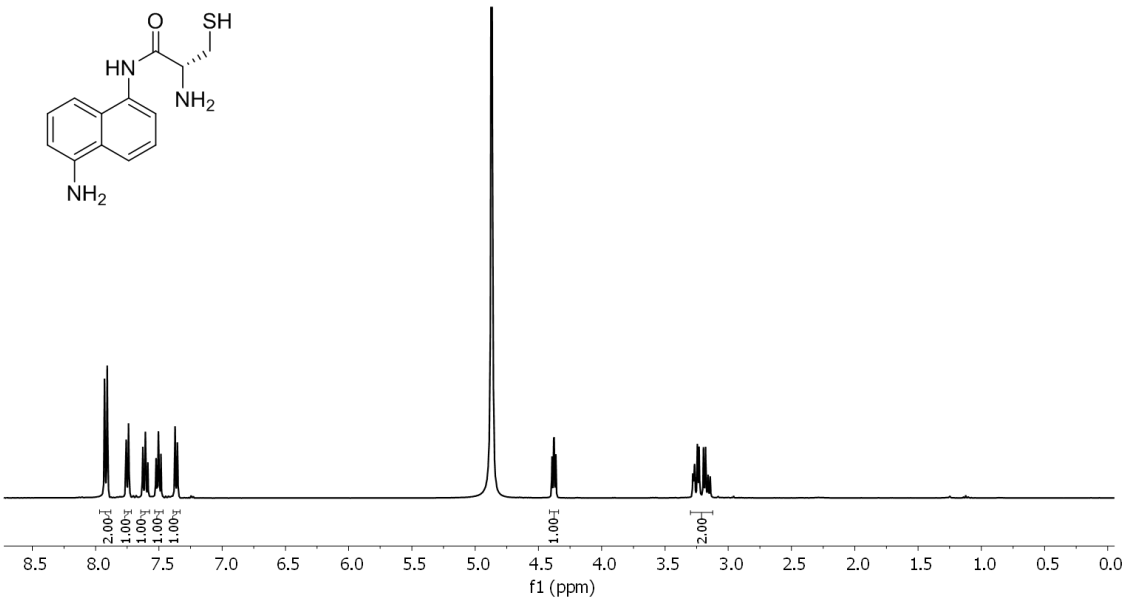


Figure S24. ^{1}H (400 MHz, 298 K in MeOD-d_4) spectrum of **3h**.

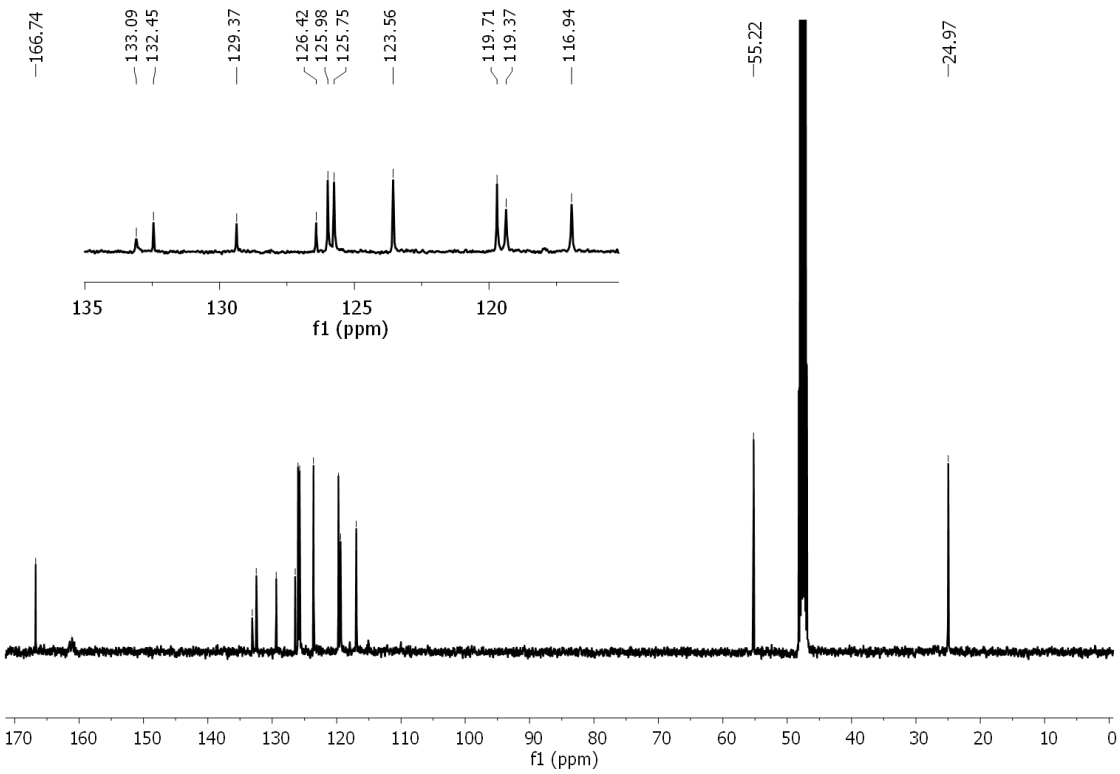


Figure S25. ^{13}C (101 MHz, 298 K in MeOD-d_4) spectrum of **3h**.

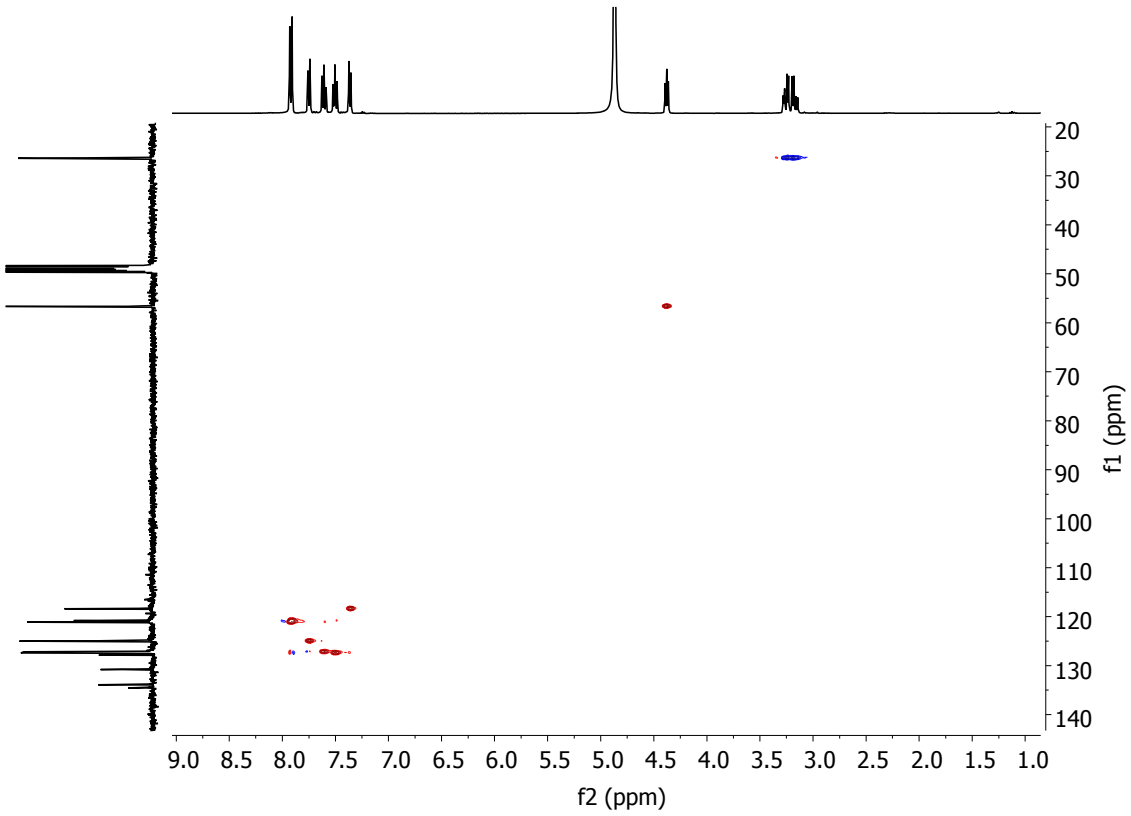


Figure S26. ^{1}H-^{13}C gHSQC (400 MHz, 298 K in MeOD-d_4) spectrum of **3h**.

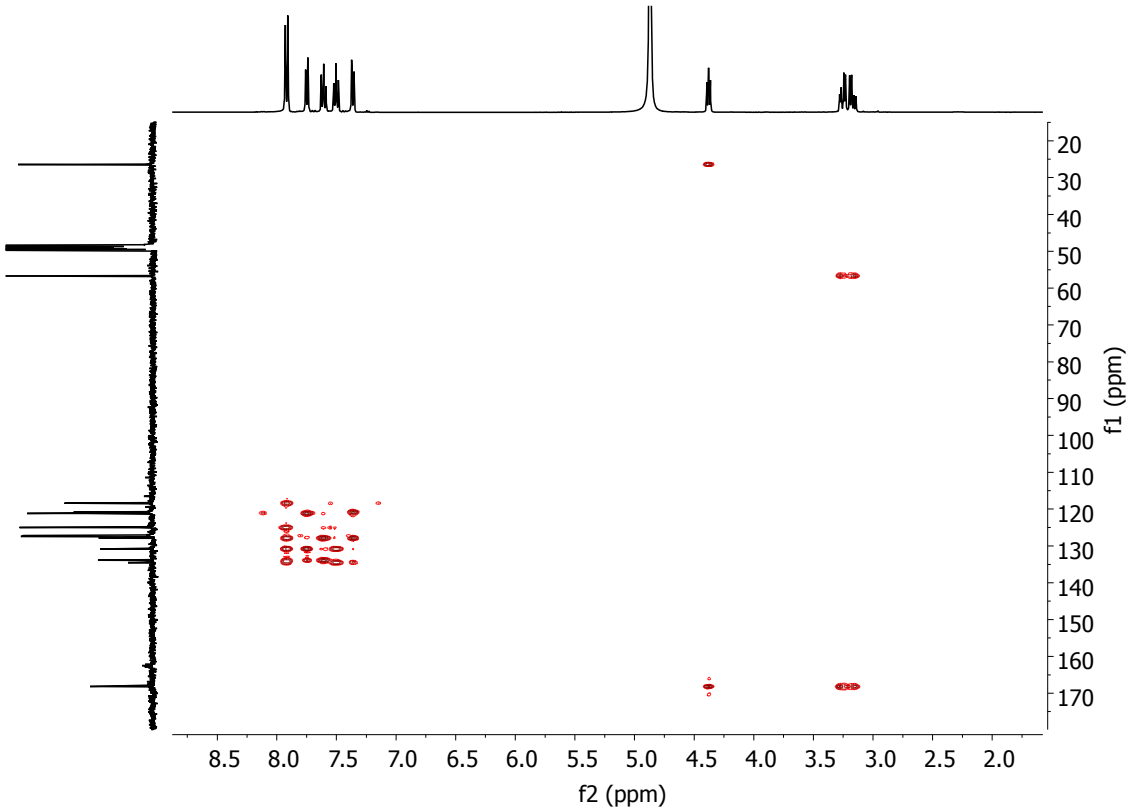


Figure S27. ^{1}H-^{13}C gHMBC (400 MHz, 298 K in MeOD-d_4) spectrum of **3h**.

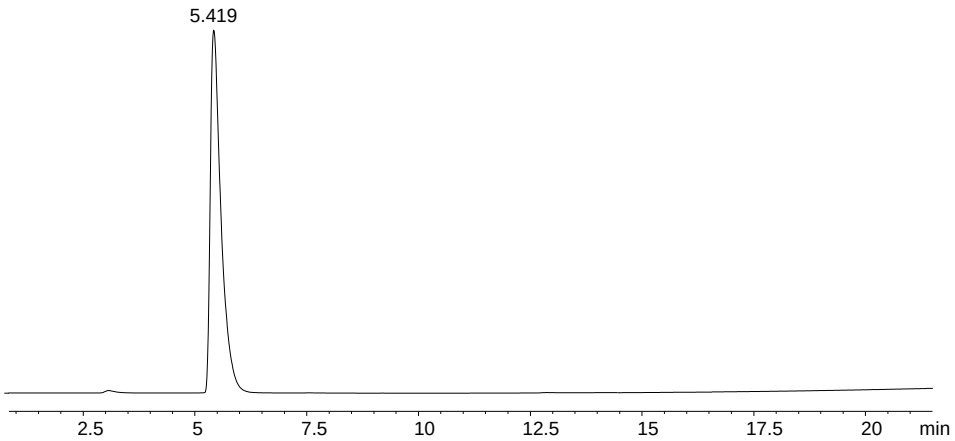


Figure S28. Analytical HPLC trace of **3h** (2 min at 5% CH_3CN in H_2O, then linear gradient from 5% to 100% CH_3CN over 18 min).

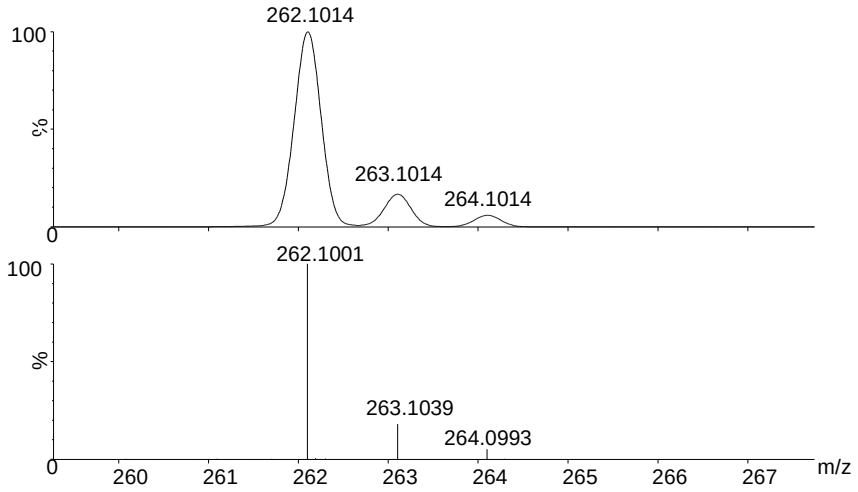


Figure S29. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **3h**.

ABSORBANCE AND EMISSION SPECTRA OF THE MONOPODAL COMPONENTS

Absorbance and emission spectra of the monopodal components

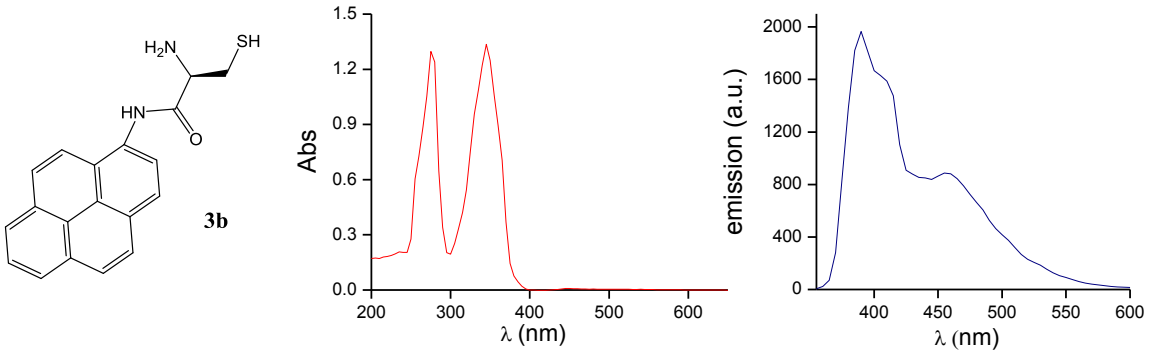


Figure S30. Absorption spectra (red) of 0.65 mM compound **3b** diluted 1:18 in 100% DMSO. Emission spectra (blue) of 0.65 mM compound **3b** diluted 1:41 in 100% DMSO; λexc = 345 nm.

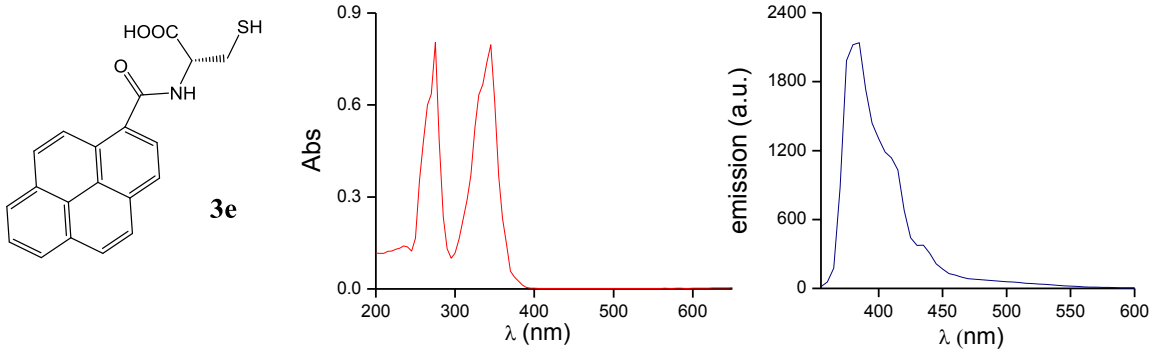


Figure S31. Absorption spectra (red) of 0.65 mM compound **3e** diluted 1:18 in 100% DMSO. Emission spectra (blue) of 16 µM compound **3e** diluted 1:41 in 100% DMSO; λexc = 345 nm.

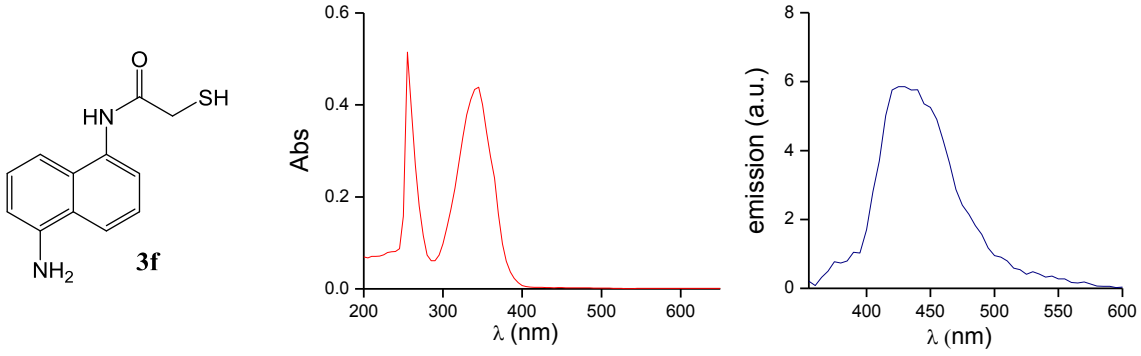


Figure S32. Absorption spectra (red) of 0.65 mM compound **3f** diluted 1:18 in 100% DMSO. Emission spectra (blue) of 0.65 mM compound **3f** diluted 1:41 in 100% DMSO; λexc = 345 nm.

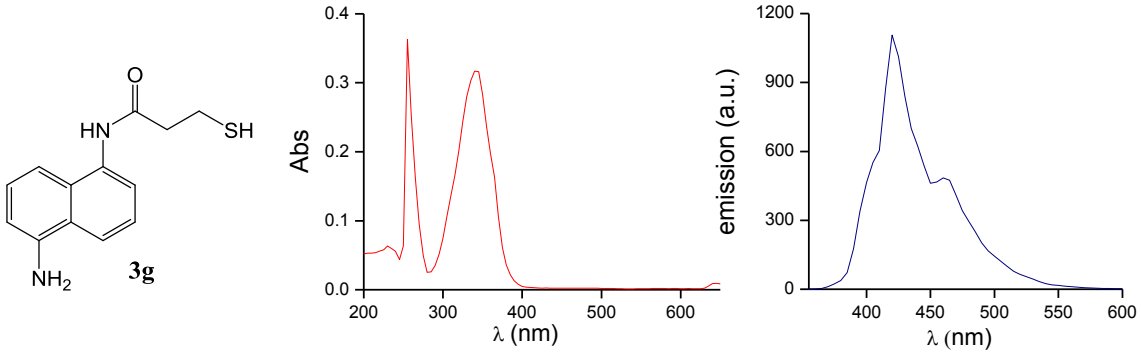


Figure S33. Absorption spectra (red) of 0.65 mM compound **3g** diluted 1:18 in 100% DMSO. Emission spectra (blue) of 0.65 mM compound **3g** diluted 1:41 in 100% DMSO; λexc = 345 nm.

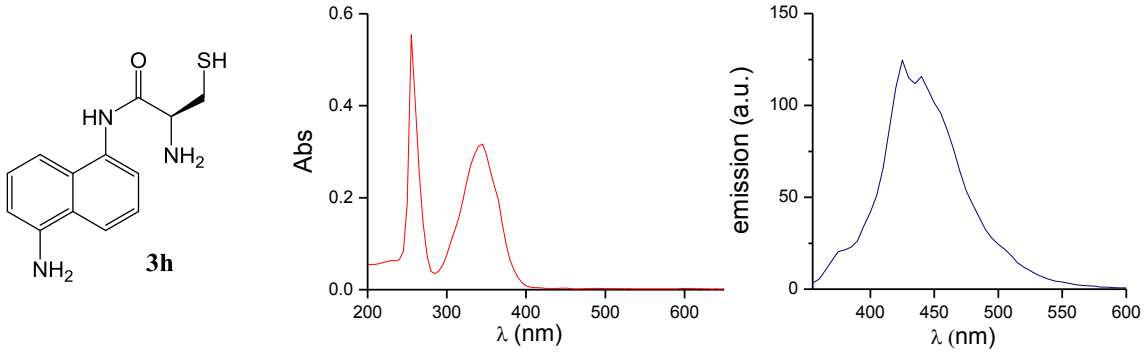


Figure S34. Absorption spectra (red) of 0.65 mM compound **3h** diluted 1:18 in 100% DMSO. Emission spectra (blue) of 0.65 mM compound **3h** diluted 1:41 in 100% DMSO; λexc = 345 nm.

Comparative absorbance and emission spectra of the monopodal components

The absorption and fluorescence emission of the five candidates with fluorescent aromatic groups were measured. After analyzing the fluorescence and the respective solubility of each compound, we selected **3b** as ideal moiety to develop our sensing network.

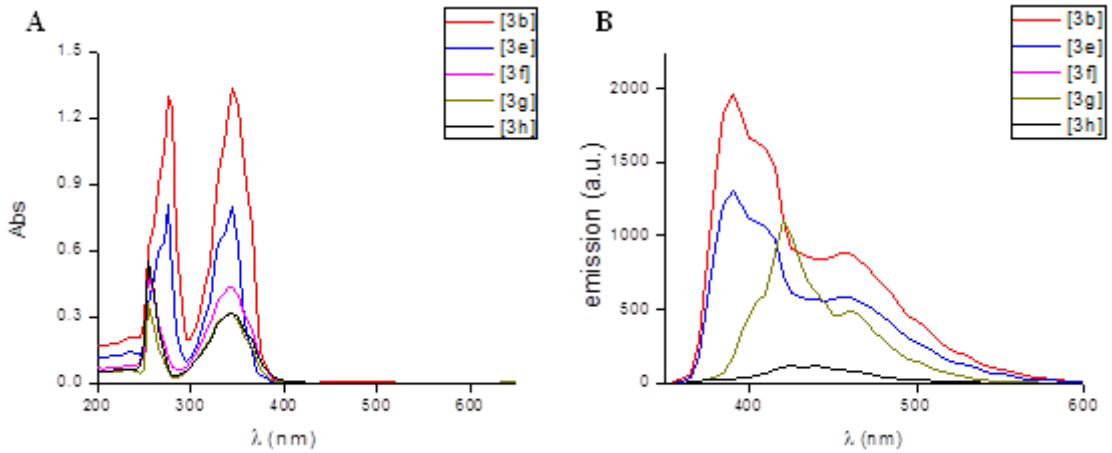


Figure S35. A) Absorption spectrum of the monopodal compounds (0.65 mM) diluted 1:18 in 100% DMSO. B) Emission spectrum of the monopodal compounds (0.65 mM) diluted 1:41 in 100% DMSO; λexc = 345 nm.

DYNAMIC COMBINATORIAL LIBRARIES

General procedure for the quantification of the free thiols

A 4 mM solution of 5,5'-Dithiobis(2-nitrobenzoic acid) (DTNB or Ellman's reagent) was prepared by dissolving 7.93 mg in 5 mL of a 0.1 M phosphate buffer (pH 8.0). Then 200 μL of this freshly prepared solution were placed in a well of a 96-well plate. Finally, 2.5 μL of the sample were also added in the well and, after 2 min of incubation at room temperature, the absorbance of the well was measured at 412 nm. The microplate reader was set to shake the samples for 5 seconds before each measurement.

For all the batches and reaction times, the absorbance of a blank sample was also measured. The blanks were prepared by adding 2.5 μL of DMSO or milli-Q water (depending on the corresponding sample) to 200 μL of the Ellman's reagent solution. The net absorbance was calculated by subtracting the absorbance of the corresponding blank. All the studied oxidations were carried out at room temperature, in capped vials and without any stirring.

Calibration curve

A 40 mM stock solution of cysteine was prepared by dissolving 242.32 mg in 50 mL of milli-Q water. From this, the rest of the stocks (36, 32, 28, 24, 20, 16, 12, 8, 4, 2 and 1 mM) were 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 the regression line was obtained by using the linear least square method (see Figure S39). The data showed good linear behavior within the range of working concentrations demonstrating the reliability of the quantification method.

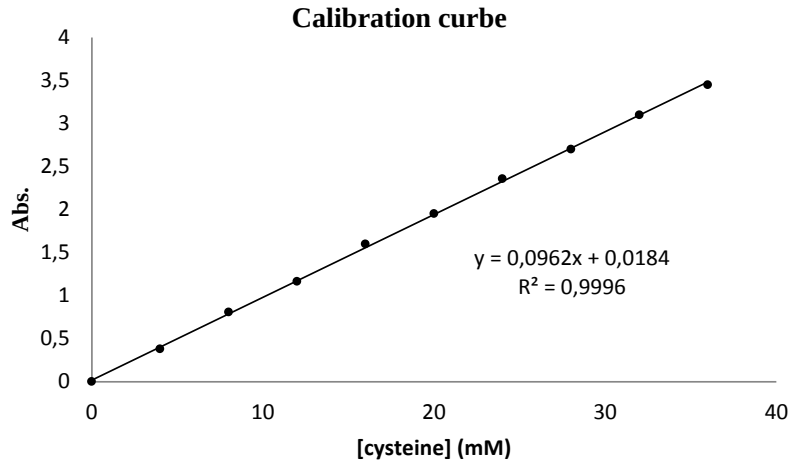


Figure S36: linear least square calibration curve.

General procedure for the preparation and HPLC analysis of the DCLs

General procedure for the preparation of the DCLs

A 66.7 mM BIS-TRIS methane buffer was prepared by dissolving 1.39 g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq).

Individual concentrated stock solutions for the different building blocks (BBs) were prepared in DMSO. Mixture stock solutions containing the necessary BBs for the generation of the libraries (**1**, **2** and **3a-g**) were prepared from these individual solutions. The reaction mixtures were then prepared by dilution of the stock solutions ensuring no differences in concentration between the reaction mixtures of the same batch. The reaction mixtures samples were prepared by adding 60 μL of the corresponding BBs stock mixture in DMSO to 180 μL of a solution of 66.7 mM BIS-TRIS methane buffer. Unless otherwise specified, the DCLs were prepared at final concentrations of 0.1 mM for the di- and tripodal BBs (**1** and **2**) in a 50 mM BIS-TRIS methane buffer (pH 6.5) with 25% DMSO. The concentrations for the monopodal BBs **3a-g** are specified for each of the experiments.

General procedure for the analysis of the DCLs

The mixtures were analysed by means of HPLC or UPLC-MS at different reaction times. Complete oxidation is achieved after 24h h for reactions containing 25% DMSO. The HPLC samples were prepared by adding 45 μL of the corresponding reaction mixture to 75 μL of a solution of 89% H_2O, 10% CH_3CN and 1% TFA. Eluent used: 2 min at 10% CH_3CN in H_2O, then linear gradient from 10% to 50% CH_3CN over 58 min.

Mixture of 1, 2 and 3b at different concentrations

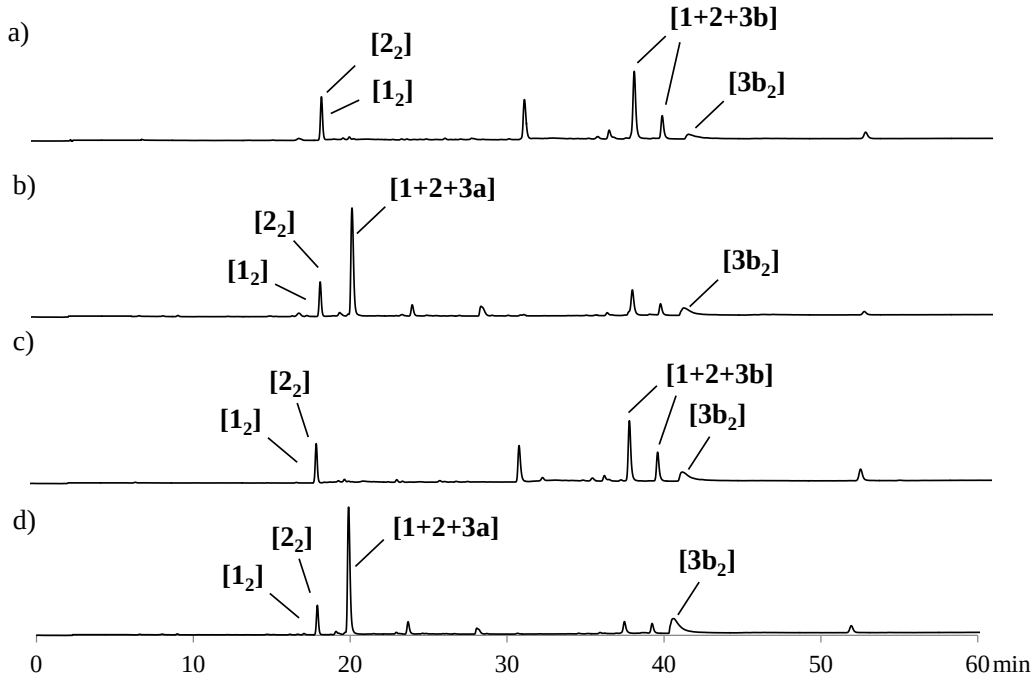


Figure S37. HPLC traces of the mixture of **1+2+3b** at pH 6.5 with 25% of DMSO at different concentrations: (a) 0.5+0.5+0.25 mM, (b) 0.5+0.5+0.25 mM with 2.5 mM L-Cys, (c) 0.5+0.5+0.5 mM, (d) 0.5+0.5+0.5 mM with 2.5 mM L-Cys.

Mixture of 1, 2 and 3g/3e with and without L-Cys

The monothiols **3e** and **3g** were less selective than **3b** in order to liberate a fluorescent response**,** thus **3b** was the candidate chosen for building the Cysteine sensor. After analyzing the fluorescence and the respective solubility of each compound, the HPLC traces confirmed that **3b** was the best option.

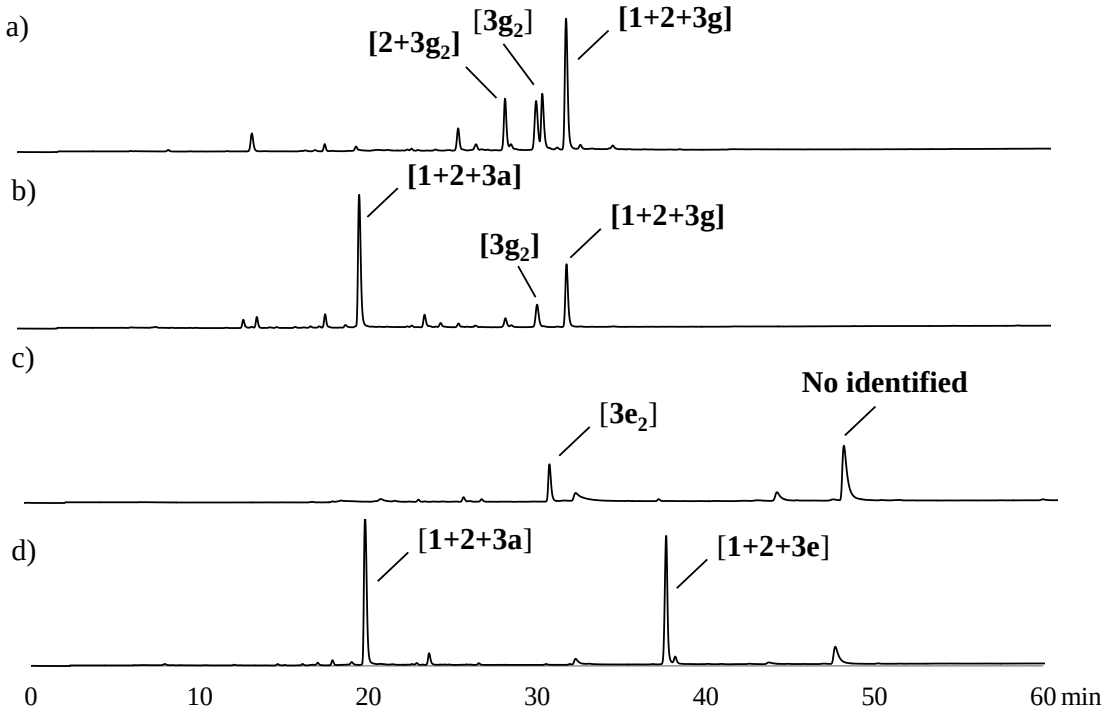


Figure S38. HPLC traces of the mixture of **1+2+3g/3e** (0.5 mM each) at pH 6.5 with 25% of DMSO: (a) **1+2+3g**, (b) **1+2+3g** with 2.5 mM L-Cys, (c) **1+2+3e**, (d) **1+2+3e** with 2.5 mM L-Cys.

Mixture of 1, 2 and 3b with different biothiols

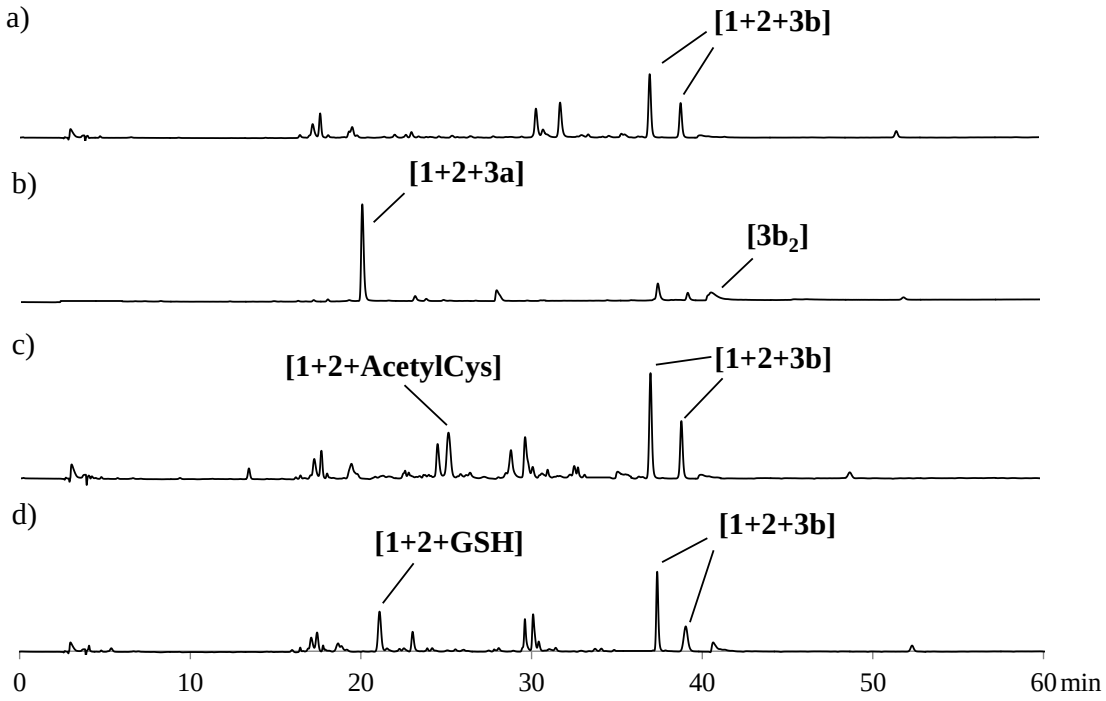


Figure S39. HPLC traces of the mixture of **1+2+3b** (0.5+0.5+0.25 mM each), at pH 6.5 with 25% of DMSO with different biothiols: (a) without biothiol, (b) 0.25 mM L-Cys, (c) 0.25 mM AcetylCysteine, (d) 0.25 mM GSH.

MASS SPECTROMETRY OF THE DCL's

General procedure for the analysis of the DCLs by HRMS

The HRMS (UPLC-ESI-TOF) samples were prepared by adding 20 μL of the corresponding reaction mixture to 40 μL of a solution of 89% H_2O, 10% MeCN and 1% TFA. Eluent used: 2.5 min at 10% CH_3CN (+ 20mM HCOOH) in H_2O (+ 20mM HCOOH), then linear gradient from 10% to 55% CH_3CN over 37.5 min.

Mixture of 3b (0.5 mM, pH.6.5, 25% DMSO)

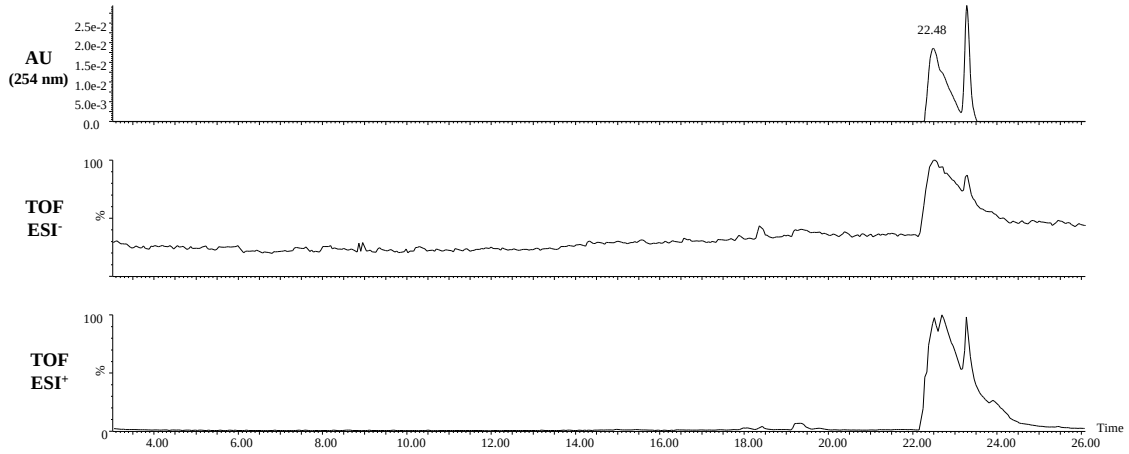


Figure S40. UPLC_ESI_TOF for **3b** (0.5 mM, pH.6.5, 25% DMSO)

Identification of the products:
[3b$_2$]
Retention time: 22.48 min.
Chemical Formula: $C_{38}H_{30}N_4O_2S_2$
Exact Mass: 638.1810

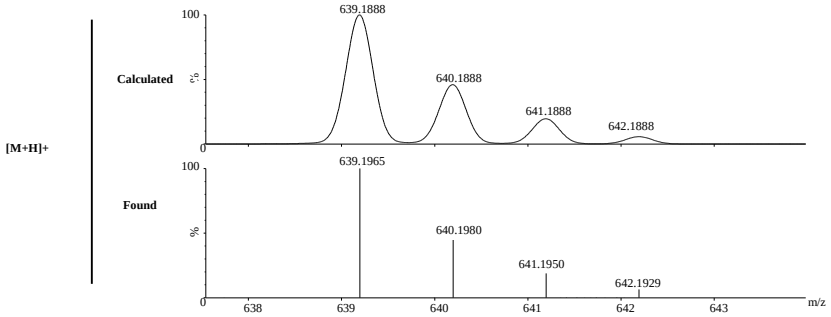


Figure S41. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**3b$_2$**]

Mixture of 1+2+3b (0.5 mM each, pH.6.5, 25% DMSO)

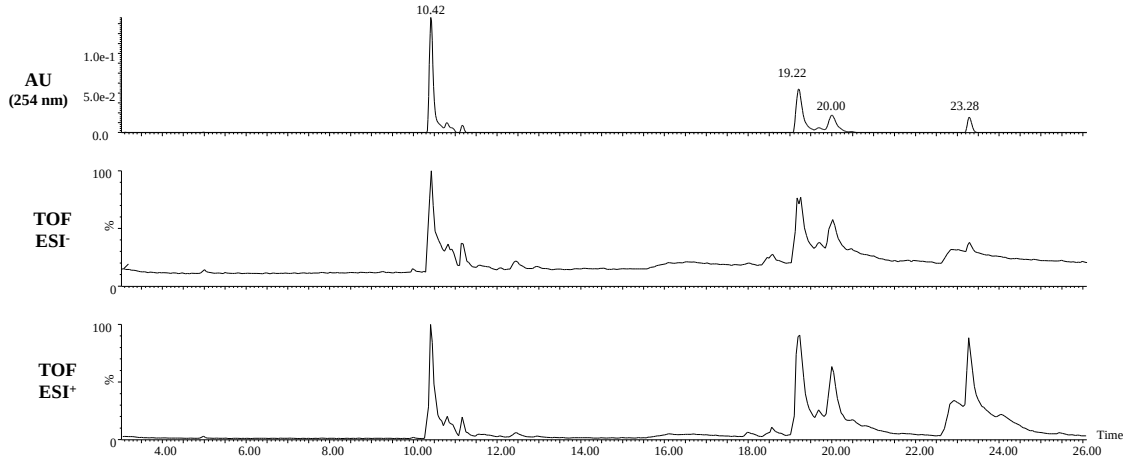


Figure S42. UPLC_ESI_TOF for **1+2+3b** (0.5 mM each, pH.6.5, 25% DMSO)
Identification of the products:

[2_2]
Retention time: 10.42 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

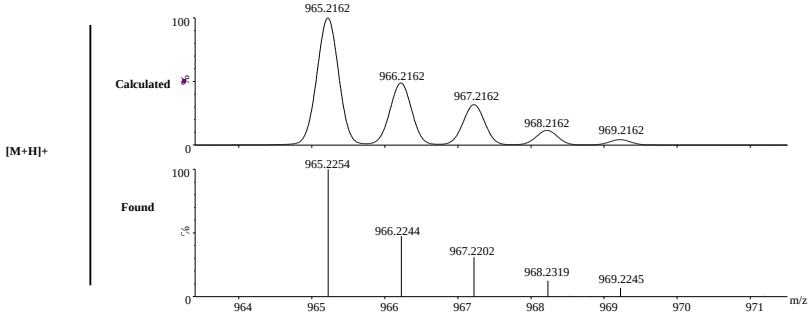


Figure S43. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[2_2]**.

[1-2-3b].
Retention time: 19.22 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

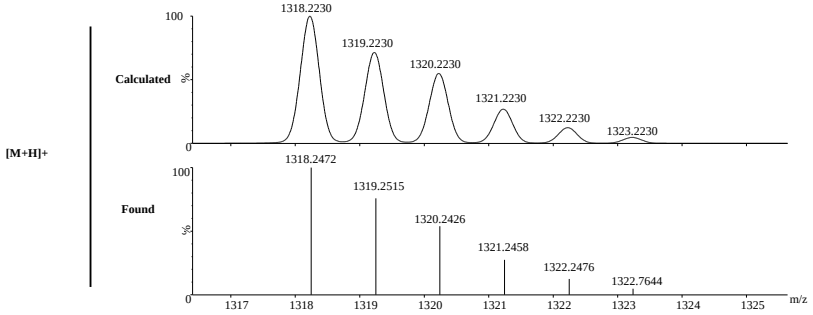


Figure S44. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1-2-3b**].

[1-2-3b].
Retention time: 20.00 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

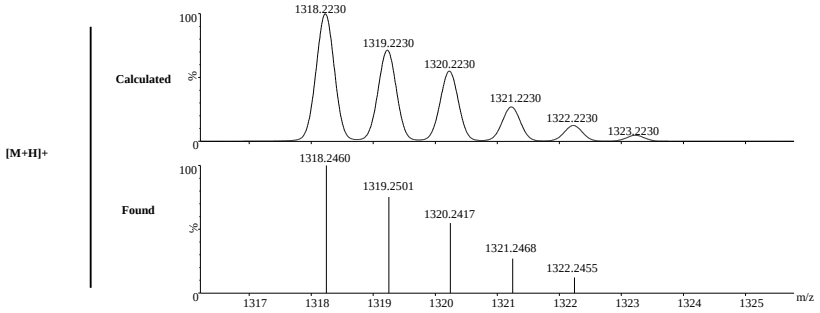


Figure S45. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1-2-3b**].

[$3b_2$] *(previously identified)*
Retention time: 23.28 min.
Chemical Formula: $C_{38}H_{30}N_4O_2S_2$
Exact Mass: 638.1810

Mixture of 1+2+3a+3b (0.5 mM each, pH.6.5, 25% DMSO)

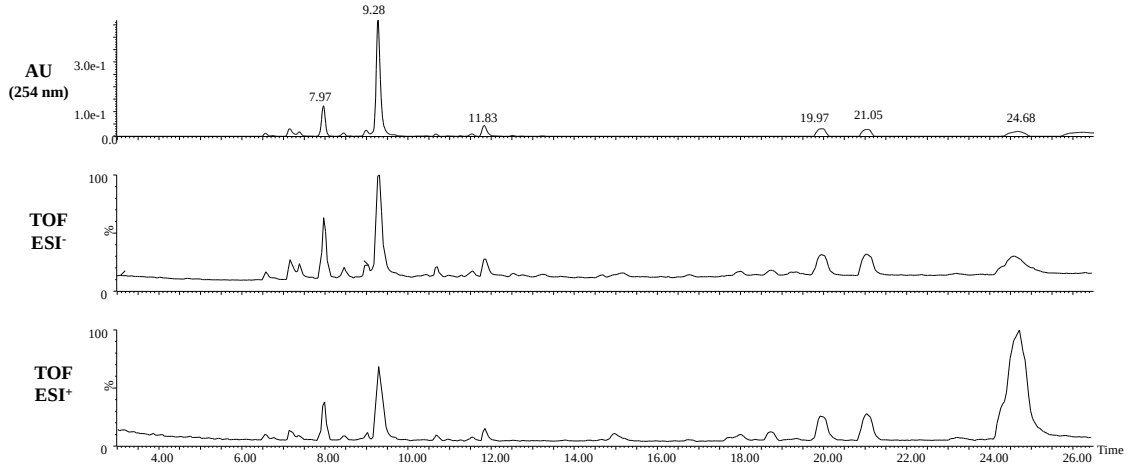


Figure S46. UPLC_ESI_TOF for **1+2+3a +3b** (0.5 mM each, pH.6.5, 25% DMSO)
Identification of the products:

Identification of the products:

[2_2] *(previously identified)*
Retention time: 7.97 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[1-2-3a]
Retention time: 9.28 min.
Chemical Formula: $C_{39}H_{46}N_{10}O_{17}S_6$
Exact Mass: 1118.1367

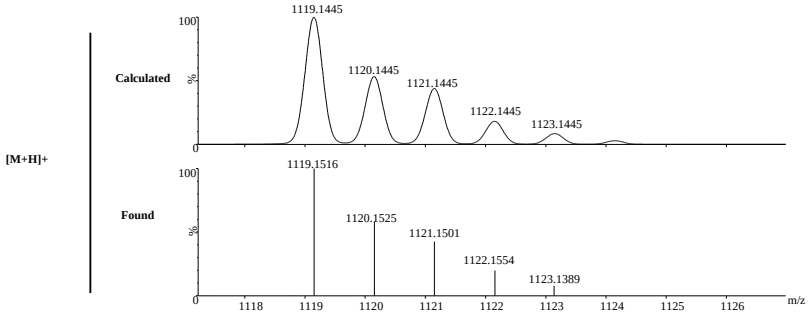


Figure S47. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1-2-3a**].

[1-2$_2$-3a]
Retention time: 11.83 min.
Chemical Formula: $C_{57}H_{68}N_{16}O_{23}S_8$
Exact Mass: 1600.2409

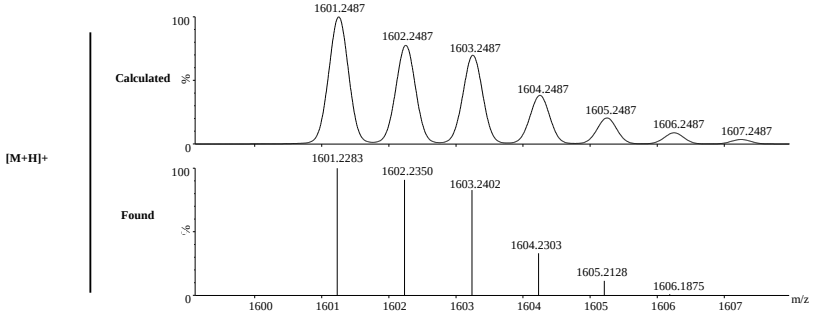


Figure S48. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1-2$_2$-3a**].

[1-2-3b]
Retention time: 19.97 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

[1-2-3b]
Retention time: 21.05 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

[3b$_2$] *(previously identified)*
Retention time: 24.68 min.
Chemical Formula: $C_{38}H_{30}N_4O_2S_2$
Exact Mass: 638.1810

Mixture of 3e (0.5 mM, pH.6.5, 25% DMSO)

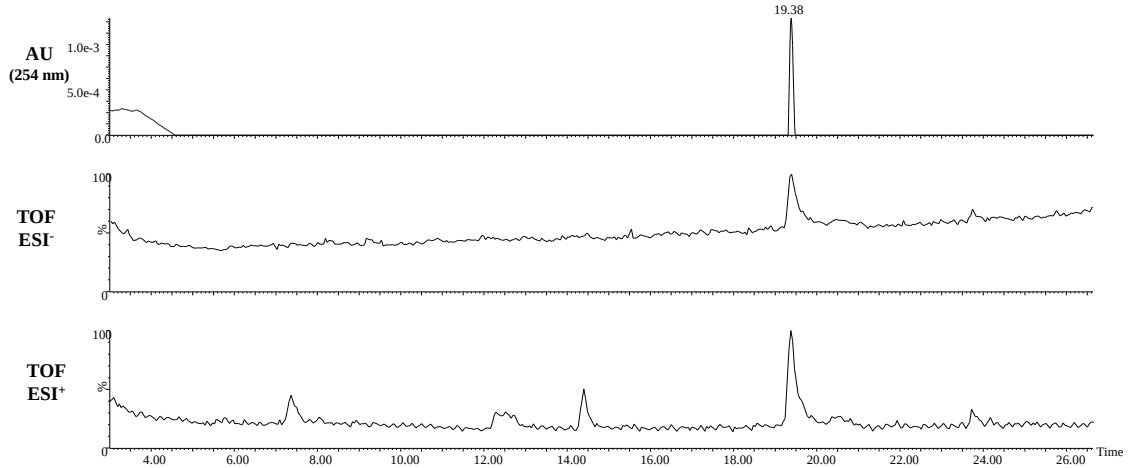


Figure S49. UPLC_ESI_TOF for **3e** (0.5 mM, pH.6.5, 25% DMSO)
Identification of the products:

Identification of the products:

[3e$_2$]
Retention time: 24.68 min.
Chemical Formula: $C_{40}H_{28}N_2O_6S_2$
Exact Mass: 696.1389

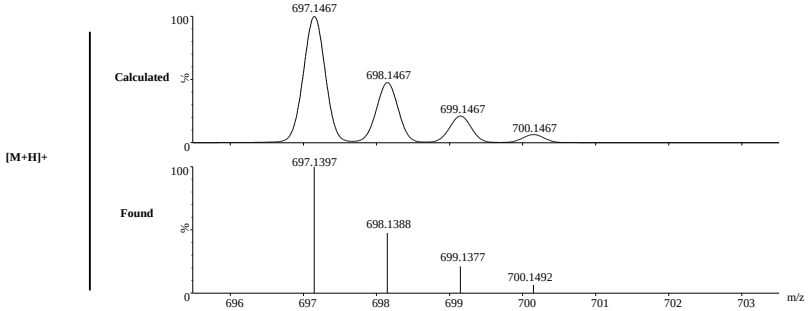


Figure S50. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**3e$_2$**].

Mixture of 1+2+3e (0.5 mM each, pH.6.5, 25% DMSO)

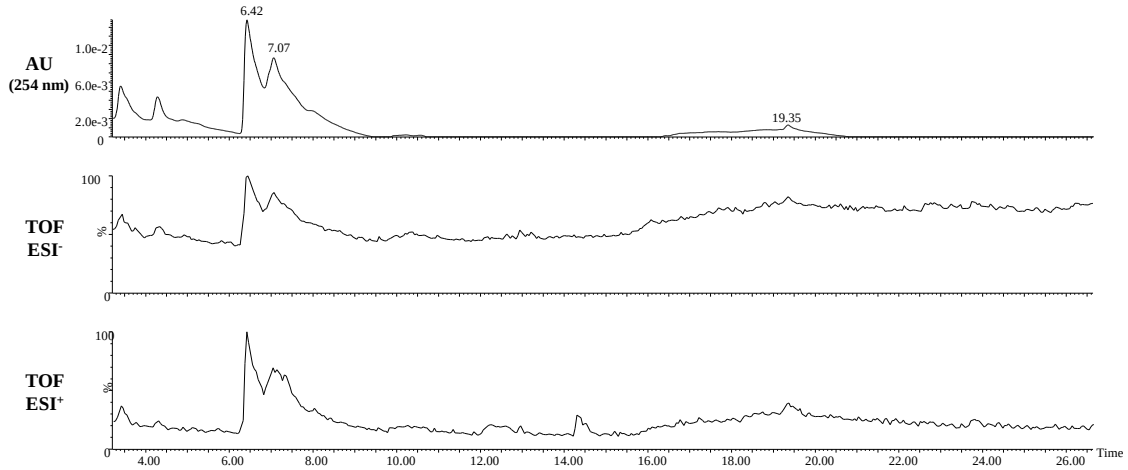


Figure S51. UPLC_ESI_TOF for **1**+**2**+**3e** (0.5 mM each, pH.6.5, 25% DMSO).
Identification of the products:

Identification of the products:

[1a-2a-SH]
Retention time: 6.42 min.
Chemical Formula: $C_{36}H_{41}N_9O_{15}S_5$
Exact Mass: 999.1326

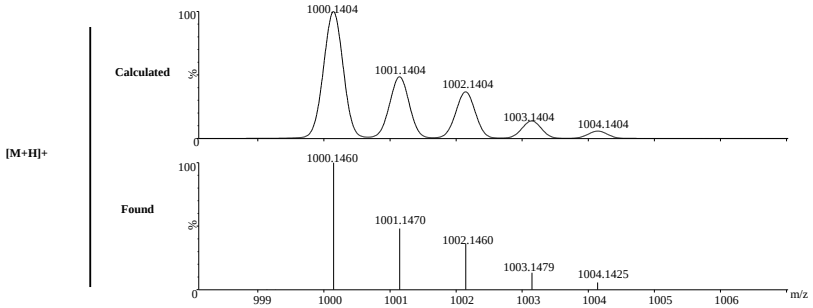


Figure S52. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1a-2a-SH**].

[1a-2a-SH] *(previously identified)*
Retention time: 7.07 min.
Chemical Formula: $C_{36}H_{41}N_9O_{15}S_5$
Exact Mass: 999.1326

[1a-2a-3e]
Retention time: 19.35 min.
Chemical Formula: $C_{56}H_{54}N_{10}O_{18}S_6$
Exact Mass: 1346.1942

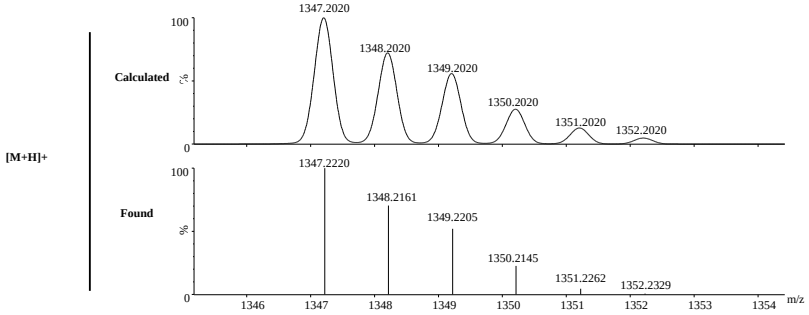


Figure S53. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**1a-2a-3e**].

Mixture of 3g (0.5 mM, pH.6.5, 25% DMSO)

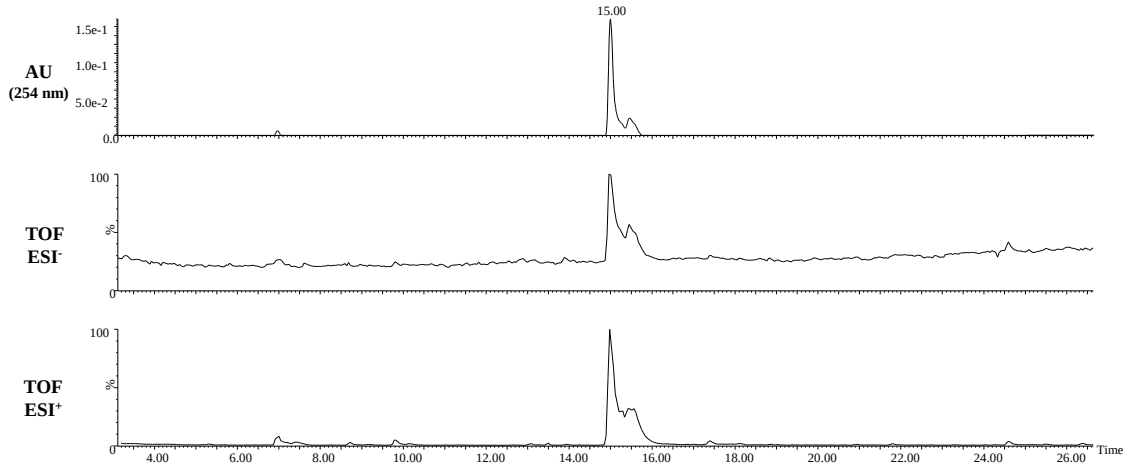


Figure S54. UPLC_ESI_TOF for **3g** (0.5 mM, pH.6.5, 25% DMSO).

Identification of the products:

[3g$_2$]
Retention time: 15.00 min.
Chemical Formula: $C_{26}H_{26}N_4O_2S_2$
Exact Mass: 490.1497

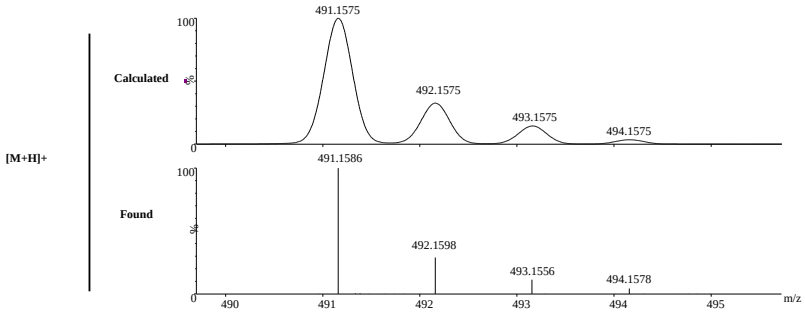


Figure S55. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**3g$_2$**].

Mixture of 1+2+3g (0.5 mM each, pH.6.5, 25% DMSO)

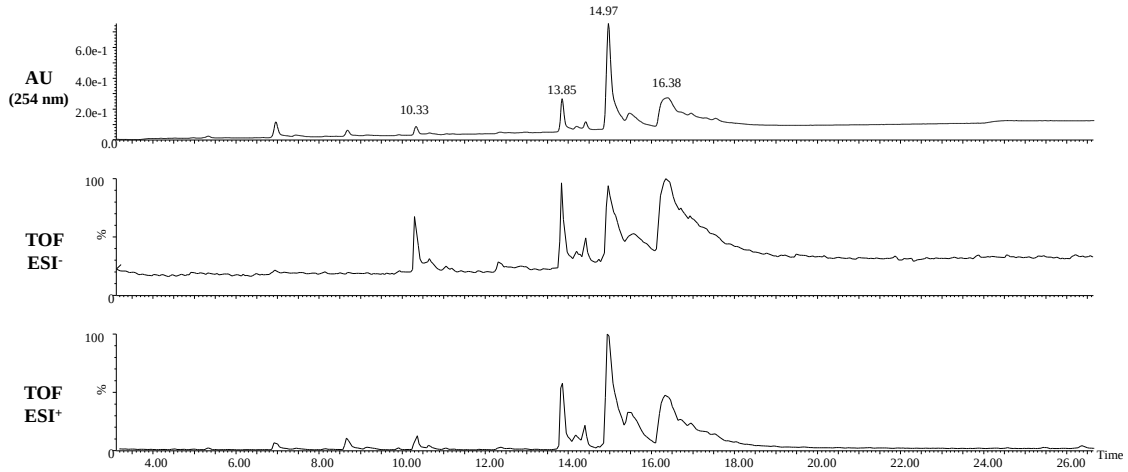


Figure S56. UPLC_ESI_TOF for **1**+**2**+**3g** (0.5 mM each, pH.6.5, 25% DMSO)

Identification of the products:

[2$_2$] *(previously identified)*
Retention time: 10.33 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[2-3g$_2$]
Retention time: 13.85 min.
Chemical Formula: $C_{44}H_{48}N_{10}O_8S_4$
Exact Mass: 972.2539

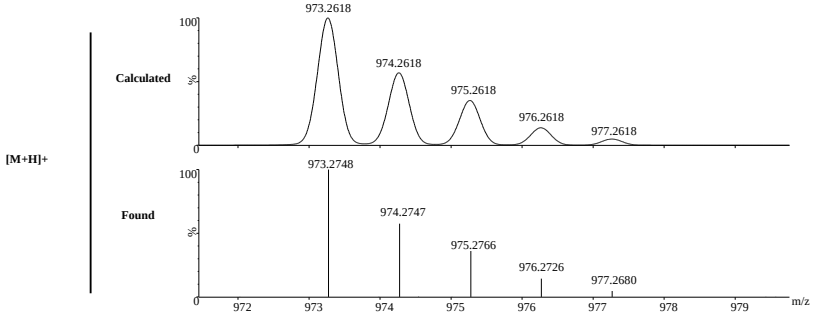


Figure S57. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of [**2-3g$_2$**].

[3g$_2$] *(previously identified)*
Retention time: 14.97 min.
Chemical Formula: $C_{26}H_{26}N_4O_2S_2$
Exact Mass: 490.1497

[1-2-3g]
Retention time: 16.38 min.
Chemical Formula: $C_{49}H_{53}N_{11}O_{16}S_6$
Exact Mass: 1244.3840

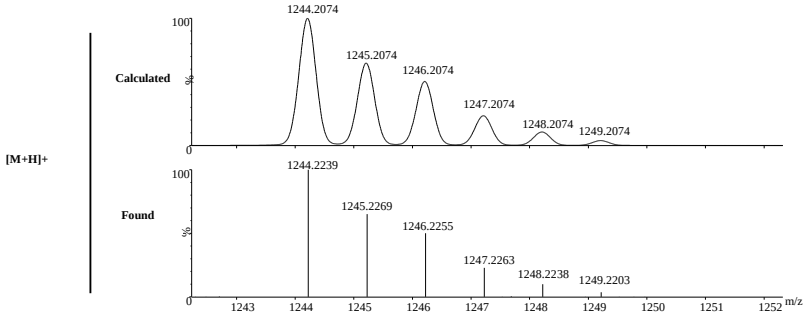


Figure S58. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2-3g]**.

Mixture of 1+2+3b+AcetylCys (0.5+0.5+0.25+0.25 mM each, pH.6.5, 25% DMSO)

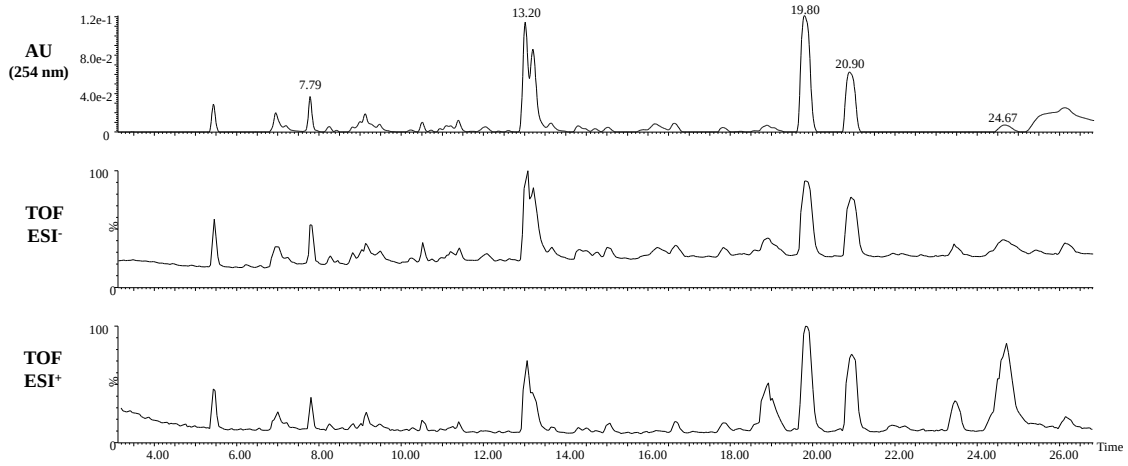


Figure S59. UPLC_ESI_TOF for **1+2+3b+AcetylCys** (0.5+0.5+0.25+0.25 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2_2] *(previously identified)*
Retention time: 7.79 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[1-2-AcetylCys]
Retention time: 13.20 min.
Chemical Formula: $C_{41}H_{48}N_{10}O_{18}S_6$
Exact Mass: 1160.1472

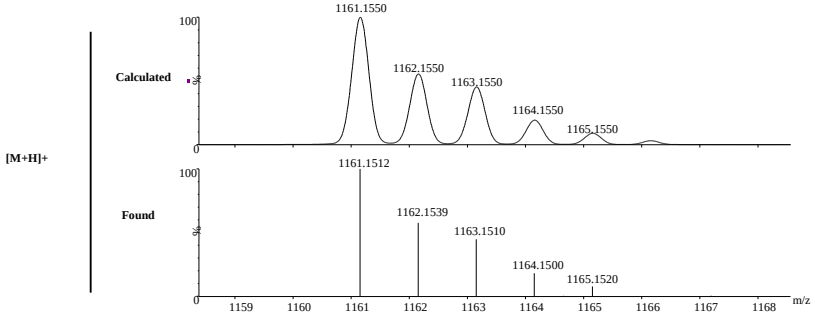


Figure S60. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2_2-AcetylCys]**.

[1-2-3b] *(previously identified)*
Retention time: 19.80 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

[1-2-3b] *(previously identified)*
Retention time: 20.90 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

[3b$_2$] *(previously identified)*
Retention time: 24.67 min.
Chemical Formula: $C_{38}H_{30}N_4O_2S_2$
Exact Mass: 638.1810

Mixture of 1+2+3b+GSH (0.5+0.5+0.25+0.25 mM each, pH.6.5, 25% DMSO)

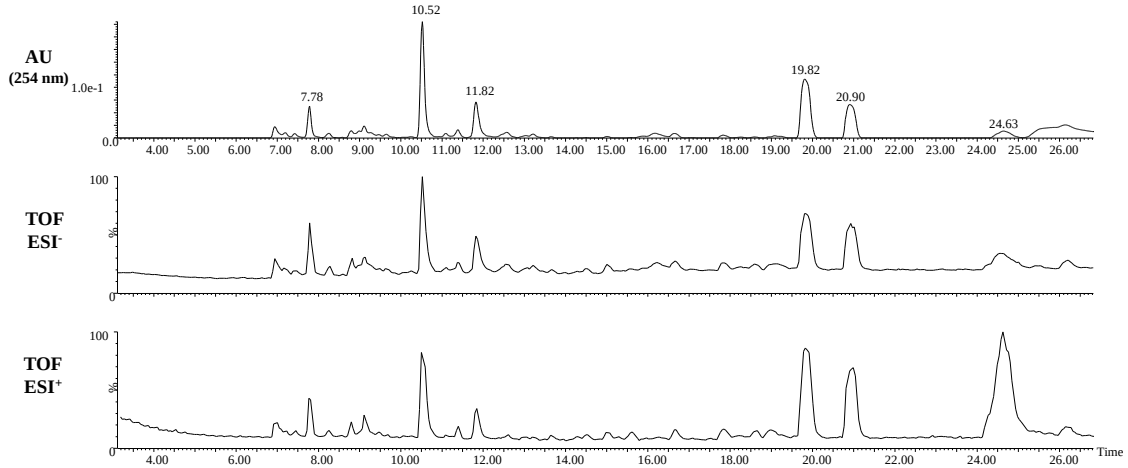


Figure S61. UPLC_ESI_TOF for **1**+**2**+**3b**+**GSH** (0.5+0.5+0.25+0.25 mM each, pH.6.5, 25% DMSO).

Identification of the products:

[2$_2$] *(previously identified)*
Retention time: 7.78 min.
Chemical Formula: $C_{36}H_{44}N_{12}O_{12}S_4$
Exact Mass: 964.2084

[1-2-GSH]
Retention time: 10.52 min.
Chemical Formula: $C_{46}H_{56}N_{12}O_{21}S_6$
Exact Mass: 1304.2007

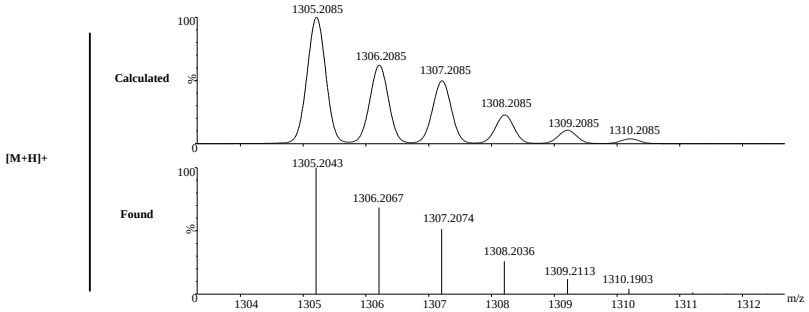


Figure S62. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2-GSH]**.

[1-2_2-GSH]
Retention time: 11.82 min.
Chemical Formula: $C_{64}H_{78}N_{18}O_{27}S_8$
Exact Mass: 1786.3049

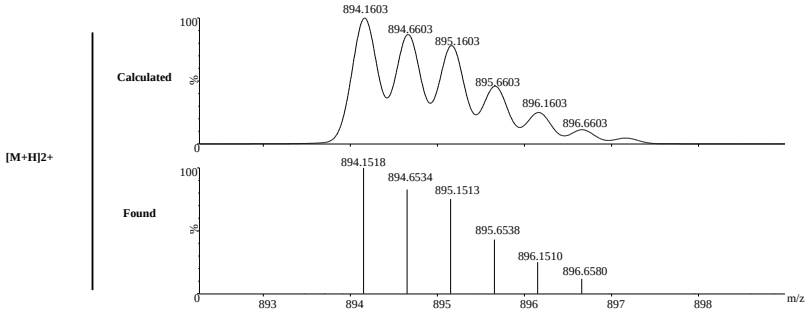


Figure S63. Experimental (lower trace) and simulated (upper trace) ESI-TOF mass spectra for $[M+H]^+$ of **[1-2_2-GSH]**.

[1-2-3b] *(previously identified)*
Retention time: 19.82 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

[1-2-3b] *(previously identified)*
Retention time: 20.90 min.
Chemical Formula: $C_{55}H_{55}N_{11}O_{16}S_6$
Exact Mass: 1317.2152

[3b_2] *(previously identified)*
Retention time: 24.63 min.
Chemical Formula: $C_{38}H_{30}N_4O_2S_2$
Exact Mass: 638.1810

EMISSION SPECTRA OF THE SENSING SYSTEM

General procedure for the preparation and fluorescence analysis of the DCLs

General procedure for the preparation of the DCLs

BIS-TRIS buffer was used to regulate the pH. A 66.7 mM BIS-TRIS methane buffer was prepared by dissolving 1.39 g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq). Individual stocks of the building blocks (BBs) **1**, **2** and **3b** in DMSO were prepared. Then stock mixture solutions of the BBs were prepared from the individual stocks ensuring no differences in concentration between the reaction mixtures of the same batch. Unless otherwise specified, the DCLs were prepared in the reaction mixtures at final concentrations of 0.1 mM for the tri- and bipodal BBs (**1** and **2**) and 0.05 mM of **3b** in a final 50 mM BIS-TRIS methane aqueous buffer (pH 6.5) with 25% DMSO. The samples were prepared by adding 15 µL of the stock mixture in DMSO to 45 µL of the 66.7 mM buffered solution, to a final volume of 60 µL. The concentrations of the monopodal BBs **3a-g** are specified for each of the experiments.

General procedure for the fluorescence analysis of the DCLs

Once the oxidation of the free thiols was complete (after 24 hours at room temperature) the reaction mixtures were analyzed by fluorescence spectroscopy. For the DCLs performed at 0.05 mM of **3b**, the fluorescence samples were prepared by diluting the reaction mixture to a final volume of 2060 µL with a solution of 50% 100 mM BIS-TRIS aqueous buffer and 50% DMSO, reaching a final theorical concentration of 1.5 µM for **3b** (Figure S74). The mixtures containing a different concentration of **3b** were diluted to reach a final concentration of 1.5 µM . Therefore, for measurement purposes the samples were diluted 1:34 with H_2O/DMSO (1:1 (v:v)) prior to the fluorescence data acquisition.

After studying the effect of the DMSO in the sensing system we found that using a buffered solution at 50 mM with 50% DMSO to perform the measurements was optimal to avoid precipitation/aggregation processes and bleaching of the excimer signal.

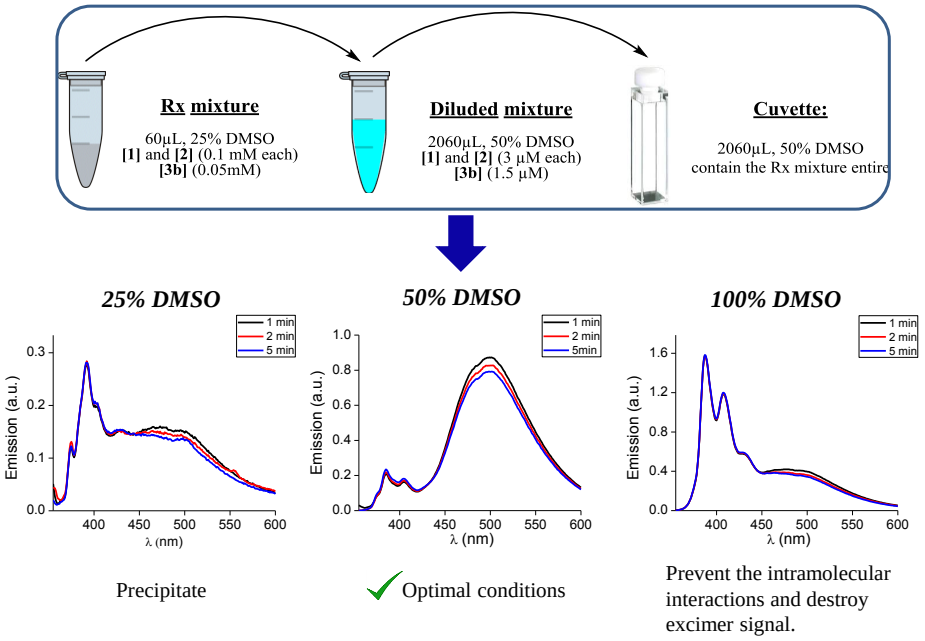


Figure S64. Preparation of the reaction mixtures (DCLs) and sample preparation for the fluorescence analysis.

Mixture of 1, 2 and 3b at different concentrations

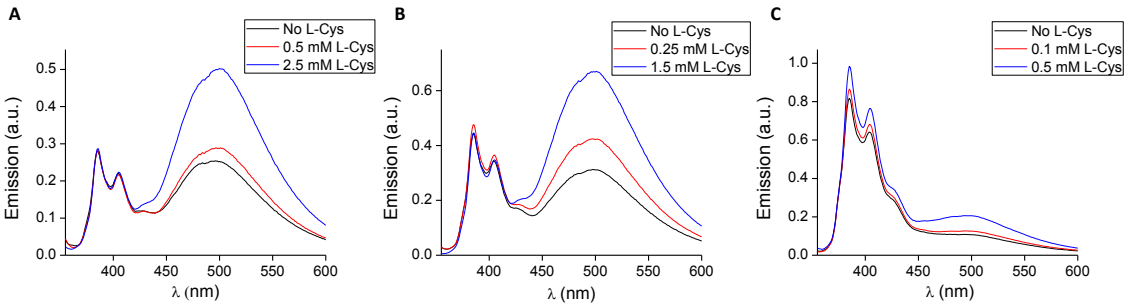


Figure S65: Emission spectra of a mixture containing **1+2** (0.5 mM each) at pH 6.5 with 25% of DMSO at different concentrations of **3b:** 0.5 mM (A), 0.25 mM (B), 0.1 mM (C); λexc = 345 nm.

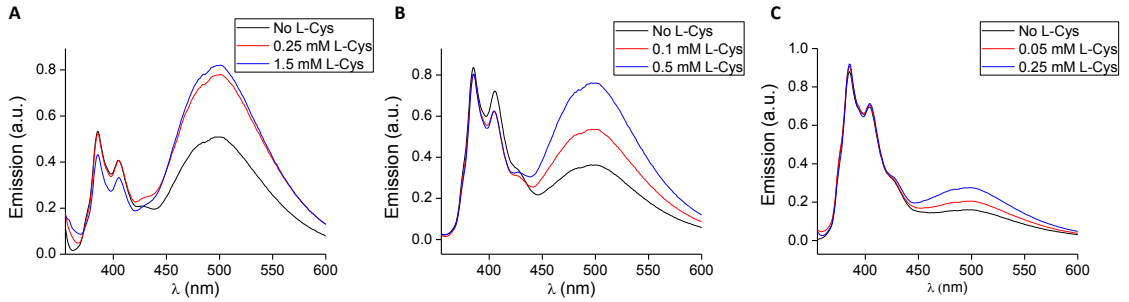


Figure S66: Emission spectra of a mixture containing **1+2** (0.25 mM each) at pH 6.5 with 25% of DMSO at different concentrations of **3b:** 0.25 mM (A), 0.1 mM (B), 0.05 mM (C); λexc = 345 nm.

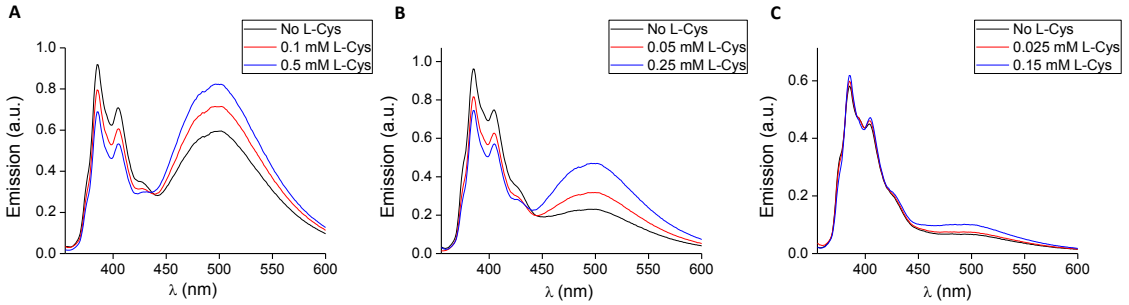
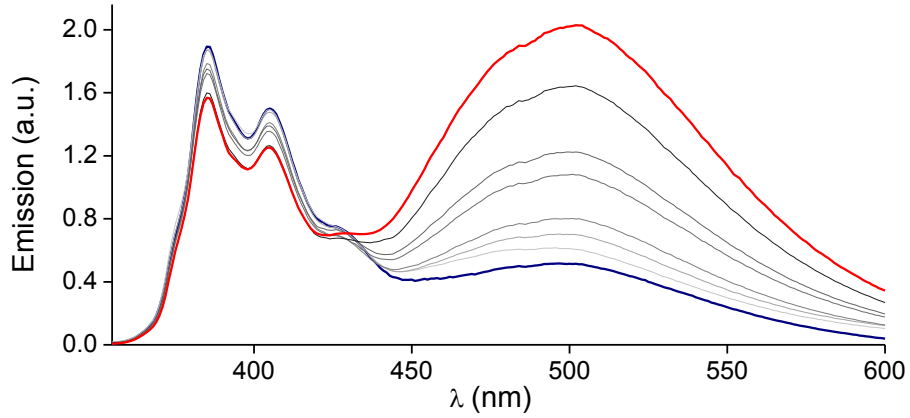


Figure S67: Emission spectra of a mixture containing **1+2** (0.1 mM each) at pH 6.5 with 25% of DMSO at different concentrations of **3b:** 0.1 mM (A), 0.05 mM (B), 0.025 mM (C); λexc = 345 nm.

Mixture of 1, 2 and 3b with different concentration of L-Cysteine or Cystine

A



B

$[L\text{-}Cys]_{Rx}$/ mM	$[L\text{-}Cys]_{measure}$/ µM	I_{501}/I_{385}
0.0	0.0	0.27
0.025	0.73	0.32
0.05	1.5	0.37
0.1	2.9	0.45
0.25	7.3	0.63
0.5	14.6	0.70
1	29.1	1.03
2.5	72.8	1.29

Figure S68: (A) Emission spectra of a mixture containing **1+2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO with increasing concentrations of L-Cysteine (blue to red)**:** No L-Cys, 0.025 mM, 0.05 mM, 0.1 mM, 0.25 mM, 0.5 mM, 1 mM, 2.5 mM); λexc = 345 nm. (B) Table of the reaction and measurement concentrations of L-Cysteine and the excimer/monomer emission ratio signal.

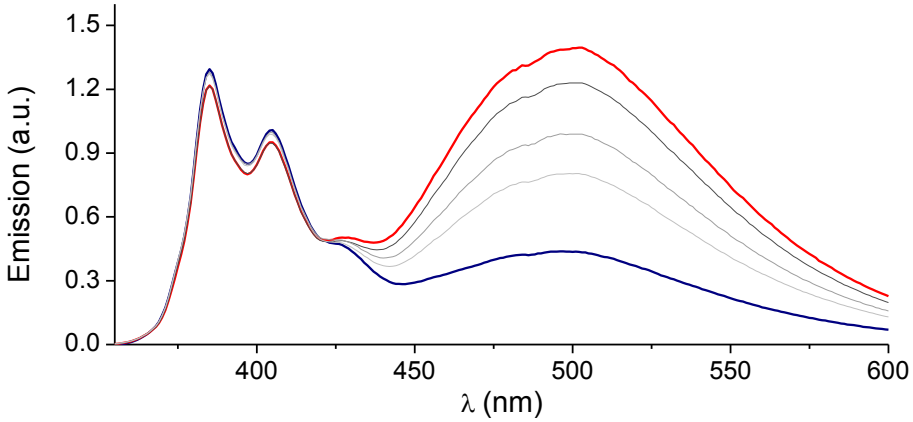


Figure S69: Emission spectra of a mixture containing **1+2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO with increasing concentrations of L-Cystine (blue to red)**:** No L-cystine, 0.05 mM, 0.1 mM, 0.25 mM, 0.5 mM; λexc = 345 nm.

Mixture of 1, 2 and 3b with different biothiols

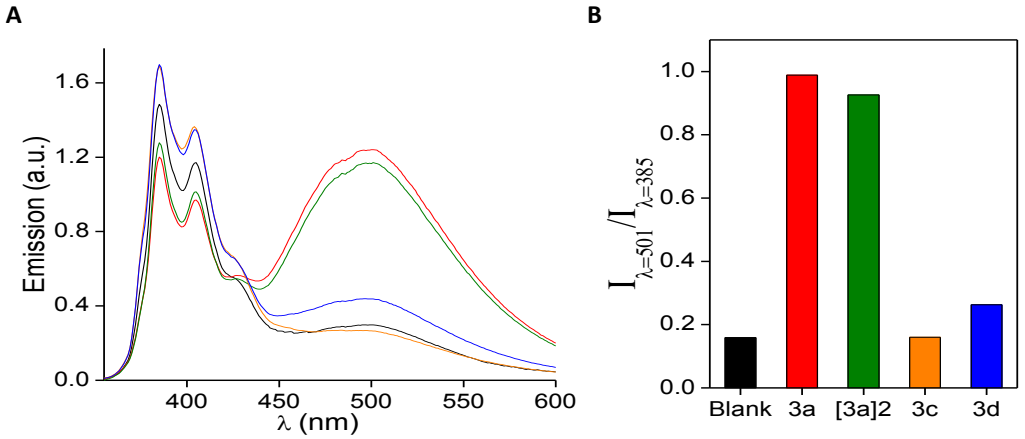


Figure S70: (A) Fluorescence emission spectra of the mixture of **1**+**2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO alone (black) and in the presence of 0.5 mM Cys (red), 0.25 mM cystine (green), 0.5 mM N-acety-Cys (orange) and 0.5 mM GSH (blue); λexc = 345 nm. (B) Excimer/monomer emission ratio plot.

Mixture of 1, 2 and 3b in different competitive media

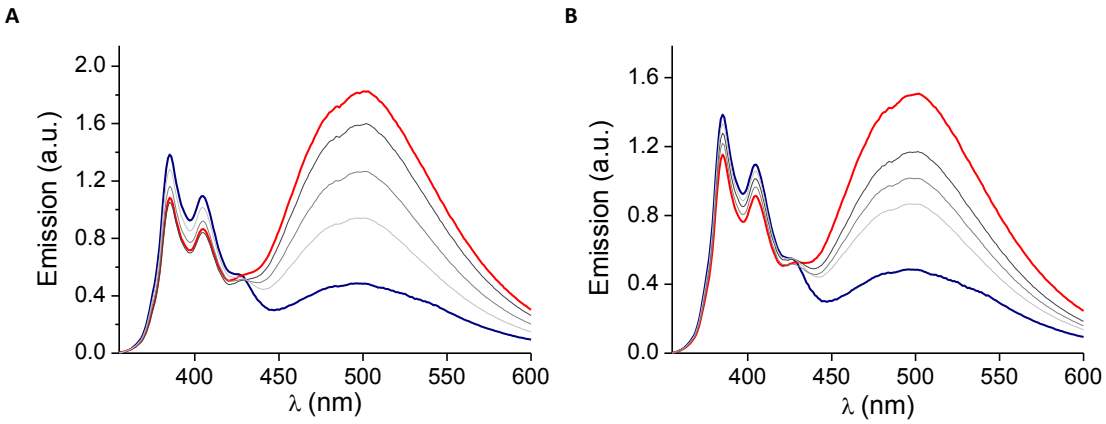


Figure S71: Fluorescence emission spectra of the mixture of **1**+**2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO and the sensing of (A) L-Cysteine or (B) L-Cystine (0.05 – 0.5 mM) in the presence of the three basic amino acids (2.5 mM Lys, 0.7 mM Orn, and 0.6 mM Arg) at the non-pathological upper limit concentrations found in urine.

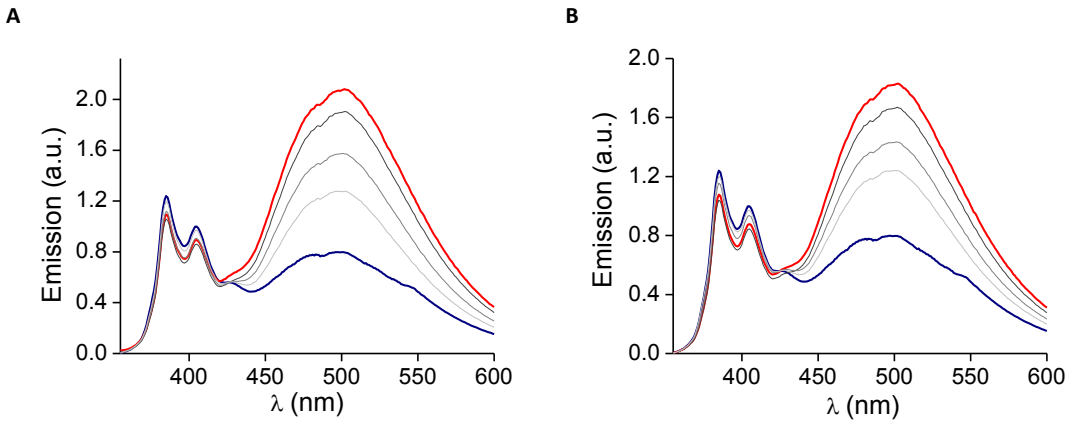


Figure S72: Fluorescence emission spectra of the mixture of **1**+**2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO and the sensing of (A) L-Cysteine or (B) L-Cystine (0.05 – 0.5 mM) in the presence of the three basic amino acids (5 mM Lys, 1.4 mM Orn, and 1.2 mM Arg). These are considered pathological concentrations in urine samples.

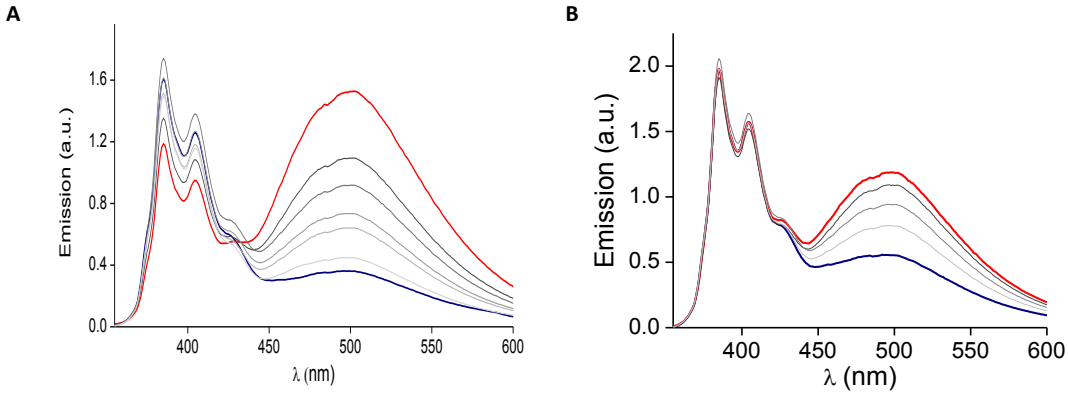


Figure S73: Fluorescence emission spectra of the mixture of **1**+**2** (0.1 mM each) + **3b** (0.05 μM) at pH 6.5 with 25% of DMSO and the sensing of (A) L-Cysteine or (B) L-Cystine (0.05 – 0.5 mM) in a buffer containing (8 mM Asp, 9 mM β-Ala, 6 mM Ser, 0.75 mM Tyr, 0.3 mM Met), these concentrations are typically found in urine samples of healthy people.

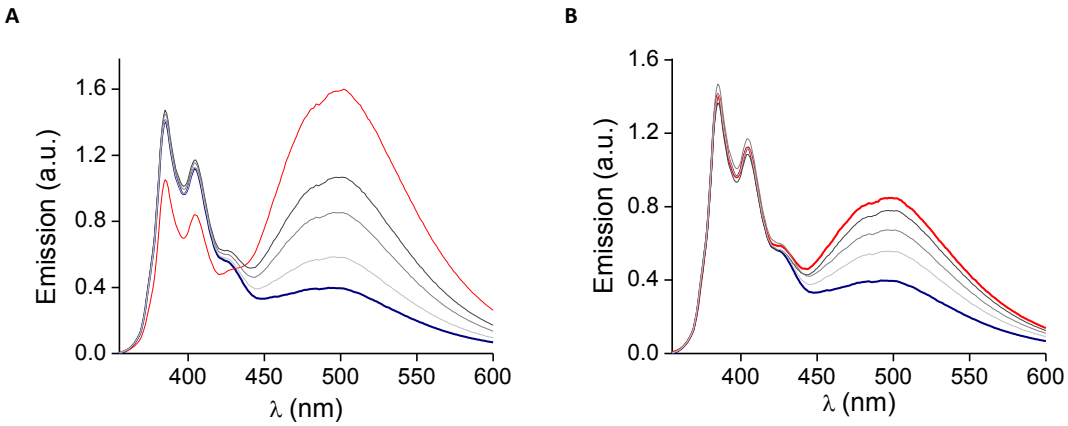


Figure S74. Fluorescence emission spectra of the mixture of **1**+**2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO and the sensing of (A) L-Cysteine or (B) L-Cystine (0.05 – 0.5 mM) in a buffer containing (0.6 mM Arg, 2.5 mM Lys, 0.7 mM Orn, 8 mM Asp; 9 mM Beta-Ala, 6 mM Ser, 0.75 mM Tyr, 0.3 mM Met), these concentrations are typically found in urine samples of healthy people.

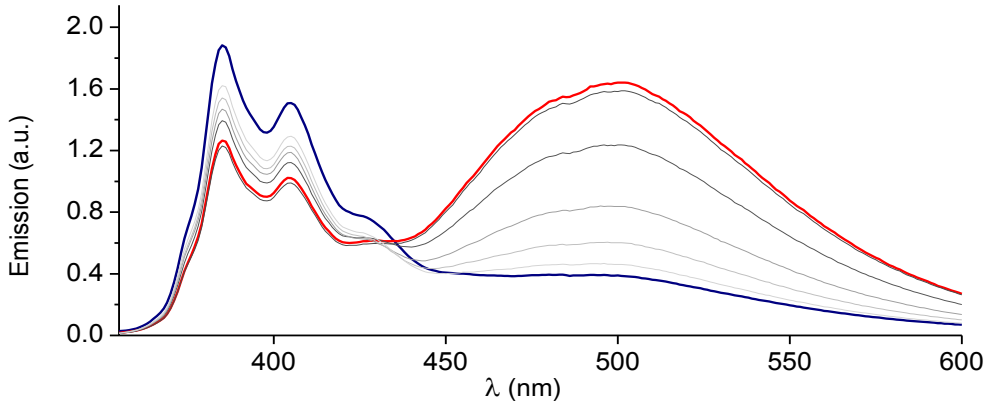


Figure S75. Fluorescence emission spectra of the mixture of **1**+**2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO and the sensing of L-Cysteine (0.05 – 0.5 mM) in a buffer containing (0.6 mM Arg, 2.5 mM Lys, 0.7 mM Orn, 8 mM Asp, 9 mM β-Ala, 6 mM Ser, 0.75 mM Tyr, 0.3 mM Met, 0.1 mM Asn, 5 mM Ala, 20 mM Gly, 10 mM His) these concentrations are typically found in urine samples of healthy people.

Mixture of 1, 2 and 3b in urine sample sensing L-Cys or Cystine

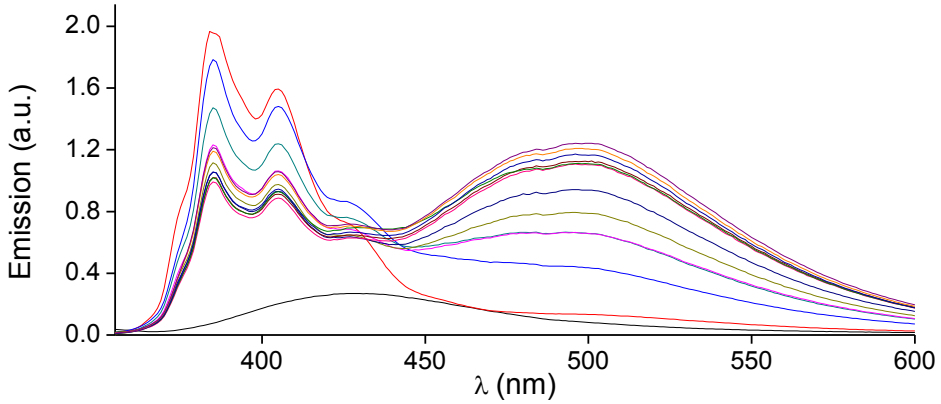


Figure S76. Fluorescence emission spectra of the mixture of **1+2** (0.1 mM each) + **3b** (0.05 μM) at pH 6.5 with 25% of DMSO and the sensing of L-Cysteine in a real sample of urine. Selected normalized fluorescence spectra of urine (black), sensor alone (red), sensor + urine (light blue) and after increasing concentrations of added Cys (0.25-2.5 mM).

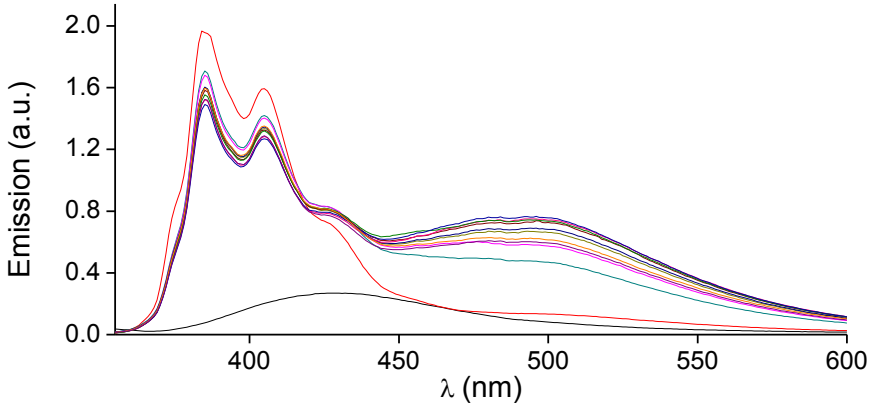


Figure S77. Fluorescence emission spectra of the mixture of **1+2** (0.1 mM each) + **3b** (0.05 mM) at pH 6.5 with 25% of DMSO and the sensing of Cystine in a real sample of urine. Selected normalized fluorescence spectra of urine (black), sensor alone (red), sensor + urine (light blue) and after increasing concentrations of added Cystine (0.25-2.5 mM).

Mixture of 1, 2 and 3b in different urine samples of healthy volunteers

Eleven samples of different volunteers were tested, and three extra measures of the combination of two samples for avoid cross response were tested too (Table S1). The fluorescence spectra of the urine without added sensor were measured to confirm that no other metabolites in this fluid could interfere with the analysis. The positive response of the sensor to the naturally excreted cysteine was also measured.

Sample's Number	Volunteer Age	Gender	Observations
1	29	F	-
2	24	F	-
3	34	F	-
4	38	F	-
5	29	F	-
6	30	M	-
7	23	M	-
8	62	F	-
9	65	M	-
10	34	F	Pregnant
11	27	F	Antecedents of kidney malfunction
12	29+30	F+M	Mixture
13	24+23	F+M	Mixture
14	62+65	F+M	Mixture

Table S1: Selected real samples of urine from healthy volunteers used for test the sensing system described.

General procedure for the preparation of the DCLs in human urine

For the experiments performed in human urine samples 300 mM BIS-TRIS methane buffer was prepared by dissolving 6.28 g of the free amine in 100 mL of milli-Q water and adjusting the pH of the solution to 6.5 by the addition of HCl (aq).

Individual stocks of the building blocks (BBs) **1**, **2** and **3b** in DMSO were prepared. Then, stock mixture solutions of the BBs were prepared from the individual stocks ensuring no differences in concentration between the reaction mixtures of the same batch. The DCLs were prepared in the reaction mixtures at final concentrations of 0.1 mM for the tri- and bipodal BBs (**1** and **2**) and 0.05mM of **3b** in a final 50 mM BIS-TRIS methane aqueous buffer urine solution (pH 6.5) with 25% DMSO. The samples were prepared by adding 15 μL of the stock mixture to 45 μL of the 300 mM buffered urine solution (containing 35 μL of urine sample and 10 μL of 300 mM buffered aqueous solution), to a final volume of 60 μL. The concentrations of the monopodal BBs **3a** are specified for each of the experiments.

General procedure for the fluorescence analysis of the DCLs

Once the oxidation of the free thiols was complete (after 24 hours at room temperature) the reaction mixtures were analyzed by fluorescence spectroscopy. The fluorescence samples were prepared by diluting the reaction mixture to a final volume of 2060 µL with a solution of 50% 100 mM BIS-TRIS aqueous buffer and 50% DMSO (Figure S74), reaching a final theorical concentration of 1.5 µM for **3b.** Therefore, for measurement purposes the samples were diluted 1:34 with 1:1 H_2O/DMSO (v:v) prior to the data acquisition.

Fluorescence spectra of urine samples of different volunteers

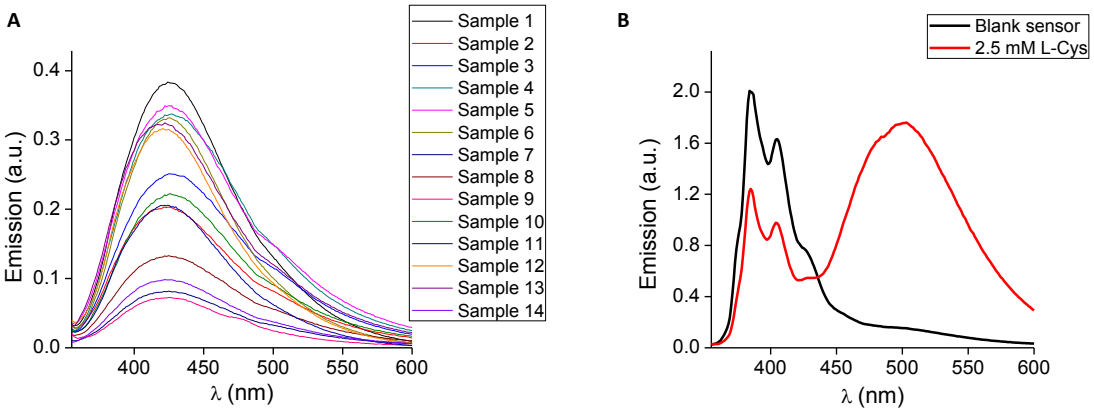


Figure S78. Control Experiments (A) The fluorescence spectra of the different urine samples without added sensor (B) Fluorescence spectra of the sensor without urine with and without 2.5 mM L-Cysteine.

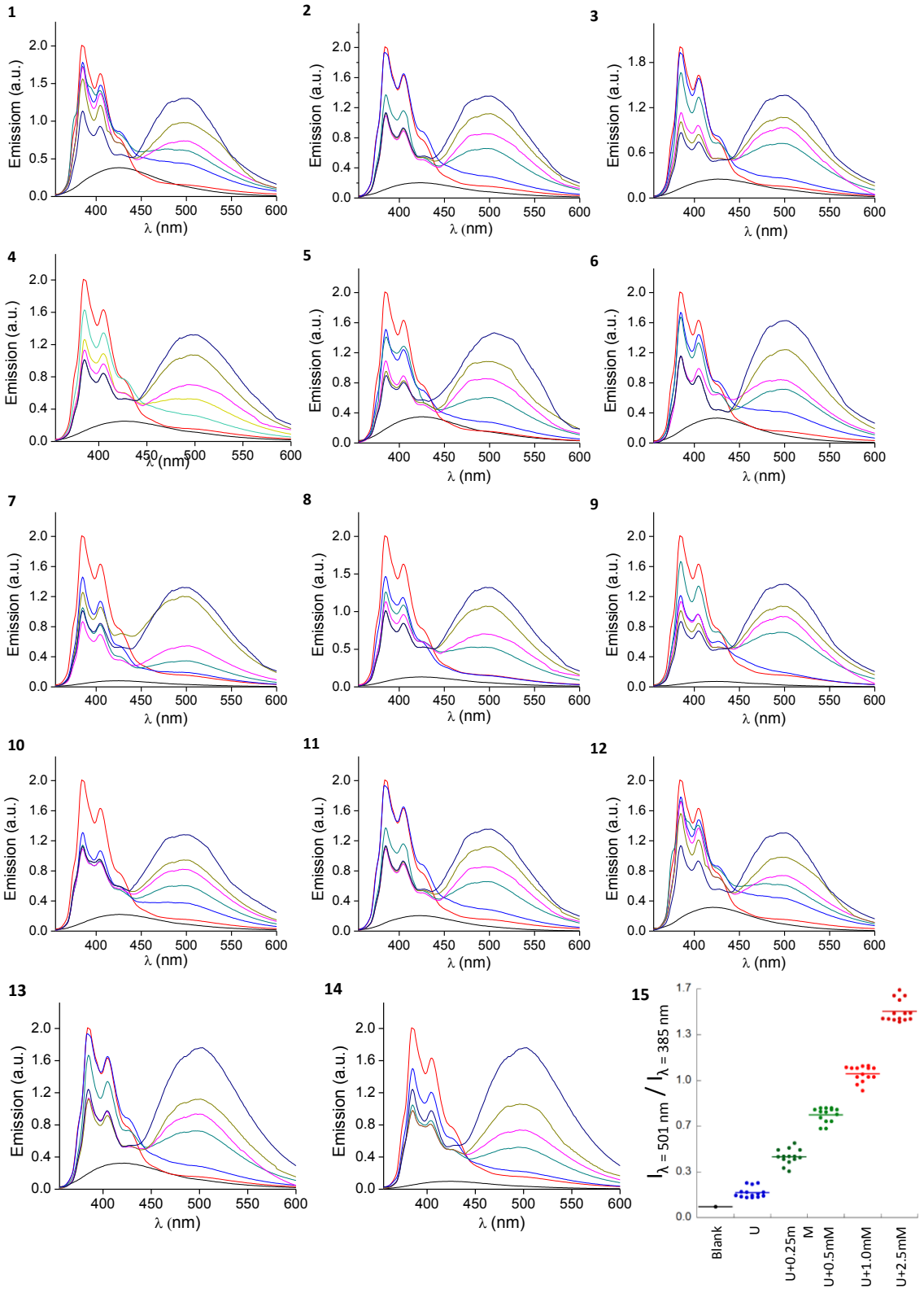


Figure S79. (1-14) Selected normalized fluorescence spectra of urine (black), sensor alone (red), sensor + urine (light blue) and after increasing concentrations of added Cys (0.25-2.5 mM) for the different samples of healthy volunteers. (15) Scatter plot of the excimer/monomer emission of the library alone (Blank), for different urine (U) samples of healthy volunteers and for the urine samples plus additional Cys.

REFERENCES

1. K. R. West, K. D. Bake and S. Otto, *Org. Lett.*, **2005**, *7*, 2615-2618.
2. J. Solà, M. Lafuente, J. Atcher and I. Alfonso, *Chem. Commun.*, **2014**, *50*, 4564.
3. A. P. Kozikowski, Y. Chen, A. Gaysin, B. Chen, M. A. D'Annibale, C. M. Suto and B. C. Langley, *J. Med. Chem.*, **2007**, *50*, 3054-3061.

CHAPTER IV

Discussion and concluding remarks

CHAPTER IV: Discussion and concluding remarks

4.1. General Discussion

In the final part of this Thesis, I will briefly discuss the main results obtained and their relevance in the field of DCC: from the study of the pseudopeptidic dynamic libraries (Chapter 2) to the possible practical applications of these dynamic libraries (Chapter 3).

In the first part of this work we focus on deeping in the study of the behavior of dynamic libraries formed by pseudopeptidic structures through disulfide chemistry, a relatively new research line in our group previously explored by Dr. Atcher. As stated in the general introduction the disulfide-based chemistry is an excellent tool for the formation of dynamic libraries and is one of the most used linkers in DCC. However, most of the systems described are based on molecules with two reacting sites (bipodals). In order to go a step further and to expand the structural and topological diversity of supramolecular assemblies by DCC, we decided to study the effect of combining BB with different valence on the same DCL. The novelty of the concept lies in the preparation of different chemical structures (monopodals, bipodals, tripodals), which can potentially lead to a large structural diversity (linear, cyclic, polycyclic) by suitable combination of the corresponding pseudopeptidic BBs in dynamic processes. Thus, by combining a series of BB with different valence, in principle a broad structural variety is expected. Nevertheless, it was observed that a self-recognition event within the BBs during the assembly process lead the equilibrium to the formation of a major, more stable compound. This allowed us to evaluate the effect of the molecular topology/geometry in the dynamic systems. We found that with the appropriate cooperative supramolecular interactions, BBs with different valence could result in the self-recognition and selective amplification of a singular constitution even when it is statistically disfavored. In figure 4.1a, we can observe that the combination of tripodal **1a** and bipodal **2a** BBs results into complex mixtures of compounds but, after the addition of Cys **3a** to the mixture, a single product containing one of each BB with a single topology is formed **[1a-2a-3a]**. This remarkable assembly is dictated by the cooperative action of weak polar interactions leading to a folded structure, as confirmed by NMR analysis (figure 4.1b). The most stable conformer shows a structure where the pendant Cys folds over the macrocyclic cavity. The ammonium cation is stabilized by the three carboxylates that arise from the mesic acid moiety. The carboxylate group of the Cys residue would as well be stabilized by hydrogen bonding interactions with phenylenediamine amide protons. More importantly, ROESYs experiment and the environmental and structural effects explained below, support the proximity between the pendant Cys and the bipodal moiety, so the presence of a folded structure in solution. All these interactions can only take place in the folded structure. We show the importance of the synergic action of subtle recognition events within the components of a DCL. Those findings make us suppose that the generation of this molecular diversity will open the possibility of creating highly complex and efficient molecular recognition systems, starting from structurally simple BBs. This hypothesis was corroborated in Chapter 3 and discussed below.

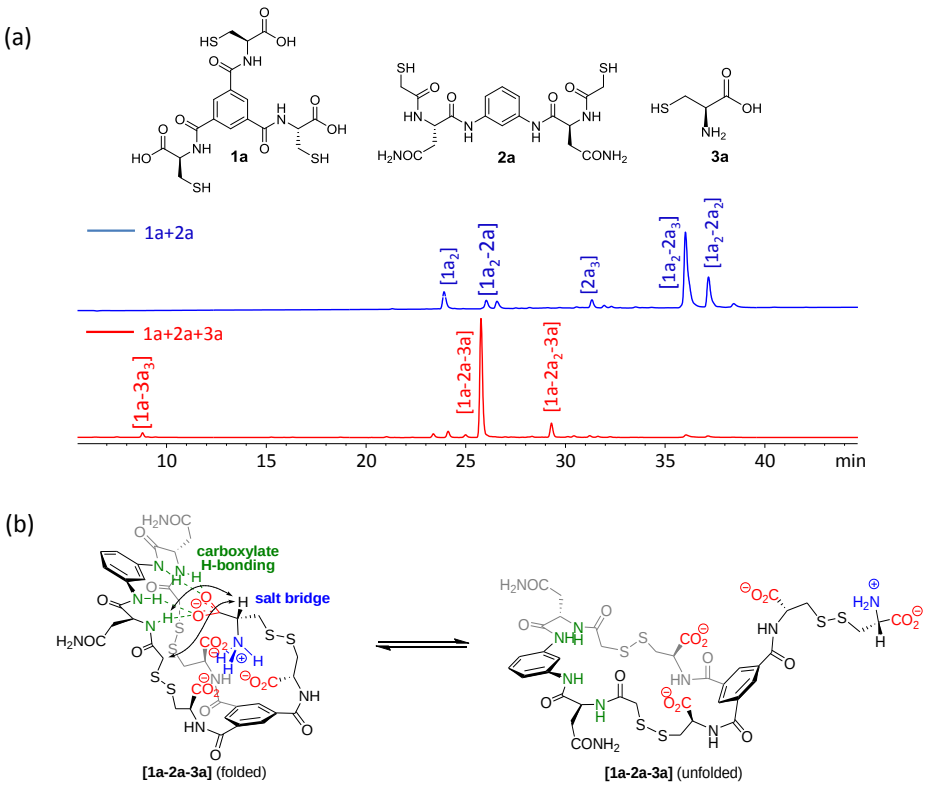


Figure 4.1. (a) HPLC traces of DLCs formed by **1a** and **2a** (0.5 mM each) in aqueous solution (25% DMSO, pH 6.5) in the presence (down) or absence (top) of Cys (2.5 mM). (b) The most stable conformer of **[1a-2a-3a]** shows a folded structure confirmed by NMR experiments and also supported by control experiments in different environments and with structurally different BBs.

After investigate the formation of libraries that arise by combining BBs with different valence, a series of DCLs formed by the combination of a monopodal, bipodal and tripodal BBs were studied. We extend our investigations to discover the variables that drive the system to its final composition. We have set up complex disulfide-based dynamic covalent libraries of chemically and topologically diverse pseudopeptidic compounds. Firstly, we monitored the formation of **[1a-2a-3a]** over the time and was observed that the reaction evolves from very complex mixtures at short reaction times to the almost exclusive formation of a major compound, through the establishment of intramolecular noncovalent interactions (figure 4.2). Our experiments demonstrate that the systems evolve through error-check and error-correction processes. The nature of these interactions, the importance of the folding and the effects of the environment are also discussed below.

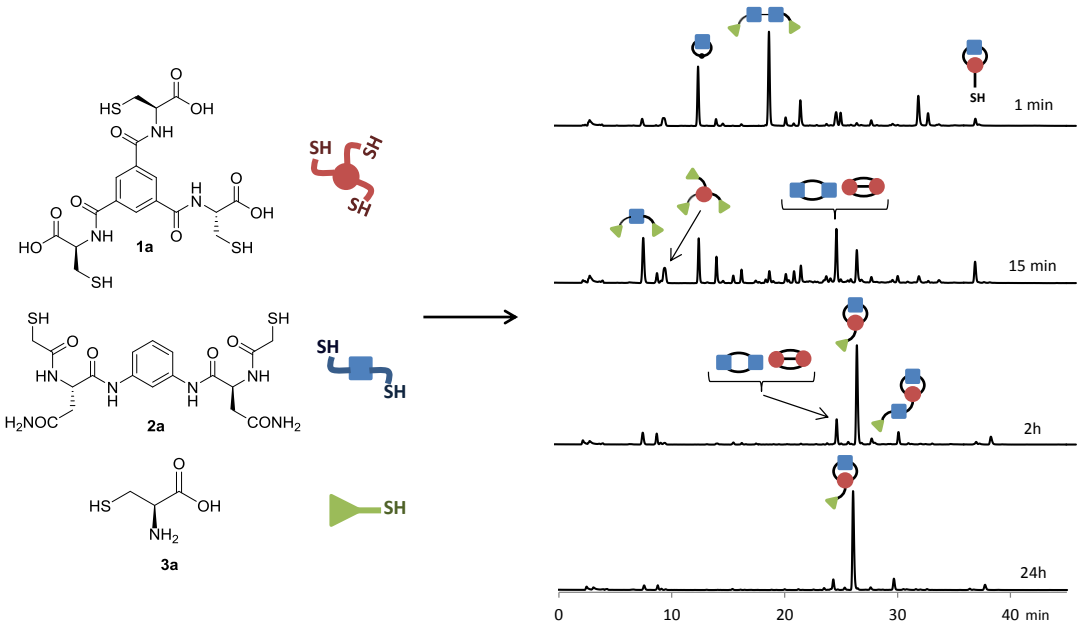


Figure 4.2. HPLC traces (UV detection at 254nm) for the evolution of a sample containing **1a, 2a** (0.5mM each) and **3a** (2.5mM) over time (pH 6.5, 25% aqueous DMSO).

After suppose that the main product formed is highly amplified from the dynamic mixture because of the attractive interactions that are established between the different moieties of the BBs during the assembly process, we decided to study the adaptive behavior of these dynamic combinatorial libraries to support this fact. The next step was to explore the mixture composition when we change the intrinsic and extrinsic factors for know more about the non-covalent intramolecular interactions that drive towards the formation of almost a single product.

We improve the comprehension of the behavior of complex dynamic libraries of compounds deeping in the adaptive properties of the pseudopeptidic dynamic libraries in response to different stimuli (effect of the pH, ionic strength or polarity of the solvent) and how these affect to the final distribution of the library products. The information stored in the different components of the library can be expressed through an adaptive process that determines the structurally fittest assembly under given environmental conditions (figure 4.3.). The importance of the pH was verified with repeating the DCL at different pH values, the optimal pH in which the recognition is effective is around 6.5. At this value the carboxylic acids are deprotonated and the amine group from the Cys moiety is protonated so the proposed salt-bridge can take place efficiently. As we expected the members of the libraries are affected by the amount of salt, therefore the ionic strength was found to be a crucial parameter. The higher ionic strength shielded the polar interactions and the selection is less-effective, this is a further indication of the polar nature of the interactions driving the formation of almost a single compound. After was studied and ascertained the effect of these stimuli, a methodology has been developed to carry out the future studies of our research group, which will be applied in all the new systems prepared as a control methodology. One of the advantages of our dynamic combinatorial approach is the possibility of screening a high number of structural possibilities with a minimum synthetic effort.

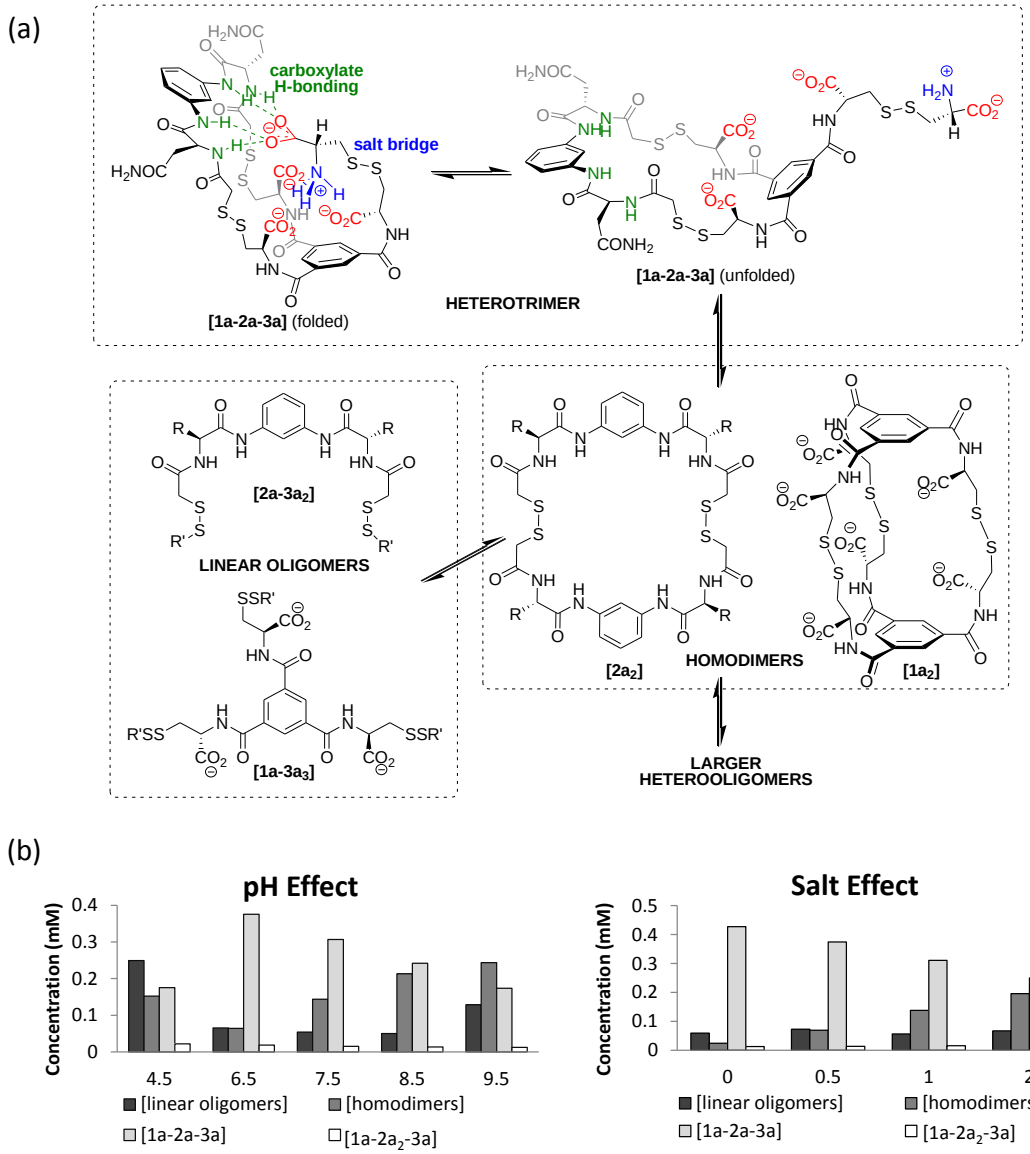


Figure 4.3. (a) General equilibrium pattern proposed for the DCLs obtained by mixing BBs with different topology. (b) Compound distribution of the DCLs of **1a**, **2a** (0.5mM), **3a** (2.5mM) at different pH values and at different salt (NaCl) concentration in aqueous buffer.

Subsequently was analyzed the structural effects arising from the scaffold of the BBs. It was expected that the suppression of the possible interactions or the establishment of new ones will clearly modify the relation of products in the library. Were studied the effects of different commercial monothiols, the influence of the presence of charged residues in the bipodal moiety and tripodals with a different geometrical disposition unity. These experiments were realized changing only one by one unity and all the changes confirmed the importance of having the correct interactions for the selection of a single geometry.

Following this research line, we focus on how structural parameters affect dynamic systems. We showed that within a DCL with BBs of different valence, the distribution of the products can follow two different behaviors of self-assembly. Introducing the appropriate BB a single species is selected from a virtual very complex mixture resulting in a very stable structure, due to the suitable non-covalent intramolecular interactions even when it is statistically disfavored. On the other hand if we varying one of the species of the same library

which may only differ in one methylene unit, results into the formation of wide diversity of macrocycles as expected. In our case, a perfect chemical and structural match between a tripodal BB **1** and a bipodal moiety **2c** results into an unexpected cage-like compound formation that emerges selectively from a potential complex mixture as a single product (figure 4.4b). Non-surprisingly, the pH and the ionic strength play a key role again in the nature of the species formed. Also, NMR, MALDI-TOF MS-MS and molecular mechanics calculations confirmed the formation of a cage-like structure instead of the other possible constitutional isomer. Additionally to this data the cage-like formation was monitored over time, at short reaction times the formation of **[1-2c]** was observed, but remarkably, it took longer than other libraries previously reported in our group to reach a thermodynamic equilibrium (i.e. formation of **[1a-2a-3a]**). This means that the formation of the main species is fairly complex and again, this fact supports the formation of **Ib** structure in detriment of the other constitutional isomer, **Ia**, which would be formed rather quickly.

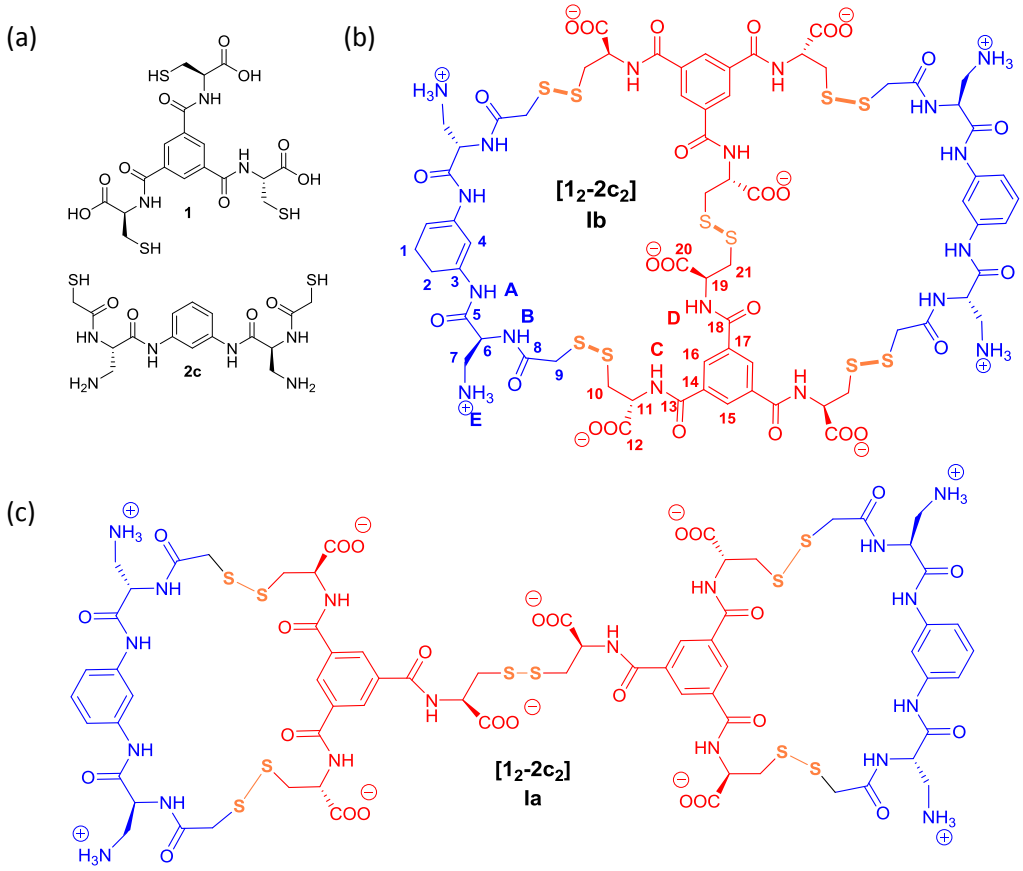


Figure 4.4. (a) Chemical structures of the different BBs. (b) Structure for compound **[1_2-$2c_2$]**, **Ib.** (c) Structure for compound **[1_2-$2c_2$]**, **Ia**.

Finally, we decided to investigate the structural elements that are crucial for the selection of the cage-like compound **[1_2-$2c_2$]** in order to get more insights on the mechanism of self-selection. With that idea in mind a competition experiment was designed and two homologous amino bipodal compounds of **2c** with different side chain length (2 or 3 methylene units, corresponding to **2d** and **2e**) were prepared and tested (figure 4.5a). Tripodal **1** (0.5 mM) was mixed with either compounds **2d** or **2e** (0.5 mM) in the usual conditions and no predominance towards the corresponding cage-like structure was observed for these homologous compounds. Surprisingly when **1** (0.5 mM) was reacted with a mixture of **2c, 2d**, and **2e** (0.5

mM each) in a competition experiment under the usual conditions (aqueous solution at pH 6.5, 25% DMSO) compound **[1_2-$2c_2$]** was formed in a way that consumed all **1** and **2c** BBs (Figure 4.5a). Therefore, the chain length plays a decisive role for the stabilization of a major compound through supramolecular interactions and homologous compounds differing as much as one methylene unit behave in a divergent manner. The delicate structural complementarity was further shown by an analogous experiment using bipodal compound **3** instead of the tripodal unit **1**. When an equimolar mixture containing the homologous amino components **2c, 2d** and **2e** and the bipodal carboxylate **3** was reacted no selectivity towards any of the amino derivatives was observed. Accordingly, compound **3** was combined with the different positively charged bipodal units, obtaining as a result a wide range of products in the final distribution of the compounds (figure 4.5b). Once again these results indicate the delicate structural complementarity necessary for the emergence of self-recognition processes. In particular, this experiment demonstrates the need of the third arm to form selectively **[1_2-$2c_2$]** and it is a further evidence for the **Ib** isomer as it forms a bigger macrocycle which brings the opposite charges in close proximity. This disposition would explain the dissimilar behavior for the other bipodal units where a single chain extension prevents establishing the adequate stabilizing interactions.

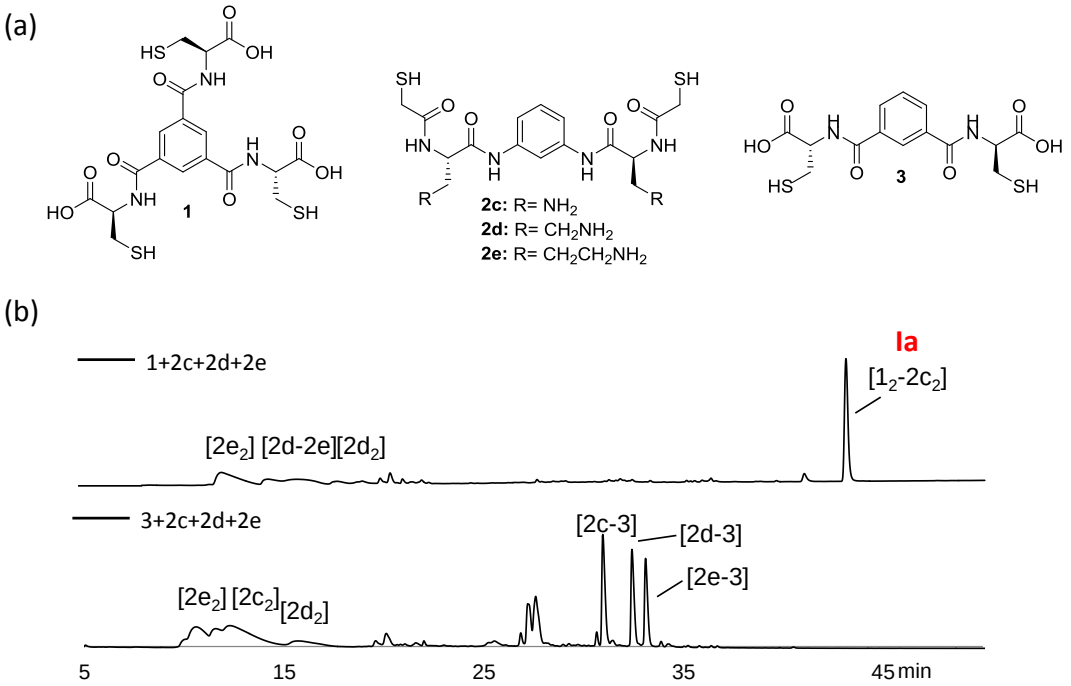


Figure 4.5. (a) Chemical structures of the different BBs. (b) Top: DCL formed by **1** and **2c, d, e** (0.5mM each) in aqueous solution (25% DMSO, pH 6.5), Bottom: DCL formed by **3** and **2c,d,e** (0.5mM each) in aqueous solution (25% DMSO, pH 6.5).

The structural importance of these interactions reminds the folding of proteins where the system is reorganized by acquiring a specific structure that is maintained by covalent bonds and other non-covalent interactions that by themselves can be weak but in a combined way they manage to maintain a compacted and stable structure. The system can be affected by external factors and being able to direct the DCL towards a desired final distribution too. Moreover, during the study of the behavior of dynamic libraries, we have discovered that these can be self-organized into complex structures selectively and according to well-defined

structural parameters. This has led to the study of these structural parameters by different techniques, some of them already explained previously as HPLC-MS or NMR.

Tom sum up, the systematic study of the intrinsic (chemical structure and valence of the BB) and extrinsic (pH, ionic strength and organic co-solvent) factors rendered important information about noncovalent intramolecular interactions. If we understand the behavior of the interactions that are responsible of the formation of different products and learn what factors are leading to their formation or not, we will decode the key to direct the thermodynamic equilibrium towards a desired formation of the products. Systematically planned variations of the BBs structures and of the experimental conditions can provide a basis for a better understanding of supramolecular associations in natural and synthetic systems and for the theoretical prediction of noncovalent interactions and result into an excellent tool for predict the formation of compounds and for contribute to the understanding of biologically important associations.

Finally, after the knowledge acquired and the results obtained in the recognition of molecules in aqueous media through DCC, we consider modify the stable product **[1a-2a-3a]** self-selected in previous DCLs to perform dynamic systems with a practical application. Specifically, we focus on the synthesis and study of specific chemical sensors of amino acids in aqueous media. The aim was to achieve for the first time the preparation of a dynamic system capable of producing a fluorescent response in presence of Cys or cystine in biological fluids. Different investigations have already been carried out on the detection of biothiols by fluorescence spectroscopy using small molecule probes. Nevertheless, many of these sensors have a relatively low sensitivity, they need long response times and a complicated synthesis process. In addition, given that these biothiols have a similar structure and reactivity, but are associated with different diseases, the design of a highly selective detection system that differentiates them remains an important challenge in the clinical diagnosis. Moreover, most of these sensors require the presence of the free thiol, precluding the detection of cystine.

As we explained in the previous section, Cys is associated with several diseases, such as cystinuria, an illness characterized by the presence of kidney stones produced by a deficient reabsorption and subsequent precipitation of the urine's cystine at neutral pH. Our goal was the development of a molecular network able to give a selective and sensitive fluorescence response for the diagnosis of cystinuria.

Our target was to modify the structure of the heterotrimer **[1a-2a-3a]** to incorporate a fluorescent moiety. The first step was to replace the Cys unit (**3a**) with a monothiol component containing a chromophore. This new fragment should also contain a hydrogen bond donor group that allows an interaction with the central macrocycle of the molecule but in a weaker way than the one given in the case of Cys. Once the new heterotrimer containing a chromophore is formed the addition of Cys releases the chromophore as its disulfide homodimer, producing a response by fluorescence. This conversion could be read in the fluorescence spectra due to the formation of excimers (figure 4.6.).

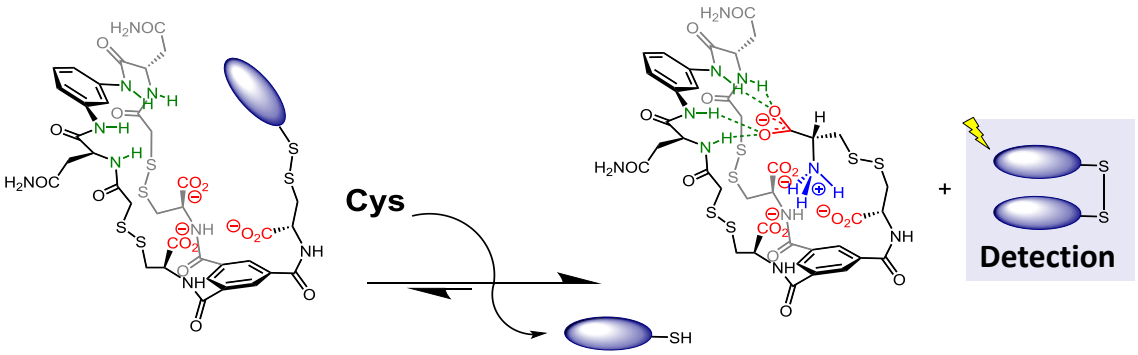


Figure 4.6. Proposed mechanism of action for a Cys sensor through the fluorescence response.

We synthesized 5 candidates with suitable fluorescent aromatic groups to replace the monopodal component (figure 4.7). It is worth mentioning that solubility presented one of the main problems as these aromatic compounds are poorly soluble in aqueous media. After analysing the fluorescence and solubility of each candidate we selected compound **3b** as the most promising component for the development of the sensor, **3b** has a suitable fluorescence spectra of its reduced and oxidized forms: **$3b_2$** presents an excimer emission band at about 500 nm that is absent in the monomer, which is characterized by a low-wavelength band with fine structure (c.a. 350-450 nm).

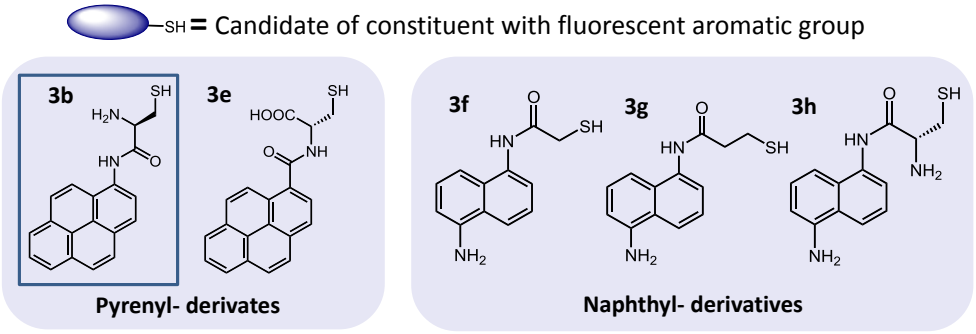


Figure 4.7. The 5 candidates synthetized and analyzed to substitute the monopodal unit.

We performed a set of experiments varying the concentration of the three components (tripodal **1a**, bipodal **2a**, monopodal-chromophore **3b**) with and without Cys to optimize the conditions in which our sensor was more selective. Finally, the optimal conditions to minimize the formation of **[$3b_2$]** homodimer, with almost no excimer emission are the ones which components **1a** and **2a** have an equimolar concentration (0.1mM each) and 0.05 equivalents of the monopodal (figure 4.8a, blue trace). As we expected the addition of Cys restructured the dynamic library. The fluorophore is released (due to the formation of the more stable **[1a-2a-3a]** species), and the homodimer of the fluorescent monopodal is formed and generate a readable response, the emission of the excimer was detected at 501 nm, so when Cys was added to the reaction the excimer band increased due to the formation of **[$3b_2$]**. The system shows a significant turn-on fluorescence response upon addition of Cys concentrations upper to 50 μM demonstrating the efficiency of the sensor. Since the normal presence of Cys in urine is ~35 μM and the occurrence of stone-producing cystinuria starts from approximately 0.8 mM, our method fits within these values.

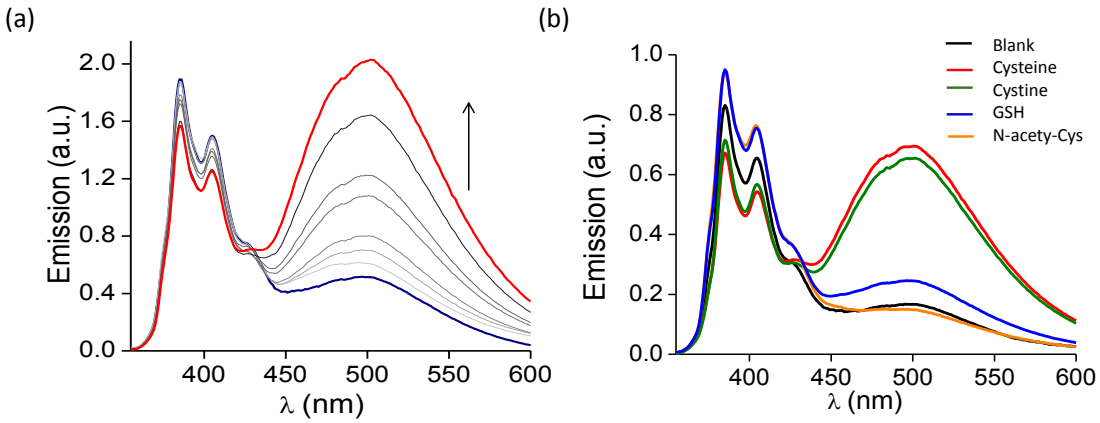


Figure 4.8. (a) Fluorescence emission spectra of the dynamic library in aqueous media and under increasing concentrations of L-Cys (0.025-2.5 mM). (b) Fluorescence emission spectra of the library alone (black) and in the presence of L-Cys (red), cystine (green), NAC (orange) and GSH (blue).

The selectivity against different biothiols was tested and the system gives almost no response to the presence of other biologically relevant Cys derivatives like GSH (figure 4.8b, blue) or N-acetylcysteine (NAC) (figure 4.8b, orange). Additionally, we check the individual and crossed responses, the interference from the presence of other amino acids that can be found in urine with the sensor. The dynamic sensor is able to detect **3a** and [**3a**$_2$] in the presence of the basic amino acids (Lys, Orn, Arg) at upper limit and pathological concentrations in regular urine samples. This highlights the robustness of the network sensor, which is able to detect the analytes in the presence of high concentrations of basic amino acids. The sensor also responded to concentrations of both Cys and cystine in very competitive media containing most of the amino acids that can be found in urine (Lys, Orn, Arg, Asn, Gly, Ala, His, Asp, β-Ala, Ser, Tyr, Met) at their normal concentrations.

Once the sensitivity and the selectivity of the method were evaluated, the sensing system was tested in real urine samples from healthy volunteers. The fluorescence spectrum of urine was measured without added sensor to confirm that no other metabolites in this fluid could interfere with the analysis (figure 4.9a, black spectrum). The positive response of the sensor to the naturally excreted Cys in the urine samples was also measured (light blue spectrum in figure 4.9a and U in figure 4.9b). The system shows a significant response upon addition of Cys, demonstrating the efficiency of the sensor. Thus, we could also sense abnormal concentrations of Cys that would not yet cause calculi (green).

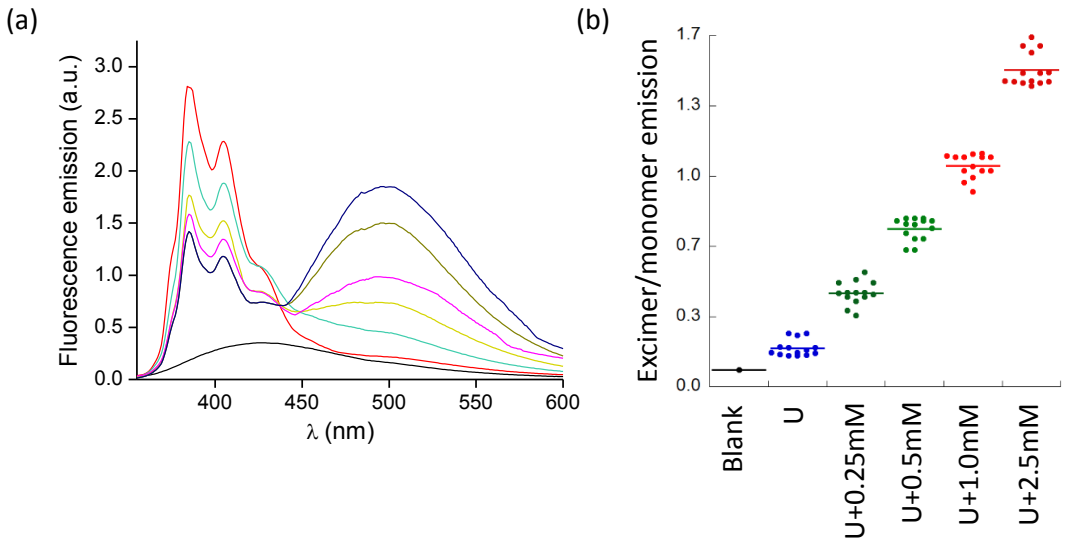
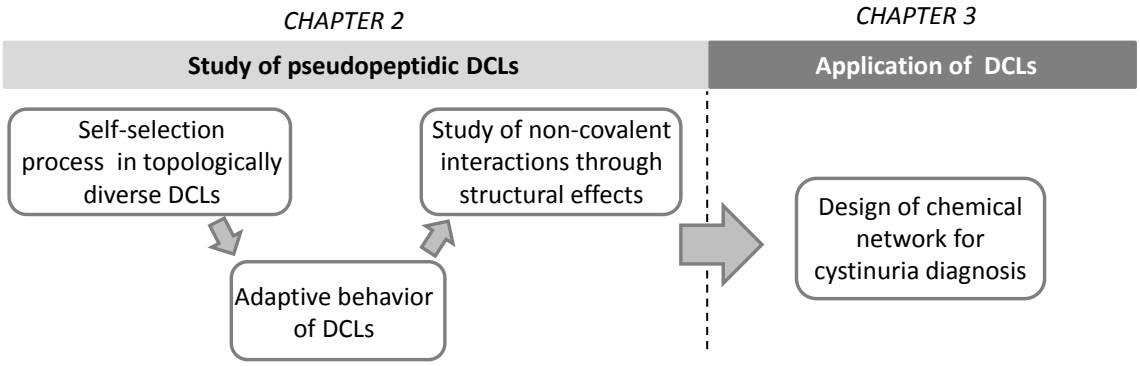


Figure 4.9. (a) Selected normalized fluorescence spectra of urine (black), sensor alone (red), sensor + urine (light blue) and after increasing concentrations of added Cys (0.25-2.5 mM). (b) Scatter plot of the excimer/monomer emission of the library alone (Blank), for different urine (U) samples of healthy volunteers and for the urine samples plus additional Cys.

In conclusion, we described a dynamic molecular network able to selectively and sensitivity sense biologically relevant molecules. This system works in aqueous media to selectively sense Cys and cystine against other biothiols, even more importantly in human urine. The dynamic nature of the sensing system based on disulfide formation and exchange, allows detecting Cys in its reduced or oxidized forms and does not show any interference with other amino acids that are in the medium. Therefore, an effective system has been created for the diagnosis of cystinuria disease.

Our work shows the potential of chemical networks to achieve advanced applications and we believe that this is an excellent result that represents a good starting point to develop other selective sensor in biological environments.

4.2. General Conclusions



In this Thesis I worked to expand the study of disulfide-based DCLs formed by pseudopeptidic structures and apply the knowledge acquired in possible applications of these dynamic libraries. I aimed to contribute to the bridge gap between the conceptual study of DCLs and the practical applications of these dynamic systems in real life. The detailed study of the sequence of topics listed above has led to the following general conclusions:

1) The possibilities of combinatorial chemistry have been expanded, mixing in the same library BB with different valence. The combination of different BB could result in a plethora of possible structures. However, cooperative supramolecular interactions between the BBs result in the self- recognition and selective amplification of a singular constitution even when it is statistically disfavored. Structural studies, including molecular modeling and NMR experiments, confirm a folded conformation for the main specie formed. The synergic action of subtle recognition events within the components of a DCL has a major effect on the product distribution of the DCL. Molecular interactions play a key role in the self-assembly processes in Dynamic Covalent Libraries.
2) A widely study of the adaptive behavior of pseudopeptidic libraries has been developed. The systematic study of their time-evolution and equilibrium composition depending on intrinsic (chemical structure and valence of the BBs) and extrinsic (pH, ionic strength and organic co-solvent) factors rendered important information about non-covalent intramolecular interactions. Our experiments demonstrate that the systems evolve through error check and error correction processes. Also, the suitable experimental conditions for the generation of useful disulfide-based DCLs have been successfully developed and established for future studies. A deep understanding of the interactions that govern the product distribution of DCLs will help to improve the design of new molecular receptors that operate in aqueous environments.
3) The intricate relationship in a potentially complex dynamic library that drives the equilibrium towards the formation of a particular species has been shown. Introducing the appropriate BB a single species is selected from a virtual very complex mixture resulting in a very stable cage-like structure, due to the suitable non-covalent intramolecular interactions even when it is statistically disfavored. These interactions are very sensitive to the size of the BBs being just an increment of a methylene unit detrimental for the efficient formation and sufficient to change the size of the macrocycles formed from 26 to 42 members; this size is only observed when the

negatively charged species has three side chains but it is not observed for the bipodal homologous compound which could, in principle, adopt the same cycle. The formation of the major compound is therefore, directed by the nature of all counterparts and the stability of the newly formed compound is higly dependent of subtle structural changes. The structural importance of intramolecular interactions in aqueous media is shown in many biological processes and knowing the rules that guide self-assembly will allow us to design more complex systems and predict their behavior opening the door to the synthesis of molecular devices with a tailored function.

4) A methodology based on extends pseudopeptides diversity in dynamic libraries has been proposed with the aim of detection molecules of biological interest. We focus on the design and study of specific chemical sensors of amino acids in aqueous media. As a proof of concept we develop a system able to give a selective response in the detection of the L-Cys, and its oxidized dimer, L-cystine, able to give a readable response through spectrophotometric measurements (fluorescence). This molecular framework acts as a selective fluorescent sensor of Cys, even in the presence of other amino acids and in biological fluids, such as human urine, with the consequent potential application for the diagnosis of cystinuria, a metabolic disease related to the transport of said amino acid.

Overall, the results comprehended in this thesis represent a contribution to expand the study of combinatorial dynamical libraries and in the consequent utilization of these for practical purposes. The combination of pseudopeptidic BBs in DCvC processes has allowed us to investigate a wide variety of structural diversity. The generation of this molecular diversity and the exhaustive study of its adaptive properties open the possibility of creating highly complex and efficient molecular recognition systems, from structurally simple BBs. In this regard, this knowledge has been applied to develop a molecular network capable of providing a selective and sensitive fluorescence response for the detection of Cys and cystine in urine samples.